CAPCO

Journal

THE CAPCO INSTITUTE JOURNAL OF FINANCIAL TRANSFORMATION

RISK

Available to download at CAPCO.COM/INSTITUTE

#43
05.2016

RECIPIENT OF THE APEX AWARD FOR PUBLICATION EXCELLENCE

EMPOWERING THE [FINANCIAL] WORLD

Pushing the pace of Financial Technology, together we'll help our clients solve technology challenges for their business – whether it's capital markets in Mumbai or community banking in Macon.

We leverage knowledge and insights from our clients around the world:

20,000 clients in towns everywhere are becoming more efficient, modern and scalable.

27 billion transactions processed help solve clients' challenges — big and small.

$9 trillion moved across the globe in a single year empowers our clients' communities to build storefronts, homes and careers.

55,000 hearts and minds have joined forces to bring you greater capabilities in even the smallest places.

Empowering the Financial World

FISGLOBAL.COM

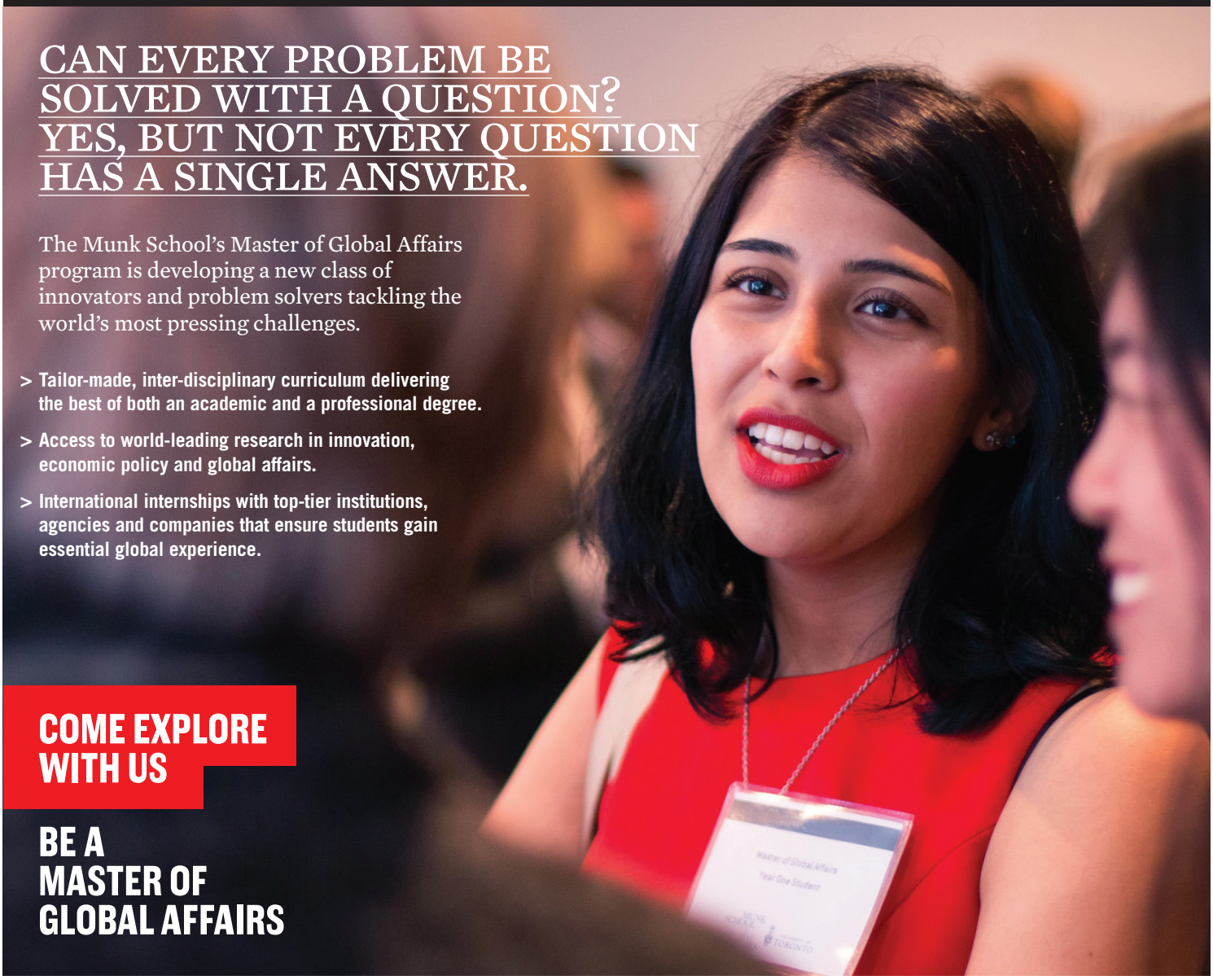
FIS

Journal

The Capco Institute Journal of Financial Transformation

Recipient of the Apex Award for Publication Excellence

WHAT ARE THE DRIVERS AND DISRUPTIONS THAT DETERMINE INNOVATION AND PROSPERITY?



CAN EVERY PROBLEM BE SOLVED WITH A QUESTION? YES, BUT NOT EVERY QUESTION HAS A SINGLE ANSWER.

The Munk School's Master of Global Affairs program is developing a new class of innovators and problem solvers tackling the world's most pressing challenges.

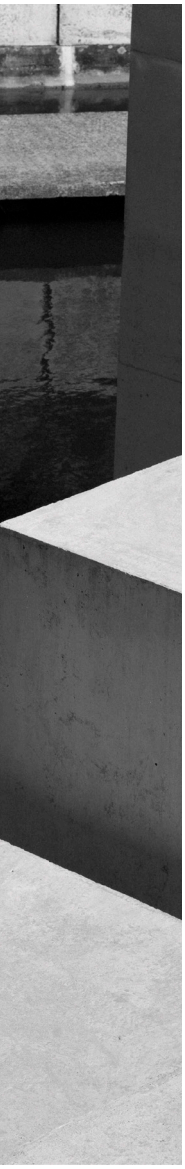
- > Tailor-made, inter-disciplinary curriculum delivering the best of both an academic and a professional degree.
- > Access to world-leading research in innovation, economic policy and global affairs.
- > International internships with top-tier institutions, agencies and companies that ensure students gain essential global experience.

COME EXPLORE WITH US

BE A MASTER OF GLOBAL AFFAIRS

MUNKSCHOOL.UTORONTO.CA
MGA@UTORONTO.CA

MUNK SCHOOL OF GLOBAL AFFAIRS
UNIVERSITY OF TORONTO



Risk

New Entrants

Regulatory

Investment

Dear Reader,

Welcome to issue 43 of the Capco Institute Journal of Financial Transformation.

The Capco Institute retains an unwavering focus on intellectual excellence to offer insight into a complex global environment. From thought leadership events to the Capco Institute Journal of Financial Transformation, our Institute brings together academics, practitioners and Capco experts who are able to provide their unique and informed perspective on the latest developments in financial services.

2016 is a significant year for those in financial services with major events set to have a significant global impact. The U.K. goes to the polls in June to decide whether or not to remain in the E.U., and the result of this referendum will have major implications for both the financial industry and the broader economy. The U.S. will elect a new president in November, which will shape the next four years or more of policy in America.

Against this political environment, global output is down, and the market is nervous about a sluggish Chinese economy. In Europe and Japan, negative interest rates are in place and there is speculation that central bankers are contemplating further measures to stimulate relatively stagnant economies.

The need for innovative thinking and attention to detail is as critical as ever, and this approach is pursued strongly in the latest edition of the Journal. The theme of this edition is risk, one of the most pertinent and fastest growing topics within the industry. Our contributors examine the risk of traditional institutions not responding proactively to the offerings or flexibility of new entrants. Further, they explore the risks pertaining to financial regulations and investment risk.

Finally, I'd like to welcome back our editor, Professor Shahin Shojai, who returns to lead the Capco Institute. Many will recall that Shahin held this position from the Institute's inception in 2001 and it's great to welcome him back as the Journal and the Institute continue to develop.

I hope you enjoy this edition of the Journal. These are exciting times and I am pleased to share these fascinating articles with you all.

Lance Levy
CEO, Capco

The ever-evolving face of risk

The risks facing financial institutions have evolved over the past couple of years. As well as the typical risks that financial services firms face, such as investment, regulatory, operational and economic risks, they are also confronting a genuine existential threat from new entrants that possess a much deeper understanding of how innovative technologies could be used to simplify the lives of clients, and to do so at a fraction of a cost. These players are able to apply new business models to old industries – such as generating revenues from advertisers rather than clients – and are aggressively disruptive. There is no reason to think they will treat our industry any differently.

It is for this reason that the first section of this Journal is dedicated to the risks posed by new entrants. These small and agile organizations see our industry as suitable for radical disruption – and they have just got started. While many financial services firms are learning to apply some of these new technologies, such as blockchain, to make their businesses more efficient, the speed of change and the huge financial resources supporting the newcomers might result in a few of today's established players being left behind. Our hope is that the articles in this section provide a suitable overview of some these risks.

As well as the risks posed by the new entrants, financial services firms are still coming to terms with the never-ending list of new regulations and their direct and unintended consequences. The experts who have contributed to the regulatory risk section of this edition shed light on the new regulatory environment and how financial services firms can view, and respond to, the onslaught of financial regulation. The articles cover topics such as risk culture, questioning the viability of financial services firms to actually improve their cultures rather than simply becoming compliant, the implications of the Dodd-Frank Act and the Volcker Rule, as well as providing an interesting review of how some of the major regulatory bodies came into being and how their evolution has influenced the way they regulate.

Finally, in Section 3 of the Journal we focus on investment risk. The articles touch on how institutional investors can be educated to become better informed and why we should be careful about assuming that professional investors have a deep understanding of investments and their inherent risks.

We hope that you find the articles in this edition of the Journal of interest, and that you continue to support us in our endeavor to bring the best thinking to senior financial executives by sharing your ideas with us.

On behalf of the Board of Editors,

Shahin Shojai

New Entrants

Crowdfunding: A New Disruptive Technology?

Get Bold with Blockchain

The Role of Financial Institutions in Advancing Responsible Value Chains

Robo-Advice 2.0: The Next Generation

Crowdfunding: A New Disruptive Technology?

Roy C. Smith – Kenneth Langone Professor Of Finance and Entrepreneurship, Stern School of Business, NYU
Won Jun Hong – Former Research Assistant, Stern School of Business, NYU

Abstract
The Jumpstart Our Business Startups (JOBS) Act was passed by Congress with bipartisan support and signed into law in 2012. Many regulators and investor advocates opposed the new law because the securities it was enabling the sale of were very risky, and the public at large was unlikely to fully understand these risks, which include over-promotion, misrepresentation, mispricing, and manipulation of prices in aftermarket trading.

The first IPO under the new crowdfunding rules, a U.S.$17 million issue by Elio Motors, has now been completed successfully. Between the SEC's new rules and new procedures developed in the market, a different way to access investors in start-up companies has been created which could provide an alternative pathway for many companies to raise early state capital. If it catches on, then much of what we know about start-up financing could be changed forever; the new pathway could disintermediate the risk capital industry, just as Uber has done to taxis, and Amazon has done to retailing. The change could be very big.

The Jumpstart Our Business Startups (JOBS) Act was passed by Congress with bipartisan support and signed into law in 2012. Its purpose was to enable "emerging growth" startup companies to raise capital directly from the public through the internet in order to help them grow and create more jobs. Many regulators and investor advocates opposed the new law because the securities it was enabling the sale of were very risky, and the public at large was unlikely to fully understand these risks, which include over-promotion, misrepresentation, mispricing, and manipulation of prices in aftermarket trading.

EARLY FORMS OF CROWDFUNDING

The basic idea behind crowdfunding is to establish internet sites through which companies can announce themselves and their business plans and solicit funding without going through expensive venture capital or underwriting processes.

Kickstarter, founded in 2009, is a for-profit "benefit corporation" that is permitted to consider benefits to society as well as profits in its business activities. Kickstarter enables artists and other creative people, and companies with interesting new products or designs, to demonstrate their products usually through a 30-day online "campaign," and to receive funding in the form of "donations" in exchange for rewards, premiums, or an opportunity to acquire the product as soon as it becomes available to the public. Kickstarter does not sell or broker stocks in companies, but it does enable startups to raise small amounts of initial funding to launch their first batch of products.

Since 2009, Kickstarter has raised over U.S.$2 billion in campaigns for 100,000 projects from nearly 10 million people, including 135 campaigns that raised more than U.S.$1 million for companies like Oculus Rift (virtual reality), Pebble Time (electronic smart watch), Dash (Wireless Smart Headphones), and the Micro (3D Printer).[1]

Kickstarter is not the only player in this crowdfunding sub-market. GoFundMe has raised over U.S.$1 billion for personal advocacies, such as education, environment, and minority empowerment related issues. Indiegogo competes directly with Kickstarter and has a presence in Canada, U.K., France, and Germany. Smallknot offers similar features as well.

STARTUP FINANCING

New "startup" businesses typically obtain their initial financing from their founding investors, friends and family members, angel investors or corporate partners, and from venture capital investment firms (VCs). Many startups, failing to have access to the other sources of funds, appeal to VCs for their initial financing in what is essentially an asymmetrical market. VCs reject most of the proposals made to them, and only invest in companies at pricing levels that could provide a high return to them, and thus involve considerable dilution of their founders' equity, along with other terms that reduce the founders' powers of control, and leave the company committed to the VC relationship indefinitely.

According to The National Venture Capital Association, in 2015, VCs invested U.S.$58.8 billion, about half of which was in seed capital or early stage investments. 1,400 companies raised venture capital for the first time in 2015. VC funds were also selling shareholders in 45 IPOs (27% of all such issues), valued at U.S.$17.4 billion.

Thus it is clear that VCs control an important pathway to startup financing in the U.S., but, at the same time, the total number of companies able to access this pathway is quite small relative to the number seeking startup funds.

INITIAL PUBLIC OFFERINGS (IPOS)

IPOs are an important part of the market for new issues of equity securities. In 2015, 169 companies raised around U.S.$30 billion via IPOs (a 35% decline from 2014 and the lowest volume in six years). Of these issues, only about 30 were IPOs of U.S.$50 million or less.[2] Consequently, the IPO market is also not a major source of startup or early stage financing, relative to the demand for such funds.

IPOs are filed with the SEC on standard Forms S-1 (but with reduced disclosure and exemption from audits of internal controls available to "emerging growth" companies, thanks to the JOBS Act). According to the SEC, around 75% of all smaller company IPOs issued after April 2012 were undertaken by companies that identified themselves as emerging growth companies. These issues had to be sold only to "accredited investors," meaning that the investor must have more than U.S.$1 million in net worth (excluding their primary

1 https://www.kickstarter.com
2 http://www.renaissancecapital.com/ipohome/press/ipopricings.aspx

residence) or income of over $200,000 per year ($300,000 if married) for the past two years and expect the same level of income in the current year.

On Oct. 30, 2015, the SEC released further rules applicable to exemption from registration for certain smaller crowdfunding transactions (under Title III of the JOBS Act), and provided a framework for the regulation of "registered funding portals" and broker-dealers that issuers are required to use under the new rules.

The requirements for obtaining the exemption under Title III are that (a) the amount raised not exceed U.S.$1 million in a 12-month period, (b) that individual investments in all crowdfunding issuers in a 12-month period are limited to the greater of U.S.$2,000 or 5% of annual income or net worth, if such is less than U.S.$100,000, or 10% of net income or net worth (not to exceed U.S.$100,000) if annual net income or net worth of the investors is U.S.$100,000 or more, and (c) transactions are conducted through a registered broker-dealer or a registered funding portal.

SEC REGULATION A+

On June 19, 2015 (as required by the JOBS Act), the SEC announced new crowdfunding rules for sales to "non-accredited issuers" (i.e., more or less ordinary retail investors) under its "Regulation A+." These issues are divided into Tier-1 and Tier-2 offerings. Tier-1 offerings are limited to U.S.$20 million within a 12-month period and require compliance with State "Blue Sky" securities laws that authorize the sale of possibly speculative securities in the state.

Tier-2 offerings are capped at U.S.$50 million within a 12-month period. Tier-2 offerings, however, preempt State Blue Sky securities laws. This provides a significant advantage as it eliminates the burden and expense associated with compliance with numerous individual State Blue Sky securities laws.

For Reg A+ offerings, issuers must file registration statements on (a new) Form 1-A that, though abbreviated compared to Form S-1, still must contain all information material to an investment decision. These issues may be sold to non-accredited investors if the investors certify that their investment in the issue being offered will not exceed 10% of the greater of their net income or net worth.

StartEngine Crowdfunding, a for-profit affiliate of a corporate "accelerator" founded in 2011,[3] was formed to assist startup companies issue new stock under Regulation A+. It does not underwrite issues or take positions in them. It is not licensed as a broker; it is, however, "an SEC registered funding portal" that connects startups to unaccredited investors via the Internet. In 2015, Startengine Crowdfunding was charging issuers U.S.$20 per investor, regardless of the size of the purchase; however, this was raised to U.S.$100 per investor after the Elio offering.

ELIO MOTORS IPO

On August 28, 2015 Elio Motors, a startup manufacturer of a slick looking, U.S.$6,800 two-passenger, three-wheeled minicar that gets 84 miles per gallon, filed the first Form 1-A registration statement for a Type-2 IPO.

Elio's founders invested U.S.$5 million in the company at an average price per share of U.S.$0.26. Accredited investors have also purchased an additional U.S.$9 million of shares at an average price of U.S.$1.48 per share through private placements. In 2015, the company issued U.S.$3 million of subordinated secured notes convertible into common stock at U.S.$5.98 per share. By the end of 2014, the company had also raised U.S.$58.6 million of long-term debt.[4]

This financing was done with no investment by VCs.

It is not that the founder, Paul Elio, did not reach out to VCs. Every time he pitched his idea to one of them, he encountered skepticism that there would ever be a mass-market for the tiny, three-wheeled commuter car. No single small-sized vehicle has ever had a material success in the U.S.; even the globally successful small cars, such as Daimler Benz's Smart and Fiat's 500C, ended up being uninteresting economically in the U.S. Small cars that are popular in European markets did not appeal to American consumers, who are used to larger vehicles. Besides, the VCs said, the costs of just testing the idea were quite high relative to the expected payoff. The automotive industry was not really suitable for VC investments, they said, Tesla Motors notwithstanding.[5]

To demonstrate market demand and raise some startup funds, in January 2013 Elio began introducing a two-tier (refundable and non-refundable) vehicle reservation system. A potential buyer can reserve a spot for future delivery of a vehicle by depositing

3 Its founders are Paul Kessler, a prolific venture capitalist who has invested in and/or worked with over 500+ companies, and Howard Marks, a founder of Activision (now a part of Blizzard-Activision), known for its blockbuster games Call of Duty.
4 Elio Motors Form 1-A, November 20, 2015
5 http://online.wsj.com/mdc/public/page/2_3022-autosales.html

an amount from U.S.$100 to U.S.$1,000. Non-refundable depositors will have priority for vehicle delivery and receive a discount that amounts to 50% of the committed deposit.[6] Tesla employed such a two-tier scheme for its Model S reservation.

By January 1, 2016, Elio had collected more than 50,000 reservations for vehicles on its website, locking in more than U.S.$340 million in advanced product orders, and U.S.$21.1 million in deposits (80% of which are nonrefundable). Advanced reservations for vehicles are different from equity crowdfunding. Kickstarter and similar crowd-accessing donation platforms are closer to advanced reservations than to actual equity fundraising.

Elio hoped to raise sufficient funds from its equity crowdfunding issue to fund prototype building and testing of 25 vehicles to be used to demonstrate various performance and safety features required to obtain a major loan from the U.S. Department of Energy.

If the Elio car meets the required performance tests it may be able to tap into the U.S. Energy Department's Advanced Technology Vehicles Manufacturing (ATVM) loan fund.[7] If so, this could provide up to U.S.$185 million of additional capital for the company.[8] The most noteworthy recipient of ATVM loan was Tesla Motors with a loan of U.S.$465 million in January 2010.[9] ATVM loans are highly attractive to the automotive industry since the loans carry low interest rates with long maturities and minimal fees.

In order to qualify for ATVM loans, the company must show that it is financially and technologically capable of vehicle production. In addition, the company's vehicles must meet the Energy Department's fuel efficiency, component quality and manufacturing location standards, which requires that all vehicles be assembled in the U.S.

Thus, in order to raise the U.S.$235 million of startup capital it requires to begin production, Elio must be able to qualify for an Energy Department loan, and to obtain the loan it must first raise about U.S.$20 million in additional equity capital.

Enabled by Startengine Crowdfunding, Elio Motors sought non-binding "indications of interests" for up to U.S.$25 million of equity from accredited and non-accredited investors over a three-month, online testing period. The idea was to gauge market sentiment to determine an appropriate price level and number of shares to be sold from the indications of interest.

For traditional public offerings of equity shares in startup companies, a registration statement containing voluminous information about the company and the risks involved in investing in it is filed with the SEC, at considerable expense to the company. (Regulation A+ does allow emerging growth enterprises to offer shares with less cumbersome SEC disclosures than are required for traditional IPO processes). Underwriters must also be retained to purchase and resell the stock being offered based on demand estimated by pre-offering marketing and sales efforts.

In August 2015, Elio Motors closed its non-binding, three-month market test with over U.S.$42 million of interest in purchasing the shares indicated by 11,000 (mostly non-accredited) investors with an average order of U.S.$3,820.[10]

On August 29, Elio Motors filed a registration statement on the newly approved abbreviated Form 1-A to obtain SEC approval for the offering. The proposed offering was to be of a minimum of 1 million and a maximum of 2 million shares. The expected offering price, set by the Company, was U.S.$12 per share.

The registration statement disclosed that Elio had not yet sold any vehicles, and in 2014 had lost U.S.$20.7 million and ended the year with a cumulative shareholder deficit of $45 million. For the six months ended June 30, 2015, the unaudited results showed a net loss of U.S.$8.8 million and an accumulated deficit of U.S.$53.8 million. The SEC reviewed the registration statement over a period of about three months, requesting changes or additions as appropriate. After a series of amendment filings, Elio Motors finally obtained approval for the offering from the SEC in late November 2015.

Elio retained Fund America Securities, a registered broker dealer, to perform several administrative functions under the new rules in connection with the offering, including determining investment limits for subscribing investors, certifying that investors are qualified, applying checks against money-laundering, serving as a registered agent for Blue Sky filings, and transferring subscription information to Elio's transfer agent.

6 Investor information can be found at: http://ir.eliomotors.com/

7 "To date, the program, which is administered by the U.S. Department of Energy's Loan Programs Office, has made over $8 billion in loans, including loans to Ford ($5.9 billion), Nissan ($1.45 billion) and Tesla. The ATVM loans are made attractive to applicants due to their low interest rates (set at U.S. Treasury rates (approximately 2% to 4%), minimal fees (no application fees or interest rate spread and only a closing fee of 0.1% of loan principal amount), and long loan term life of up to 25 years (set at the assets' useful life)." Source: Security and Exchange Commission EDGAR Database

8 Wall Street Journal, http://www.wsj.com/articles/Mr. Elio-elios-quest-to-build-a-three-wheel-car-1433301222

9 US Department of Energy, Loan Programs Office, http://energy.gov/lpo/tesla

10 Eliomotors.com

RESULTS OF THE ELIO OFFERING

The offering was conducted online via the Startengine Crowdfunding website for 74 days from late November 2015 to late February 2016, during a period in which the S&P 500 stock index dropped 6.8% and VCs and other investors in many high visibility technology "unicorns" took substantial write-downs.

In February 2016, the company announced that it had accepted orders for U.S.$17 million of shares (5.3% of the company) from 6,600 investors, which capitalized the company in the market initially at U.S.$321 million.

Trading in the shares began on February 19, 2016 on OTCQX, an over-the-counter exchange. One week after the offering, shares were trading at U.S.$16.50 and soon thereafter increased to U.S.$37 per share, and by February 29 reached a high of U.S.$75 per share, before dropping sharply to U.S.$20.75 on March 4, underscoring the extreme volatility that can occur in thinly traded markets. Trading volume was only in the hundreds of shares during most of the period following the offering. The tradable "float" in the company's shares, even after a tripling of the share price, was still only U.S.$52 million, an amount too small to attract interest from large institutional investors.

WHAT IS DIFFERENT ABOUT THE ELIO OFFERING?

Since Elio had been denied venture capital financing, the offering essentially allowed the company to turn to ordinary investors as an alternative source of startup capital, and to do so at a much lower cost than VC investors would have required had they been willing to invest. The offering represented only 5.3% of total shares outstanding; VC investors, as a group, typically own 30%-50% of outstanding shares by the time of a traditional IPO.

The Elio IPO involved no underwriters or underwriting fees. Legal and other fees associated with the offering, excluding fees to Startengine and Fund America Securities, a broker-dealer acting as a sales agent, amounted to approximately U.S.$150,000, or only 0.1% of the amount raised. Total expenses of the offering, based on pro-rating Elio's estimated minimum and maximum amounts, were U.S.$1,689,000, or 9.9% of the proceeds received. Of these, per-investor fees to Startengine Crowdfunding at U.S.$20 per investors were U.S$132,000; fees paid to Fund America totaled approximately U.S.$649,000. We also estimate that Elio spent approximately U.S.$750,000 on marketing and public relations in connection with the offering. Traditional IPOs of comparable size generally involve underwriting fees and commissions of approximately 7% plus legal, auditing, and other expenses of another 2% to 3%.

Shares were priced by Elio (not by underwriters or venture capital investors) after a three-month market trial at U.S.$12 per share (up from U.S.$5.98 per share for a private placement of convertible debentures earlier in 2015).

The shares were marketed entirely thorough the Internet using user-friendly Startengine and Elio's websites, which enabled potential investors to "reserve" shares in the offering on a "non-binding" basis (as well as reserving the company's product when it became available). Once the SEC permitted the offering to be sold, investors were contacted by email to confirm their acceptance of their allocation of shares.

After the offering the shares were traded on OTCQX, an over-the-counter exchange specializing in small companies operated by OTC Markets. The shares are not being listed on NASDAQ or the NYSE. Trading in the shares is limited and in small amounts, suited to "ordinary" retail investors. However, despite that, in the after-market following the IPO, Elio shares initially rose to a 38% premium over the offering price despite a significant downturn in the stock market indices, and in preliminary pre-IPO valuations of high-visibility technology companies.

Although Elio's Reg A+ fundraising did not meet its maximum goal of U.S.$25 million, raising U.S.$17 million still impressed a lot of entrepreneurs. Following Elio's offering, over 40 companies made Form 1-A filings and 12 had received their approvals by February 20, 2016.[11] Startup companies in many different industries, including healthcare, banking, and even cannabis distributers, are now eyeing crowdfunding as an alternative to traditional venture capital or initial public offerings in order to tap a different source of funds with less equity dilution.

IS CROWDFUNDING AN "UBER?"

"Uber" refers in this context to the ridesharing company's disruptive challenge to a settled industry. Bypassing venture capital or private equity investors and the traditional Wall Street dominated IPO process to access ordinary investors through the Internet could certainly be disruptive if Elio's success is repeated by other companies.

11 Porter Wright Morris & Arthur LLP, 2015, "Diverse companies receive SEC approval to raise fund with Regulation A+," Federal Securities Law Source, December 15.

Whether it will or not will depend on the longer-term success of the investments, i.e., whether investors will be able to expect to sell shares purchased in the IPO at a later date at a profit, and whether the issuers of the shares will find the process easier and cheaper to use than the traditional methods.

It is possible, however, that the very limited liquidity in the crowdfunding stocks will cause prices to decline and exit opportunities to be constrained, which could discourage future offerings. If crowdfunding investor demand is reduced by poor performance of the initial issuers or by frustrations with the available liquidity, then crowdfunding may fail to offer a viable alternative to traditional methods.

This could happen because of the inability of ordinary investors to analyze risky investment opportunities, or their inability to endure high levels of volatility associated with risky investments, or because of over-promotion or mispricing by the companies involved.

Bypassing the skilled financial intermediaries also deprives companies of their experienced advice and counsel that many VCs and underwriters highlight as a compelling reason for using them.

In the traditional early stage financing methods, venture capitalists or underwriters vet the companies thoroughly and agree to pricing levels at which they are willing to risk their own money. It has long been thought that this screening process generates value to investors, and that investors are prepared to reject alternative processes that do not include it. Crowdfunding now presents this unscreened alternative, and the Elio Motors offering suggests that the perceived value of the vetting may have been exaggerated.

For many years, angel investors (individuals investing in startup situations) have grown to become significant players in the venture finance area, with 316,000 investors funding 73,000 companies in deals worth U.S.$28 billion in 2015.[12] Angel financing assists more startups than traditional VCs do. Angels do not rely on VCs or other traditional intermediaries to do their screening, they do their own. Crowdfunding increases angel investors' knowledge of and access to deals well beyond what they might encounter on their own.

In addition, for many years the U.S. IPO market has included many smaller companies issuing shares in modest amounts. Biotech companies, in particular, tend to use the IPO and follow-on equity markets as a substitute for additional rounds of venture capital finance, despite limited liquidity in their shares. On the whole, there has been enough success in smaller size IPO market for it to continue to attract investment capital.

Further, ordinary investors have been able to purchase shares in traditional IPOs for years, but rarely get a chance to do so because underwriters allocate shares in the IPOs to hedge funds and favored high-net-worth clients. Even then, despite considerable vetting by underwriters, many IPOs disappoint investors in the aftermarket. Crowdfunding certainly removes barriers to entry that prevented ordinary investors from participating in the IPO market.

What crowdfunding does is to bring the power of the Internet to the startup funding market. Between the SEC's new rules and Startengine's new procedures, a different way to access investors in startup companies has been created that, after some early learning experience, should provide a viable pathway for many companies to raise early state capital.

If it does work, then much of what we know about startup financing could be changed forever. Most startups have to fight for VC attention, and submit to tough pricing demands and loss of control when they do get noticed. By the time a company is ready for an IPO (something usually decided by the VCs), they must submit again to considerable legal and accounting expenses and the 7% commissions that have been demanded by underwriters for generations, and considerable legal and accounting expenses and then take their chances that aftermarkets will provide adequate liquidity to sustain the price level.

If it catches on, platform companies like Startengine will expand, improve, and attract competition which may lower fees. They will be able to disintermediate the risk capital industry, just as Uber has done to taxis, and Amazon has done to retailing. The change could be very big.

12 Torres, N., 2015, "What angel investors value most when choosing what to fund," Harvard Business Review, August 6. Available at: https://hbr.org/2015/08/what-angel-investors-value-most-when-choosing-what-to-fund.

Get Bold with Blockchain

Benjamin Jessel – Managing Principal, Capco
Tommy Marshall – Partner, Capco

Abstract
Distributed ledger is a technology that offers the potential to significantly disrupt the financial services industry through a new paradigm that could ultimately result in the trade and settlement cycles across many assets shrinking from days to seconds. As with many new technologies, the hype cycle is in full effect, with many highlighting the huge potential of the technology without any references to the fact that market infrastructures evolve through decades not years, and where regulation often lags several years behind technology advances.

The technology for distributed ledger has been – for the most part – proven, and many financial institutions have begun to understand that it can be applied as a technology design pattern that can enable a small network of invited participants to collaborate over a secure network. Key recent developments such as the DTCC's announcement of a Repurchase Agreement (Repo) proof of concept and the trail of the technology by the Australian Stock Exchange and Japanese Stock Exchanges demonstrate that this technology is beginning to be taken seriously.

The debate fundamentally comes down to interoperability and the "network effect" – can banks cast off their conservative philosophy, and risk averse approach to new technology and work together to build distributed ledgers or will we have to wait until there are outside pressures from regulators or the fintech community before this evolution starts in earnest? In this article, we urge consultancies and financial institutions to be "bold" about blockchain; specifically to develop their thinking away from the headlines and high level narrative and objectively assess the use cases for the technology in detail.

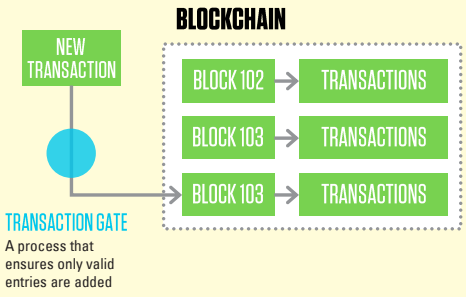
There is no doubt that distributed ledgers and the technology that fuels it, "blockchain," is a hot topic. There is also a growing consensus amongst credible, senior members of the financial services industry that this technology, be it centralized, decentralized, or federal ledger, will have a profound and lasting effect on their industry.[1] Many say we are witnessing the beginning of a true democratization of finance – a period during which the closed network of correspondent banks and counterparties could be replaced by an open, more transparent, and perhaps even safer system (as depicted in Figure 2).

This approach has the potential to reduce costs through commoditization of contracts, increased process speeds and reduced settlement risk while increasing trade transparency. But when will banks put blockchain to the test? Who is going to Get Bold with Blockchain?

IS BLOCKCHAIN PAYING ITS WAY? BLOCKCHAIN AND PAYMENTS

Initially, the focus on distributed ledger and blockchain use cases was firmly on cross border payments, on the basis that bitcoin was a technology that enabled payments transfer. This is not surprising given that the recent rise in bitcoin price has been predominantly cited on capital flight from countries such as China.

A blockchain is a database shared by every participant in a given system. The blockchain stores the complete transaction history of a cryptocurrency or other record keeping system.



Transactions aren't recognized until they are added to the blockchain. Tampering is immediately evident, and the blockchain is safe as record because everyone has a copy. The source of discrepancies is also immediately obvious.

Source: http://www.zdnet.com/article/how-blockchain-is-likely-to-change-it-and-business-forever/

Figure 1 – How blockchain works

We have seen companies like Ripple™ striving to replace cumbersome traditional network of correspondent banks with blockchain inspired solutions that provide the promise of faster, simpler and cheaper payments.[2]

This technology has the promise of allowing users the ability to prioritize payment methods based on criteria such as optimization of speed, compliance, counterparty risk or exchange rates. This could be a game changer for financial services institutions not able to fine tune business decisions in this way before. We also are beginning to see banks experiment with crypto-finance payments on blockchain rails. Pioneers, like FIS's Clear2Pay,[3] and start-ups like Earthport™[4] are busy forming partnerships to explore what banks, and blockchain, could do together. Yet still, sweeping change remains elusive.

BLOCKCHAIN FROM FRONT TO BACK OFFICE: WHERE'S THE POTENTIAL IN CAPITAL MARKETS?

In trading, clearing, and settlement, proponents of blockchain technology predict a complete replacement of what many feel is an antiquated system, full of human processes and trapped funds. Imagine one sleek design, which mitigates current challenges/risks, such as settlement time and custodial risk (as identified in Figure 3) while putting to rest the snarl of private, bi-lateral ledgers that record asset ownership and liabilities in equities, bonds, and derivatives markets. Industry focus then centers around whether the present structure of capital markets trading gives way to a centralized shared database of trades built on crypto-finance principles. Some industry experts believe that the concise clarity of blockchain technology, and the options we have in adopting it, could have saved us from the lethal crisis of confidence that fueled our last market collapse by providing a full auditable view of ownership of all assets and liabilities. Distributed ledger and the blockchain could potentially take out systemic market risk and make markets safer.

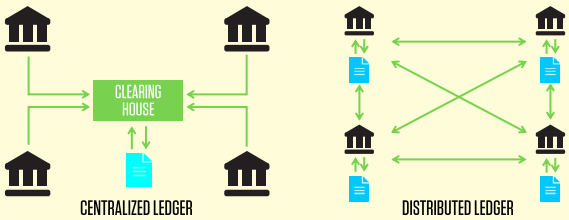
1 Centralized ledger: transactions are recorded centrally by one party. This is most analogous to the current banking system where there is a central bank /depository/ custodian recording transactions and ownership (e.g., the DTCC today). Distributed ledger: each participant owns a copy of the ledger, which is updated each time a transaction is made and then confirmed. A consensus (egg 51%) needs to be reached in the group to confirm ownership. Federated ledgers (sidechains/altchains): similar to distributed ledger but with a degree of exclusivity built in. Depending on how they are set up, either these can take the form of private blockchains only available to a selected few, or alternatively access is open but participants can choose whom they transact with.

2 https://www.ripplelabs.com

3 http://www.fisglobal.com/C029864

4 http://www.earthport.com

A distributed ledger is a network that records ownership through a shared registry

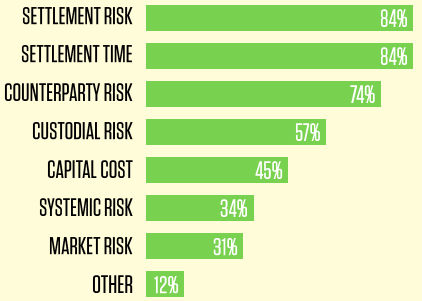


In contrast to today's networks, distributed ledgers eliminate the need for central authorities to certify ownership and clear transactions. They can be open, verifying anonymous actors in the network, or they can be closed and require actors in the network to be already identified. The best known existing use for the distributed ledger is the cryptocurrency Bitcoin.

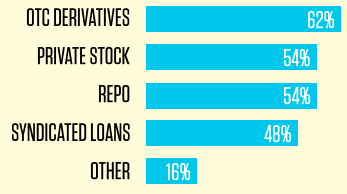
FT graphic. Source: Santander InnoVentures, Oliver Wyman & Anthemis Partners

Figure 2 – Embedding distributed ledger technology

Which of the following do you believe distributed technology could help reduce?[a]



Other than payments and digital currency, what area/products could most benefit from the technology?[b]



[a] Based on 58 respondents in 2015.
[b] Based on 61 respondents in 2015. Source: Greenwich Associates 2015 Bitcoin, Blockchain and The Capital Markets Study

Source: http://www.coindesk.com/survey-47-of-finance-pros-say-firms-exploring-blockchain-tech/

Figure 3 – Distributed ledger technology benefits

The ability to be able to reduce – and in some cases eliminate – clearing is appealing to financial institutions given the current cost of reconciliation of trades. Furthermore settlement could be significantly speeded up. A key new paradigm of the technology is that "the trade is the settlement." In other words, because the change in ownership of assets can occur at trade time, so can the change of ownership of the assets involved in settling that trade.

That is significant implications for capital – the velocity of capital on banks' balance sheets accelerates, meaning that capital no longer has to sit on the sidelines awaiting the settlement of a trade. The capital impact of moving from T+3 to T+0 across asset classes would be a significant prize. Yet, many counter this argument and claim that it is not necessarily a technology constraint that we cannot move to shorter settlement times, but how the market is structured today. Moving to shorter settlement times in many cases actually introduces more problems than it solves.

POTENTIAL ENERGY

But these are scant few examples when compared to the size of the financial services market as a whole. So what is less clear is when – and specifically how – blockchain technology will work with traditional core systems at global banks. 47% of finance professionals say their firms are exploring opportunities in the area, including a number of the use cases referenced in Figure 4.[5]

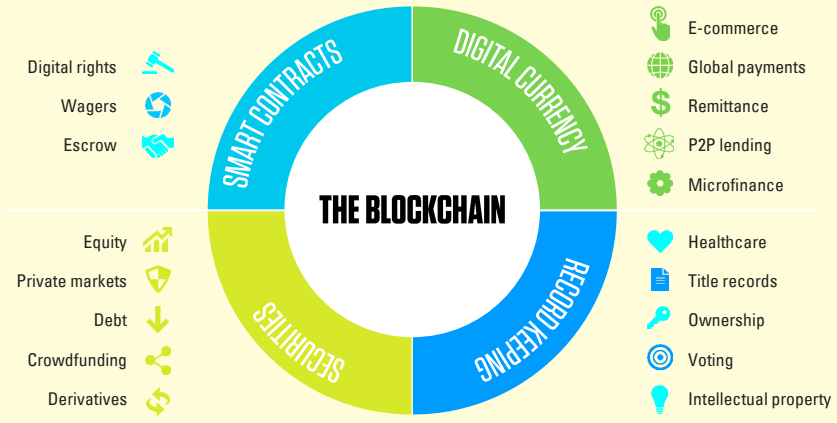
Most top tier banks now have distributed ledger labs or are actively engaged in the R3CEV forum or the open ledger forum group, but few have been able to move to an actual business proposition or actionable plan. Many are waiting for the outcome of the tests from the R3CEV group.

With preliminary testing having proved the potential of the technology, the financial services industry is now poised to put blockchain and distributed ledger to the test.

WHAT'S IN THE WAY?

There are a number of things holding back the cataclysmic changes predicted to transform financial services due to blockchain technology.

5 http://www.coindesk.com/survey-47-of-finance-pros-say-firms-exploring-blockchain-tech/



Source: http://www.capco.com/uploads/articlefiles/635/file_0_1445593031.pdf

Figure 4 – Blockchain use cases

First, the short-term market opportunity appears small for large banks. Tight margins and strained budgets encourage IT executives to tweak the status quo until the returns on investments for blockchain are clearer. In the clearing and settlement space, there is as yet not the appetite – or clear financial incentive – to dismantle and replace the current banking infrastructure, a costly and risky undertaking.

Second, a key adoption hurdle for the technology is the requirement for a number of industry participants to adopt a common standard and technology. With the large number of market participants involved in equities trading, clearing, and settlement, the likelihood of these organizations all adopting a new data and technology standard soon is low.

The automation potential of the distributed ledger and blockchain is significant. Financial instruments can be digitized into "smart contracts" that are effectively financial contracts translated into computer programs that could remove vast swathes of operational roles currently conducted by humans. However, not only does this require an agreement and adoption between financial institutions to a standard, but also an acceptance by compliance contracts to run computer programs between financial institutions that directly integrate the financial systems of financial organizations. Even if computer code is audited, concerns about compromise and security will be hard to overcome. Another key challenge is to determine to what extent contracts are similar enough to enable them to be encoded without the need for many different types of contracts to be built.

Finally, the industry market infrastructure has been designed around a trading and settlement process of days rather than seconds – financial institutions are simply not yet designed and geared for concepts such as real-time netting of positions, and having capital to settle transactions within seconds of trade rather than days.

But where we are seeing interesting opportunities is where there is less of a market infrastructure established, where it is consolidated, or where there are few market participants involved in the trade lifecycle. For instance, R3CEV – a distributed ledger consortium of over 40 banks – is actively trialing commercial lending use cases, which is a market without a significant infrastructure in place today, with few market participants. Other examples include the Australian Stock Exchange, which is an exchange that is unique in the degree to which it has a consolidated hold on the market – it operates a central securities record-keeping or ledger-keeping service and is responsible for the totality of the settlement and clearing process.

Finally, the DTCC announced earlier last month a trial of distributed ledger in the repo market. This is a market with a small number of market participants, with a relatively standardized asset. In this market, rapid clearing and settlement isn't the desired outcome – repos already settle on a T+0 basis. Instead, the opportunity is a more optimal allocation of capital and the ability to achieve trade compression, which should improve liquidity ratios. Given that the DTCC has been skeptical on distributed ledger in recent months, this – to us – signifies the heralding of a new age in the story of distributed ledger technology.

This development, combined with the developments in the open ledger / hyper ledger project, Digital Asset Holding's announcement of a proof of concept with the Australian Stock Exchange, and IBM's announcement of a trial of distributed ledger technology with

the Japan Exchange Group suggests that distributed ledger has moved from the theoretical into the practical.

DÉJÀ VU ALL OVER AGAIN

The catalyst for change will likely take a radical regulatory edict, such as a move to a T+ 0 (Trade + 0 day) settlement protocol. "In the 1970's 'Paper Crisis,' firms were choking on a backlog of paper orders caused by a spike in equity trading volumes," recalls a market veteran. "Regulators, like the Federal Reserve and the SEC, came in and said, 'Guys, clean up your customer order confirmations or you're out.' A lot of firms closed their doors due to lack of technology. Those who invested too late didn't make it. Those who did make it were the ones who invested in automation before it was absolutely necessary. I look at blockchain and it feels like it's déjà vu all over again."[6]

COMPLIANCE OR COMPLACENCE?

Another hurdle is that the future regulatory environment remains unclear as the pace of regulation continues to lag behind the speed of technological innovation. Because early Bitcoin exchanges demonstrated the risks involved in being on the wrong side of ambiguous regulatory guidance, would-be industry participants are spooked and cautious. Compliance departments will have to do a lot of due diligence before becoming comfortable with their financial institutions getting on the blockchain train. But with precautions understood, are financial institutions being too casual with their "wait-and-see" attitude toward blockchain? Are questions around vague compliance allowing banks to be complacent?

FIRST TO BE SECOND

The road forward will not be easy. For starters, many of the new entrants that may provide liquidity on blockchains in the future are not going to be traditional financial institutions. In addition, new models, unfamiliar processes, and unknown exchanges will pose challenges for banks trying to calculate counterparty risk in a foreign environment. As we all know, people do not like change and financial services professionals are no different. Industry consortia are failing to push potential use cases and shape regulation due to an inability to agree on standards as well as a tendency to play it safe. When it comes to blockchain, it seems as if everyone wants to be first to be second.

Other major areas of enhancement include, but are not limited to:

- Instantaneous confirmation and settlement
- Lack of need to warehouse "Physical" assets but rather record securely using blockchain technology
- Immediate pricing transparency from point of execution
- Enhanced settlement security
- Increased market liquidity with immediate settlement
- Simplified Dodd Frank reporting via real time auditing transactions posted to the blockchain
- Reduced counterparty and delivery risk
- Reduced indirect transaction costs
- Reduced in back office operations overhead

Source: http://www.capco.com/insights/capco-blog/will-blockchain-kill-clearing-part-ii

Figure 5

WHERE IS THE OPPORTUNITY?

But for the bold of heart there are opportunities that are ready to be realized by blockchain today. Innovators are excitedly pointing toward new markets to kick the tires, as Nasdaq has committed to do by leveraging the blockchain platform to facilitate the issuance and transfer of shares of privately held companies.[7] Markets where automation is still limited are also likely to see fast adoption. Industry insiders like Symbiont™ are starting to whisper about anticipated inroads into the corporate debt markets first.[8] Analyst houses like Greenwich Associates have their eye on all manners of leveraged and syndicated loans as the primary feeding ground for blockchain disruption.[9] Further, the open ledger forum, which includes IBM, the Linux foundation and Digital Asset Holdings, are now starting

6 http://www.capco.com
7 Read more: http://www.nasdaq.com/press-release/nasdaq-and-chain-to-partner-on-blockchain-technology-initiative-20150624-00446#ixzz3lLfFPhHP
8 http://symbiont.io
9 https://www.greenwich.com

to deploy distributed ledger components to enable the distributed ledger movement transition from the theoretical to the practical.

With regulation from Dodd Frank's Volcker rule pushing dealers away from taking principal positions toward a pure agency model, blockchain technology may be just what dealers are looking for to lower transaction costs in a growingly restrictive trading environment.

WINNER TAKES ALL? TACKLING BLOCKCHAIN

Consulting professionals are ringside center in trying to assist global banking clients to navigate the new territory blockchain unfolds. A number of areas of enhancements, as noted in Figure 5, are sought after, but as one proponent stated: "Everyone is scrambling to be in this field for fear of missing out...but few have a clear view as to the differentiation and competitive advantage that the technology will bring to their business, let alone a plan to actually implement and monetize its possibilities."

Of course, the full implications of blockchain on financial services will not be realized until the majority of players are using the technology, but it is becoming clear this is not a good enough reason to sit on your laurels. As another expert said: "By all means develop a proof of concept as a first step in the process, but you have to be bold. Define a proposition that you believe is truly game changing for your organization, and the industry, and proceed aggressively in that direction. Take the risk. To my mind, this is not a game of feature parity. It could just be that it's winner-takes-all."

The Role of Financial Institutions in Advancing Responsible Value Chains

Herman Mulder – Fellow, Nyenrode Business Universiteit[1]

Abstract

2015 was a landmark year for making businesses aware of the consequences of their actions on society. The adoption by the UN of the Global Goals for Sustainable Development and the Paris Climate Agreement are important steps in this regard. The private sector has an important role to play in the realization of the promises made. The financial services sector should be a key driver of public and private sectors to support (by research) and realize (by co-funding) this agenda. It is a business opportunity and a societal obligation. This article explores why and how this sector may become the key driver for the transformative momentum, what it can do itself, and how others (governments, civil society organizations) should support and enable it.

1 Herman Mulder is also an independent member of the Dutch NCP for the OECD Guidelines, Co-founder of the True Price Foundation, TEEB Advisory Board Member, Board Member of Worldconnectors; former Chairman of the Board of the Global Reporting Initiative (GRI), former Board Member of Utz Certified, former Director-General Group Risk at ABN AMRO Bank, initiator of the Equator Principles for Project Finance. This article is an update of July 2015 paper by the author as part of the Strategic Series of the Duisenberg School of Finance.

INTRODUCTION

Responsible business conduct (RBC) by financial institutions extends into their entire value chain, which includes those of their customers and investees, in line with the 2011 Update of the OECD Guidelines for Multi-National Enterprises (MNE-GLs).

These comprehensive MNE-GLs are fully aligned with the UN 2011 Guiding Principles for Human Rights (UNGPs) and international labor standards. The "value chain" definition of a financial institution relates to all activities that such institutions carry out to create positive value for their customers or investees, their direct stakeholders and society at large, as well as those activities that mitigate or reduce any "adverse impacts" by such customers and investees which they have "caused," "contributed to," or are "directly linked" to, as defined by the OECD MNE-GLs.

Enhanced due diligence, effective leverage, and responsive public accountability, key factors in responsible business conduct (RBC), are not only a societal responsibility but also a business opportunity, because these should create value for all shareholders, customers, other stakeholders, and society at large. The 2007/8 financial crisis has illustrated the importance of the financial sector for the economy and society: its resilience, standards and focus matter to us all. Financial system resilience may be defined as the capacity to continuously perform its primary functions, in particular supporting the real economy and enhancing societal wellbeing.

The international developments in 2015 are offering strong momentum and opportunity also for business in general, and the financial sector in particular, to further articulate its role in creating stakeholder, as well as societal value. The financial sector may have a role as "key enabler" due to its role in the economy and its customers' funding operations. Important drivers in this context are the aspirations and commitments of the 2030 UN Global Goals for Sustainable Development, the Paris Agreement on Climate Change, the outcomes of the 3rd UN conference on Finance for Development in Addis Ababa, as well as UNEP Inquiry Report on the design of a Sustainable Financial System.

This article proposes some initiatives for the financial services industry (including banks, pension funds, insurance companies, private equity, impact investors, and philanthropic organizations). The financial services sector may give itself more credit, but also, clearly, assume more responsibility for values-based, authoritative, value-preserving and -creating leverage it has with its customers and investees for creating positive, true societal value. It deserves more trust from society for its efforts to learn from previous mistakes and working with higher values and better practices. For further improvement the sector may benefit from policy and regulatory support from governments and from constructive engagement and input from civil society organizations.

THE EMERGING CONTEXT: TRANSFORMATIVE MOMENTUM

Values-based scaling up/speeding up is becoming a business reality, particularly among large international corporations. The RBC/ESG/CSR agenda has evolved during the last 10 to 15 years from being merely a voluntary exercise undertaken by progressive practitioners through self-regulating codes/frameworks like the Equator Principles, UN Global Compact, PRI (Principles for Responsible Investment), GRI (Global Reporting Initiative), ISO26000, and platform-organizations like the World Business Council for Sustainable Development (WBCSD), UNEP-FI. We have now entered a new stage of development, with governments taking on more active and normative roles. A more shared agenda, co-created by governments, business, and civil society organizations, is setting the stage for soft law based on emerging good practice. Businesses, and they are many, that are members of organizations such as WBCSD and PRI are supporting such trend, as it responds to urgent environmental and social needs affecting their businesses, as well as facilitating the creation of an international level playing field in business.

The 2011 update of the MNE-GLs, as a baseline framework for corporate behavior, is a prime example of this, as there was active involvement of the business sector in the development of these guidelines. The update has been co-created and, hence, has a high degree of shared responsibility from BIAC (Business and Industry Advisory Committee to the OECD), TUAC (Trade Union Advisory Committee to the OECD), and OECD Watch (NGO platform to the OECD[2]). A number of sectoral guidance notes are being developed similarly on a multi-stakeholder basis building on these MNE-GLs. Another example is the 2013-launched GRI-G4 reporting framework on corporate disclosure, as it has become recognized (also by its multi-stakeholder governance structure and its due process of worldwide public consultation) in many countries as an authoritative standard for sustainability/non-financial reporting by business, and as such recognized in the MNE-GLs and the 2014 EU Non-Financial Reporting Directive.

2 The OECD Guidelines for Multinational Enterprises are being adhered to by the 34 state members of the OECD, but also adhered to by 12 non-members (including, for example, Brazil).

Setting standards and mainstreaming with government support is accelerating, and documents of sector initiatives are increasingly linked with the generally accepted standards (for example, by the Thun Group with the UN Guiding Principles on Business and Human Rights, and by recent initiatives by leading pension funds with respect to the MNE-GLs and Climate Change). It took GRI 15 years to evolve from a good idea, a framework to become a widely applied standard, which has clearly contributed to a general acceleration of the pace of RBC standards, practice, and disclosures. Such evolution in other RBC areas will intensify and accelerate and become mainstream, with the financial services sector becoming actively involved. The only missing stakeholder is still the end-consumer, but that will also change shortly.

The recent international policy agenda has created very promising confluence and momentum for all businesses (large and small; local and international) towards increased focus on value chains, public-private sector partnerships, RBC and looking "beyond GDP/financial capital." This momentum is occurring within the context of a shared ambition for more sustainable, inclusive, global economic growth, and societal stability. Through focus on "sustainabilizing" (for environmental issues) and stabilizing (for social issues), the value chains are rapidly becoming a shared business and societal interest. Recent 2015 milestones to note are:

- The G7 and G20 Summits, chaired by respectively Germany and Turkey.
- The UN Finance for Development Conference (FfD3) in Addis Ababa, with explicit reference to, inter alia, blended and infrastructural finance, new financing instruments for sustainable development, need for policy coherence;
- The adoption of the UN post-2015 (universal) Global Goals for sustainable development in New York, with 17 goals and 169 targets.
- The UNEP inquiry report on a "sustainable financial system," in particular focusing on the role of regulators/supervisors/standard-setters.
- The Paris Agreement on Climate Change, following the COP21 conference.

The moral underpinning of the agenda was further emphasized by the Papal Encyclical "Laudato si'" on, inter alia, the environment and on climate change.

In addition, the broadening of the interest to establish global responsible business conduct standards was illustrated by China's recently growing interest in learning from, and even associating itself with, the MNE-GLs, (at least for its international business), making the latter (in due time) even more globally recognized. The ProActive Agenda for the financial services sector under the MNE-GLs is work in progress. In 2016, the OECD will recalibrate official development assistance (ODA), as well as review the recognition of new (innovative) instruments as part of the "total official support for development" (TOSD), such as development-impact related partial risk guarantees. Some large philanthropic organizations are also increasingly focusing on the Global Goals, as was also recognized in Addis Ababa.

The Dutch policy and civil society agenda also includes some promising initiatives, such as:

- The start of the 2015-2017 Sector Covenant Process of 13 high societal impact sectors (among which is banking): although initiated by the Government, it is led by business with active civil society involvement.
- The preparation of an ambitious, broad-based Dutch 2030 Global Goals' strategy by a coalition of (80) progressive business and civil society organizations (led by Worldconnectors, DSM, True Price, with support from the Dutch Ministry for Foreign Affairs: Dutch Global Goals Charter Coalition) for partnerships in the Dutch national and international value chains.
- The legislation of the EU Directive on Non-Financial Reporting into national law.
- The Dutch EU presidency during the first half of 2016, with, inter alia, special focus on aid and trade.

This context offers the Dutch financial services sector an opportunity in 2016 to domestically lead, as well as reinforce its progressive role in international policy and encourage national and international value chain practices towards an ambitious, more responsive and responsible role of the financial institutions in society.

As is my credo: "Nothing is impossible, particularly when it is inevitable." This is an opportunity to lead with ambition, and by the financial services sector in particular.

CREATING VALUE BY AN AMBITIOUS, RESPONSIVE AND RESPONSIBLE FINANCIAL SERVICES SECTOR

Micro- and macro-prudential risk management is at the core of the financial services sector. In a smaller, increasingly dynamic, "flat," transparent and stormy world, values'-based forward and integrated thinking, practicing, and reporting should support such risk management by the financial institutions and their customers, investees, insureds. Risk management is about taking informed decisions and carefully considering all that we need to know: the ever-changing

context and the not so obvious or ignored "externalities" of today (which we do not account for, or offload on society). Risk management is also quite relevant in order to prudently time-match assets with liabilities, particularly when it concerns long term and illiquid exposures, which may become potentially (sometimes sooner than you expect) value-impaired or even stranded.

Strengthening in-house risk analysis and risk management capabilities is of the essence and should serve as an "insurance premium" against major liabilities and risk-losses. We seem to always have the wisdom to explain with the benefit of hindsight why a crisis, a failure or an incurred loss was foreseeable and even preventable, yet we often lack the foresight or the preparedness to take the collective, corrective early steps needed, notwithstanding hearing "the canaries in the coal-mine."

The need for integrated thinking is increasingly apparent, but also challenging for practitioners and assurance providers alike. It has four dimensions:

- **Forward thinking with advanced due diligence:** risk-based, societal context- and impact-sensitive, rights' compatible, actual and potential, initial and ongoing. It should explicitly consider, next to the traditional economic and financial capital, also natural, human, social, intellectual, and manufacturing capitals.
- **In the entire value chain (cradle-to-cradle):** a transition from "know your client" to "know your client and their value chain and emerging context."
- **Evaluating and balancing multi-stakeholder and societal interests:** impacts and benefits, including today's society and those for future generations.
- **"Materiality-based" public disclosure and integrated reporting thereon**: on the basis of the principle "report or explain why not."

Societal risk analysis and impact assessment is a new risk approach. Operating in the public domain is a complex and challenging task for business. As society is becoming more involved in business performance and its implications on others, a new risk category may be considered and organized: "societal risk and impact assessment." Such risks may be differentiated from more traditional risk categories, such as credit (including political) and operational (including reputation). This approach underlines the need for proactive, enhanced initial and ongoing due diligence on impacts that financial institutions directly or indirectly (through their value chains) have on others, including civil society at large. Knowledge of international law and standards and anticipating "the law of the future" are important new dimensions for businesses with international value chains.

Functions such as reputation management, public affairs, and issue management tend to be more defensive, inward looking, and ad hoc, while societal risk management is more dynamic, outward- and forward-looking, and strategic, supported by active societal engagement, learning, evaluating and capturing sector-, client- and transaction-specific issues. Leading financial institutions have recognized this. Early identification of potential adverse impacts on society, and value impairments or stranded assets in their own portfolios may be outcomes of such a structured risk-based scenario approach. The process of creating a materiality-matrix (as recommended under the GRI-G4 framework, and increasingly practiced) in consultation with internal and external stakeholders is quite valuable to match the interests of society with those of the corporation, and attach key performance indicators (KPIs) to the outcomes.

"Fueling the wheels of the economy" as a catalytic role for the financial services sector within a real and inclusive economy should become the new paradigm. By setting an example in its core values, comprehensive policies, responsible practices, accountability by its (materiality-based) integrated reporting and by using its individual and collective leverage towards more responsible, sustainable customers, it will become a driver for positive change.

Also relevant in this context is the fact that the "internalization rate" of unaccounted externalities in the costs and value creation of doing business (as these are, as yet, either unrecognized or, by default, off-loaded on society, but eventually and inevitably will have to be, by markets or regulations, absorbed) is accelerating by advanced businesses.

Sometimes a shock is needed to wake us all up and prepare us for a transformation: "from pain to gain." For example, the 2013 Rana Plaza accident has had a major impact on the entire garment and apparel sector. It has definitely accelerated the "internalization rate" momentum in this sector, with direct multi-stakeholder involvement across the value-chain. The severity of this accident, the circumstances around it, and the attention it received in the international public domain has demonstrated the risks of liabilities, the costs of conflicts and accidents, and the loss of reputation, throughout the value chain across the sector, even for those companies not directly linked to the accident.

The notion of "show me the money" (quick profits to the shareholder) is clearly graduating into a broader and longer term "show me the value" with a specific focus on generating long term value to all stakeholders without, at the very least, doing no harm to society, and preferably to do good. There is a wider trend towards identifying and measuring environmental (and social) externalities: UN-TEEB, WBCSD and the Natural Capital Coalition are among the

platforms and institutes that work in this field.[3] Also various accountants and consultants, including the Big 4, are contributing to this movement.[4]

True Price, a Dutch social enterprise, is among the leading international incubators working towards creating positive and negative measurements, and tools to even monetizing the environmental and social externalities by developing and testing methodologies for true pricing (for products), integrated P&L[5]/true value (for businesses), and true returns (for investment portfolios). It provides actionable insights into the value chains of companies. Pension funds are getting interested in assessing their value-creation for society. The MNE-GLs are used as the standards' baseline and take into account non-traditional forms of capital, such as natural, human, social, manufacturing, and intellectual capital, as well as financial. This is similar to the approach taken by the International Integrated Reporting Council (IIRC).[6]

The traditional, mostly legal boundaries of corporate responsibility and accountability are becoming increasingly blurred. This is reflected in the MNE-GLs, where certain responsibilities are recognized for a given company from not only the perspective of "causing" or "contributing" to adverse impacts, but also by being "directly linked to operations of a business relationship causing the adverse impacts." This notion defines an extended responsibility and accountability for impacts (and remediation in the case of "cause" or "contribute," the latter including "benefiting from"), requiring companies to deepen and broaden their due diligence. The term "being directly linked" is of particular relevance to the financial sector. In this context, the potential leverage a financial institution has is more important than just the financial interest associated with the given business relationship. Businesses are encouraged by the MNE-GLs to use their leverage within their value chain as a way of responsibly conducting business with suppliers, contractors, partners, and customers (including investors and investees). They may do this alone, in cooperation with partners, or within the sector. In the structuring of financing (and insurance) documents, addressing RBC-related issues need to be defined as precedent conditions, as well as ongoing.

The "materiality" question, relating to "what matters to whom," has become a major area of focus, especially since GRI-G4 was launched in 2013. Leading businesses are increasingly publishing their materiality matrix, linking and ranking in their annual reports the relevance of issues to stakeholders (and society).

The issue of client- and competition-sensitive confidentiality versus public accountability is one that also needs to be further explored, with a need for a greater emphasis on the "materiality to customers, to other stakeholders and society." Aggregate exposures to particular sectors, and policies and the performance therein, should definitely be regularly reported.

Reinventing structured finance may become a high priority for banks in redefining their core-business: "blended finance" in infrastructural – and/or corporate value chain – financing, inspired by the Global Goals, with the OECD Guidelines as the baseline, will offer opportunities for cooperation between arranging banks with asset-investing pension funds and impact-investors. Such financings may be complemented by independent feasibility studies and technical assistance financed by philantropies, possibly further enhanced by development-impact related credit/investment-guarantees from governments or multinational financial institutions. Such integrated, multi-stakeholder, and multi-faceted structures may make a major contribution towards realizing the Global Goals.

All of these external, societally-driven initiatives offer the financial services sector an opportunity to illustrate to their customers a broader perspective, offer practical advice, mobilize capital, and provide leadership towards a shared, long-term interest between the private sector and society at large. This requires a clear "signal from the top" and consistent engagement with stakeholders. Many Dutch companies (including financial institutions) are internationally prominent and recognized to be open to having such meaningful multi-stakeholder approaches on policies and transactions.

THE FINANCIAL SERVICES SECTOR MUST DO THE HEAVY LIFTING ITSELF

Internal culture matters: embedding core values, integrated thinking, and broad-based action by financial institutions requires awareness, creating the right internal culture, training, professional focus and self-confidence in order to engage with relevant stakeholders and society. By sharpening their own diagnostics and research, financial institutions will become more authoritative on relevant developments and issues, while their capital mobilization and advisory services are supporting long-term value creation and asset protection for customers and themselves.

3 See amongst others TEEB, 2008, "The economics of ecosystems and biodiversity interim report," and NCC, 2014, "Taking stock: existing initiatives and applications."

4 See amongst others: PwC, 2013, "Measuring and managing total impact: strengthening business decisions for business leaders," KPMG, 2014, "A new vision of value: connecting corporate and societal value creation," True Price, Deloitte, EY, PwC, 2014, "The business case for true pricing."

5 True Price (2015) Multidimensional P&L in brief.

6 IIRC, 2013, The International <IR> Framework.

The quality of available data – on the "real economy," industry developments, clients' value chains, and externalities – is a critical risk and success factor. Further, access to risk analytics, independent impact assessments, and (buy-side) rating methodologies are important factors going forward.

The financial services sector must improve its disclosure and external communications about its operating standards, policies, and related performance. Considering its commitment to the RBC mission, the Dutch financial sector has recently made significant progress, most notably through increased stakeholder engagement, policy development, active participation in (international) industry platforms, performance, and disclosure. Still, much more can be done, individually and, importantly, collectively. In addition, the financial services sector may, of course, maintain its broader societal engagement, which includes foundation support, and community services like financial education and empowerment.

The MNE-GLs are the foundation for the Dutch Sector Covenant Process and should be recognized as over-arching guidance for its own business principles and practices. The MNE-GLs are not only focused on the ambition of "sustainability" (foremost the environment, climate change, employment conditions, and social justice), but extend this by introducing the theme of "responsible business conduct" (RBC). This includes other important aspects such as disclosure, bribery, consumer protection, science and technology, competition, and taxation.

The key theme to be strengthened is the initial and on-going risk-based due diligence in the entire (supply, distribution, and service) value chain. In this context, it should consider focus on "activities directly linked to adverse impacts," the "materiality" of an issue for all stakeholders, as well as its potential impacts on, or benefits to, the society at large. It should also consider using its individual or collective "leverage" to address the impacts. Stress-testing exercises on its portfolios, clients, and transactions may be expanded beyond the traditional, purely economic approaches, to include societal developments (resource availability, loss of biodiversity, environmental degradation, social injustice, climate change, fundamental values). Much can be learned from progressive customers and investees. Benchmarking other, less advanced companies with such leaders offers useful insights, and can be used as input for sector policies and advice to others.

Making the financial institutions better fit for tomorrow requires:

- Fully integrating comprehensive ESG/RBC factors into the core risk and research analysis and approval processes, including in the KPIs and Risk Appetite Framework and Statement.
- Accelerating awareness, culture, and training for all "lines of defence" (customer-facing, risk-function, control and audit).
- Becoming the authoritative example/benchmark for public reporting (integrated reporting, including G4).
- Making use of emerging approaches/tools, such as entire value chain risk assessment and impact mapping, foot printing, circular economy principles, monetizing externalities, and creating pro forma integrated P&Ls.
- Taking the lead in the SDG and climate change agendas.

The term "impact-investing" may still be considered "a special asset class" or even an "oxymoron," but should, sooner rather than later, become a "mainstream standard," i.e., a "tautology." Furthermore, in an increasingly fast-paced, changing world, the risk of value-impairment and stranded assets is becoming more prominent, particularly for long-term investors.

Making financial markets better prepared for tomorrow's challenges and opportunities requires stimulating and assisting corporate customers to significantly improve their risk analysis and disclosures, and improve their public disclosure by embracing the EU Directive on non-financial reporting (applicable to any company with more than 500 employees). Organized trainings, workshops, and advisory services may also be of benefit. Making markets much better fit for purpose recognizes that markets are not perfect, and prices (and valuations) are often not right. Hence, we must raise, together with regulators, supervisors and stock exchanges, the quality of disclosure/ transparency by the sector itself and their customers and investees, with Pillar 3 of Basel III/Solvency II market-disclosure at the center. In addition, migration from combined reporting to materiality-based integrated reporting (including using GRI-G4) should be strived for. It should be recognized that not reporting the impact of potential material issues on the company itself as well as on those affected by the company (other stakeholders and the society) is a disservice to efficient markets and may even become a legal liability.

Access to remedy for affected stakeholders (Prof. John Ruggie's third Pillar, as well as "state duty to protect" and "business duty to respect") is also an important factor in an effective "social compact" of business in specific transactions, at both the corporate or sectoral level. The National Contact Points (NCP's) under the OECD MNE-GL's in the Netherlands, which is independent but operating under the political responsibility of the Minister for Aid and Trade, are also demonstrating the benefit of such mechanisms by accelerating mediated solutions and reducing the cost of conflicts.

THE FINANCIAL SERVICES SECTOR CANNOT DO IT ALONE: THE ROLE OF THE GOVERNMENT

A key contribution of the government is to create a coherent and enabling environment, making the financial services sector and financial markets systemically more resilient. A coherent, effective control environment needs to be created by regulatory and supervisory interventions (regulations, policies, guidance), which is fully aligned with societal priorities. Sector codes, regulations, and policies through organizations such as OECD, UN, BIS, EU, ECB, as well as national governments (for example, in the Netherlands we have the Dutch government, DNB, SER-ICSR (IMVO)), are important for providing more coherent regulations and policy guidance, as well as explicit recognition and capital charge incentives for structured approaches to value chain risk analyses.

"Sustainabilizing" and (socially) stabilizing value chains for defensive portfolio risk purposes may be practiced, while at the same time realizing the potential for strategic and practical (sustainable) value chain advisory services by financial institutions (banks in particular) to their customers. Such an approach may also directly strengthen their earnings capacity. New regulations and policies should not unduly increase the burden of internal controls or supervisory oversight. The materiality concept is also of the essence here.

Current regulations and codes are not sufficiently compatible, and in certain circumstances even counter-productive, to the emerging RBC/ESG agenda. Anti-trust laws, due to its traditional focus on consumer-protection, sometimes seem to get in the way, without offering space for collective action by business on sustainability. Experts from financial institutions should, perhaps at the behest of supervisors, prepare a white paper with an overview/evaluation of the specific requirements and processes that may become obstacles. They may make recommendations as to how regulations and codes may support the RBC agenda, and how a "bonus/malus" approach may be considered on risk weighting/capital charges in order to facilitate pricing (des-)incentives for customers. Perhaps this might be an opportunity for "Basel V."

Consultation within the financial sector on ESG/RBC directions and issues may be intensified through the sector organizations, with the governments and/or supervisors (such as DNB, the Dutch Central Bank, and AFM, the Netherlands Authority for the Financial Markets, in the case of the Netherlands) serving in a convening and stimulating role. This approach would be similar to the process on bank transparency that was hosted in 2014 by the Dutch Ministry of Finance and which has now been adopted for further action by the NVB (the Dutch Bankers' Association). This may be done in a structured way, say twice a year, with focus on society-strategic issues that could be related to the 2015/2016 international agenda mentioned above. The UNEP Inquiry Report is definitely contributing to this process.

The new agenda, with focus on "stewardship," "duty of care," and "universal ownership" of the financial services sector also has profound consequences for a modern, effective corporate governance framework, including but not limited to values, purpose, ownership, corporate boundaries, the board role and composition. It is recommended that "impacts on society in the entire value chain" of a corporation is better addressed in this debate and in the corporate boards, as adverse impacts on civil society (including nature) have become more relevant. The 2015 Principles of Corporate Governance, as adopted by the G20 leaders in Antalya (Turkey), unfortunately do not sufficiently recognize this.

THE FINANCIAL SERVICES SECTOR CANNOT DO IT ALONE: THE ROLE OF SOCIETY

Media, civil society organizations (including trade unions, NGOs), and academia have an important role in challenging, as well as supporting, the financial services sector to fulfil its roles. Risk management is all about taking medium- and long-term, informed, decisions (i.e., beyond a regulatory one-year horizon for the probability of default). It is important "to know what you ought to know," but actually do not. So the question is, how can we know such issues? Who can assist? Many societal issues, like inclusiveness, environmental degradation, loss of biodiversity, climate change, and social justice, reflect processes of creeping insolvency until a major disaster acts as a wake-up call. Early, effective preventive or corrective action is morally required and business-wise prudent.

Knowing and showing: an increasing number of civil society organizations recognize that the issues at stake are too important and urgent to address not to work together with business. Some of them are taking, next to their traditional advocacy role, a more strategic approach. This may also be attributable to their accredited role, since 2011, in the MNE-GLs; moving from a single issue challenger solely focused on "naming and shaming" to a new, constructive one on the basis of "knowing and showing." This implies raising an issue to be resolved, but also for putting it in a broader, even systemic, context for balanced solutions.

Do not disengage: running away from an adverse impact may make matters worse for the affected people, but is quite a challenge as public opinion may still be highly critical; intensifying dialogues and

cooperation with civil society organizations may help in this context. Companies should not disengage from the issues or impacts, but rather endeavor to address and resolve them, as is also stated in the OECD MNE-GLs. By doing so, the financial services sector will become more credible and trusted, innovate within their business and reduce the costs of conflict. While this will require significant commitment, resources, and stamina, it is strategic and even "existential." Structured dialogues by the financial services sector with civil society organizations and members of parliament to address issues and initiatives from both within the sector and civil society organizations are important. Education and ongoing training on RBC issues should be significantly strengthened and scaled up. Universities and other knowledge centers need to allocate more resources for integrated thinking and practices across disciplines, for education and research.

SOME RECOMMENDATIONS

Strengthen due diligence and leverage: financial institutions have the opportunity and responsibility to use their due diligence and leverage, whether alone or collectively, to stimulate more responsible business practices in their entire value chain, including the value chains of their business relationships, notably clients and investees. Their added value includes anticipating and balancing risks, and supporting value creation for them as well as for all relevant stakeholders and society at large. Due diligence of the value chain is at the center, and it is not just a defensive approach. Next to risk management, it is also becoming a tool for identifying business opportunities, not only for the financial institution itself, but also for their (small, medium and large) business customers and investees.

Adopt the MNE-GLs as a baseline standard: by adopting the OECD MNE-GLs as a baseline guidance in its customer-facing businesses (including capital mobilization and advisory services) the sector would be instrumental in the building of a broad-based, multi-stakeholder environment (as also aspired in the Global Goals) for responsible business conduct; thereby also contributing to a level playing field for itself, its customers, and investees.

Focus on materiality: prudent, effective, and efficient risk-management should focus on the materiality of issues to the company's stakeholders and society. Public disclosure should also take this approach. Regulators and assurance providers should consider this too as regulatory – and data – accumulation in financial and non-financial reporting is becoming costly while not serving its purpose of being meaningful to interest-holders. Use the internal and external stakeholder materiality-matrix process as a valuable platform for identification and prioritization of key issues and attach KPIs thereto.

Show ambition and courage to lead: the 2015/2016 national and international policy agendas offer a great opportunity to put the ambition of a revitalized, innovative, responsive, and responsible financial sector into practice and in the public domain. Financing is at the center of realizing sustainable development, as reflected in the Finance for Development Summit (FfD3).

Integrate new approaches, business models, processes and tools: in a world that is increasingly dynamic, inclusive, transparent, rights' sensitive, and resource constrained, it is prudent practice to improve initial and ongoing due diligence with clients, investments, and other transactions. Further integration of responsible business conduct into the main risk management and business generating processes has become an imperative for resilience and success. This integration also offers the space for innovation, such as the development of new tools, as well as business models and practices.

Read the emerging context: frequent engagement on societal and strategic issues with clients, civil society organizations, governments, supervisors, and academics may be intensified. Within the Netherlands, the recently started 2015-2017 Sector Covenant Process, initiated by the government, offers a platform for this in 13 parallel sectoral processes, including the financial services sector itself with three sub-trajectories: banks, pension funds, insurance companies. Some involvement of the financial services sector in the other sector processes would create value for the other sectors as well.

Create an enabling policy framework: governments and supervisors may provide more coherent, forward-looking policies and regulatory frameworks, which are conducive to a proactive financial sector that meets not only stakeholders' interests but also general society's priorities. A review of the prevailing regulations and policies, as part of the UNEP Inquiry Report, may be made jointly with the sector against this perspective. Define KPIs and report on the performance, including dilemma's faced.

Recognize that RBC is a journey, and you cannot walk alone: civil society organizations are encouraged to recognize RBC by business as a journey that requires trust (or at least "benefit of the doubt"), time, cooperation, balanced solutions (of which such organizations sometimes do not agree amongst themselves), and accountability. Such organizations should be prepared to constructively cooperate with the financial services sector on key issues on the basis of "knowing and showing"; recognizing that legacy issues and incidental "pain" often offer an opportunity for systemic "gain"

for everyone. In addition, universities are encouraged to more actively include the RBC perspectives and issues discussed in this paper into their curriculums.

Intensify the non-competitive dialogue and collective initiatives: it is recommended that the internal financial services sector dialogue intensifies: non-competitive exchanging of views, cooperation, and streamlining are warranted in order to further prioritize and pool resources. Anti-trust regulators need to recognize that such joint sector initiatives are intended to be created in the interest of the public good. Competition should be on the basis of client proximity, execution capability, and pricing, not on the basis of compromising fundamental individual and/or collectively agreed values.

Be honest in your internal and external communications: in today's world of data-technology and (social) media, increasingly, public perceptions are driving realities affecting reputations, brands and trust. There is no time or space to hide anymore: transparency, unknowns, dilemmas, mistakes are part of reality, so accept this reality and actively listen, learn, and address with honesty and confidence.

Culture matters: perhaps the most important factor within the financial sector for rebuilding external trust and internal self-confidence and pride is by performing the right business for the right purpose: this must come from within!

CLOSING COMMENTS

The year 2015 has shown us the enormous impacts and costs of conflicts and poverty, causing an emerging social crisis with large flows of refugees and migrants (in line with my earlier view on a "forthcoming major crisis of October 2015"), but, fortunately it also demonstrated the ability of the international community to set goals and make commitments to address some of the root causes. The period 2016-2020 is critical to convert the expectations, promises, and commitments into tangible results. This cannot be done without business, with a special role for the financial services sector: "noblesse oblige."

Robo-advice 2.0: The Next Generation

Andrew Arwas – Managing Principal, Capco
Katie Soleil – Senior Consultant, Capco

Abstract

No new trend or development in the world of wealth management could have attracted more discussion, printed words, or conference speeches than Robo-advice. Hailed as the new era in providing investment advice on a large scale to the emerging, digitally savvy generation, Robo-advice has spawned new names in this most traditional of markets. Betterment, Wealthfront, Nutmeg, and others have emerged in recent years offering simple, engaging user interfaces, automated advice via algorithms and low fees. But already we are starting to consider how the next generation of such services will develop as the market matures. What is clear is that successful Robo-advice 2.0 services will focus not on the technology, but on the underlying investor and their very human needs. The key to successful proposition development continues to be a clear focus on the needs of the consumer, not on the clever technology. We believe this focus must be on three critical elements: an offer which is meaningful and relevant at all stages of the customer lifecycle, a cost structure that is transparent and low relative to traditional advice services, and genuine simplicity (in language and process) throughout. If Robo-advice is to transform into a truly mainstream and global form of digital advice, Robo-advice 2.0 must create personal, relevant, and differentiated services for real people.

As assets under management grow, robo-advice is proving more than just a flash in the pan in a bull market. How can established financial institutions successfully develop a next generation robo-advice offer? The answer lies in leveraging their long-term strengths to create personal, relevant, and differentiated services for real people. Robots may help deliver robo-advice 2.0. But it cannot be designed by them, or for them.

THE PROPOSITION

Established financial services providers have the opportunity to retain existing customers – and to acquire new ones in new market segments – by "logging on" to the latest form of digital wealth management: "Robo-advice 2.0." The term "Robo-advice" has become a catch-all term applied to investment services using online platforms to reduce – or even eliminate – the need for human advice. Yet, robo-advice is not universally impersonal. While some types of approach do include the purely formulaic, as well as forms of low cost and simple initial product selection, others also extend, in some cases, to greater depth of personal, tailored advice.

It is important to distinguish between the customer proposition and the technology when discussing robo-advice. Often, the quick, simple, and low cost investment solutions on offer are easy to buy conceptually while the technology makes the buying process accessible. But successful development of robo-advice 2.0 will **not** be based on extra technology bells and whistles. At its core will be personal, relevant, and differentiated services for real people. The technology is the medium. It is not in itself the message.

The robo-advice combination of lower fees (for an offer that typically includes rebalancing the portfolio) with engaging interfaces, appeals to people who want technology-enabled, self-directed, and affordable investing. Established players should not, however, assume the appeal is **exclusively** to the novice investor; or that it **always** culminates in simple choices, based on automated ETF selection. The offer has started to reach a wider range of age groups and levels of investment experience.[1]

OPPORTUNITY OR THREAT?

From the established brands' viewpoint, market uptake of robo-advice can be seen either as a passing fad, as a pure threat to be resisted tactically, or as a significant opportunity.

The rapid growth in the value of assets under management (AUM) suggests that, while robo-advice is a long way from cracking the market wide open, it is having a noticeable impact. This growing appeal can also be seen in the context of loyalty disintegration among millennials.[2] Its increasing uptake underlines one well-established consumer attitude: basic dissatisfaction with banks.

Robo-advice is not a threat to be countered just by putting more resources into marketing traditional investment advisory services. It cannot be resisted solely by "counter-attacks" based either on exploiting fears of smaller brands' longer term market vulnerability or by dismissing robo-advice as purely formulaic, or even inadequate. Where wealth management offerings today provide a poor client experience and/or poor value for money, the growth of robo-advice will undoubtedly present a new threat. But there is no reason to believe it will bring down the whole industry as we know it.

Few industry players today could fail to recognize both that robo-advice is a direct reflection of powerful forces across society as a whole, and that it is a major opportunity. As a result, growing numbers of established financial institutions are exploiting the opportunity to create and market the next phase of robo-advice.[3]

WHY DOES ROBO-ADVICE APPEAL?

Robo-advice grew because it is a real solution to the needs of real customers. Today, a number of key factors are helping to drive its ongoing uptake. Legislative changes to advice distribution frameworks (such as the Retail Distribution Review in the U.K.) left many potential investors unable – or unwilling – to afford up-front, conventional advice fees. Robo-advice is an accessible and affordable option for "advice gap" customers, who realize a range of life events – property purchase, education and health care funding, prolonged retirement – demand some investment for the future.

1 According to a report by Numis Securities, robo-advisors were unlikely to substantially affect the £1.1 trillion U.K. wealth management market in the next five years. However, Numis pointed out that Hargreaves Lansdown, once an upstart challenger, has risen over three decades to command significant market share in its sector; Evans, J., 2015, "Slow march of the robots," Financial Times, June 22, http://goo.gl/aKJeiZ

2 The 2015 Makovsky Wall Street Reputation Study indicates that some 49% of millennials said they would consider using financial services options from companies such as Google or Apple.

3 Many well-known brands — including Investec Wealth, Brewin Dolphin, and Barclays Stockbrokers — plan to launch robo-services in 2016, while Hargreaves Lansdown, the biggest direct-to-consumer fund shop, launched its own Portfolio+ service in July 2015 [Evans (2015)].

Proliferating device types and internet access mean that selecting financial services online is no longer restricted. Nor is it embraced exclusively by the so-called "millennials" either. Customers of all ages respond to online experiences that replace complex and expensive processes with clarity and perceived value for money.

The factors referenced above will continue to drive and shape a robo-advice market where challengers are prominent. However, they have strong potential to create opportunity for established service providers too.

Robo-advice shows that new levels of customer responsiveness – not the dominant style of many traditional offers – **are** possible, both in terms of proposition and delivery technology. The robo-advice delivery model offers older financial institutions the chance to reach existing customers with new, more personal and more responsive offers. Importantly, they can also engage with entire new market segments. And they can do all this in an operational context of very significant scalability and economies of scale.

Achieving comparable levels of scale and instant accessibility through traditional advice models – based on teams of highly qualified, but relatively scarce, advisors – would be a practical impossibility. However, cutting edge technology and efficient processes are not a proposition in themselves. Content – those real solutions for real people – has to be king.

HOW HAS ROBO-ADVICE IMPACTED TRADITIONAL SERVICE DELIVERY?

Technology-driven economies enable the transition to lower or zero fees. In addition, many current robo-advice propositions accept smaller initial investments. The minimum deposit can be just a few euros, in the case of some special offers.

The changes mean investing is no longer exclusive to high-net-worth (HNW) or ultra high-net-worth (UHNW) profiles. The robo-advice concept has "democratized" the market, appealing to younger potential investors with high incomes, and to older individuals who feel comfortable with technology and want to explore investing in a straightforward, jargon-free environment.

It is true that these developments are enabled by cost-effective delivery mechanisms, but some robo-advice service providers have gone further. They have challenged the most basic premises of financial advice, with radical **propositions** that sit above pure technology issues.

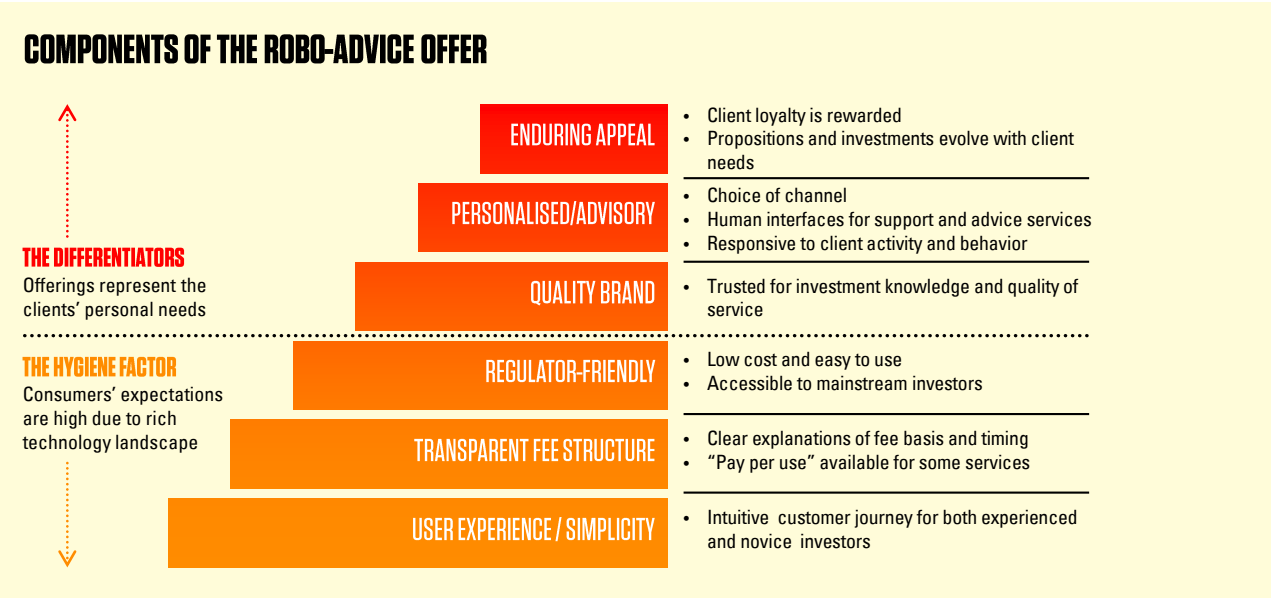


Figure 1 – Components of the robo-advice offer

For example, one is offering investment solutions on a "menu basis" where only the more complex "a la carte" services attract fees.[4] Another proposition[5] enables investors to advocate their own investment choices online. They can then earn rewards when other platform users choose their selection. In effect, an internal market is being created where individuals can become quasi investment advisors themselves.

These potentially disruptive propositions are not **intrinsic** to any core robo-advice fixed model. Their fundamental innovation value lies in their conceptual approach and their content. They are surely worth consideration from established players looking to develop robo-advice 2.0. They show that delivery technology does not have to exclude innovative and highly differentiated proposition development.

Some industry professionals have dismissed the more disruptive innovations as short-lived gimmicks. However, new thinking, stimulated by the broader possibilities of robo-advice, **has** led to real game changers for investment advice channels.

One robo-advice proposition and platform[6] has been adopted by a fund manager, who is rolling out an associated application for use by a distribution channel of 3000 financial advisors. Another is partnered with one of the biggest fund banks in Germany,[7] managing more than 600,000 customer accounts with deposits of more than €14 billion.

ROBO-ADVICE MARKET ENTRY – ADVANTAGE THE ESTABLISHED FINANCIAL INSTITUTIONS

Established financial institutions enjoy a number of powerful advantages when it comes to successful robo-advice market entry. They can leverage their greater spread of financial product offers and expertise to up-sell and cross-sell; enhancing the customer experience through a richer next generation robo-advice proposition **and** exploiting their own expertise. They can use their highly developed distribution channel networks to reach the customer base efficiently and effectively. And they have the immense leader advantage of trusted brand and market prominence already in place.

First phase competitors have done much to help the established brands. Pure-play robo-advice start-ups have done a lot of the heavy lifting – defining the initial wave of propositions, advice styles, technology platforms, target demographics, even fee structures and business models. This is an inversion of the more usual situation, where challengers cherry pick markets that long-term players have committed decades and billions to creating and developing. In fact, most of the important asymmetries between newcomers and established brands work in the latter's favor, as we highlight in four key areas below.

As far as **commercial security** is concerned, established financial institutions can resource their robo-advice offer from their own capital base. This contrasts strongly with the investment-heavy operational demands of the start-ups. Many of them have still to trade their way into profit, even if they have achieved market profile.

Trust factors provide a potentially powerful attribute. Being around for decades, or even centuries, means you do not have to commit a fortune to building a brand profile. Relationship capital is as important as the financial sort and a familiar brand commands a major potential advantage through its longevity. Banks may not always be popular with consumers, but they are acknowledged to have considerable long-term experience, when it comes to being trusted with funds. They can also justifiably claim to have survived multiple cycles in the market – downs as well as ups.

Longevity also brings deep experience of **regulatory issues** around financial advice and its compliance status. The exact regulatory position of the advice component continues to be a moving target. This will become a prominent issue in a market downturn. Some investors will inevitably challenge the appropriateness – even the fundamental status – of the advice and choices offered by their robo-advice platform.

While there are advantages to being a startup, the vast majority of new robo-advice offers are tiny in comparison to the mainstream operators. Agility is a useful attribute. However, it is more useful to be able to leverage pre-existing **operational scale** and handle large customer volumes through proven front-to-back systems.

4 WiseBanyan. https://wisebanyan.com/

5 Motif enables the user to pick their own "basket" of stocks or be inspired by others. It offers ideas-based investments, and blends advocacy with social media by allowing users to earn royalties when their "motifs" are purchased.

6 Betterment has penetrated the "traditional" market, offering services to the fund manager Fidelity. Fidelity in turn offering these services to their 3,000 financial advisors, who can provide them as an app to their clients.

7 Vaamo is partnered with one of the biggest fund banks in Germany, FFB (FIL Bank Ltd), which manages more than 600,000 customer accounts with deposits of more than €14 billion.

SUCCESSFULLY DEFINING ROBO-ADVICE 2.0 – HOW WILL IT LOOK?

The true potential of this market will not be realized by **replicating** first generation offers. There are lessons to be learned and missed opportunities to be seized. A successful next generation robo-advice proposition will build on trust values and relationship capital. However, it will explicitly signal this is not the "same old, same old" dressed up in an attractive design. Customers must discern personal aspirations-based propositions not pre-conditioned by an institutional agenda.

Each customer's sense of ownership of their investment selection will prove a strong loyalty factor, even when markets become more difficult. The adept established brand will nurture the dual sense of "**my** investment journey," which is "supported by **my** financial institution." It will be very difficult for the start-ups to sustain comparable relationship strength.

Of course, one of the subtlest challenges for established brands is to pair personal, aspirations-led investment choices comfortably with a sales imperative. They should not fight the fact that a robo-advice platform is there to sell investments. Nor should they neglect up- and cross-selling opportunities – especially since their likely first target will be their existing customer base. To achieve the right balance, highly considered and creative thinking will be needed to define and develop optimum customer experiences.

Ultimately, how will the best robo-advice platforms – truly engaging customer propositions enabled by accessible delivery mechanisms – crack the challenge of excellent customer experience? By balancing perfectly understood personal aspiration with appropriate choices – including some from the provider brand's broader portfolio. This will require an expert blend, with real integrity towards the customer's financial needs and institutional business development sitting comfortably together. (Specific examples might include automated contribution links from salary increases and bonus payments to SIPP contributions.)

All the guiding theory will culminate in a "good platform environment;" defined as making customers feel they are in the driving seat, rather than being passive recipients of pre-conditioned advice. The online information gathering approach will create the sense that they are setting their own unique advice coordinates. The "robo" element will then react to those coordinates, with the algorithm-driven investing process reinforcing the fact that the user is getting exactly what they asked for. No hidden or contrary agendas. This sense of empowerment will be a key selling point over traditional advice.

AND HOW WILL ESTABLISHED INSTITUTIONS POSITION THEIR ROBO-ADVICE?

Most likely by incorporating the new approach within a portfolio of existing solutions. They will not suddenly abandon other proven customer service and product channels. Some customers will immediately see robo-advice as the best way for them to buy certain investment solutions at key life stages. So they will want a packaged solution from a robot. At other times, they will prefer a more personal and advised conversation.

Flexibility of choice is one of the great advantages the large, established players enjoy. They can offer a true **range** of advice, while the new entrants are far more restricted – or even one-dimensional – in their capability.

Any market "mass migration" to robo-advice is improbable, however. It is well understood that customers have different attitudes to risk for different investment purposes. By extension, we can expect them to also display a variety of buying preferences for a range of investment needs.

WHAT ARE THE NEXT STEPS ON THE ROBO-ADVICE 2.0 DEVELOPMENT JOURNEY?

Long term players need to take a view on their commercial horizons, establishing the scale of market share and revenue opportunity that robo-advice presents. They must capture a clear picture of those existing customers most likely to use robo-advice, as well as thinking whether it is advantageous to attract new and different customers. Then, at the operational level, key areas to clarify early on include compliance, marketing, technology, and access to expert advice.

THE IDEAL ROBO-ADVICE 2.0 DESTINATION – REAL PROPOSITIONS FOR REAL PEOPLE

Today's entry-level robo-advice formula remains primarily transactional. It enables investment **purchase** quickly, easily, and simply. But many consumers (perhaps most) do not buy investments as an end in themselves. They are motivated by specific, life-driven goals. In reality, they want to save towards a house deposit, or to fund school and college fees, or to give a grandchild a sound financial start.

ROBO-ADVICE CHECK LIST

1 BUSINESS STRATEGY

Scale of robo-advice opportunity for your business

- Is this going to grow my business significantly?
- Is not providing a robo-advice offer going to damage perceptions of my business?
- Am I going to lose significant customer numbers without this offer?

Target customer profile

- Which of my existing customers are most likely to use this offer?
- Which customer sectors am I targeting specifically?
- How do I build a robo-advice offer that engages while remaining "on brand"?

Strategy

- Where does a robo-advice offer fit within my overall strategy?

2 GOVERNANCE

- Do I have a crystal clear understanding of the current legal/compliance status of robo-advice?
- Am I aware of all potential compliance implications of offering this type of advice?
- Do I have a complete view of emerging trends, indications and likely future regulatory developments?

3 TECHNOLOGY

- Do I have the necessary technology resources in-house to design, develop, and deliver a robo-advice platform quickly and cost-effectively?
- Do I have a good understanding of the robo-advice market from a tech perspective?
- Do I have a clear picture of the tech challenges of integrating a robo-advice offer into my existing technology ecosystem?
- Do I have access to current best practice and best-of-breed technologies and solutions?

4 MARKETING

- How will I position my robo-advice offer relative to my core brand and other elements of my product and service portfolio?
- Do I have a clear idea of the most effective channels for marketing my new robo-advice offer?
- Do I understand the relationship and key intersections between the technology rollout of the robo-advice platform and marketing milestones?

Figure 2 – Robo-advice check list

For established financial institutions developing their own robo-advice 2.0, there is one absolute imperative: the advice must be truly differentiated, with personal appeal to real people. This crucial "human dimension" must be embedded in every aspect, including online packaging, promotional activity, and – above all – the individual customer journeys that begin with the core investment product offering.

Offers reflecting real personal needs will stand out in a crowd of well-designed, but ultimately impersonal, websites. Underpinning clear differentiation, there are three critical areas to get right at all stages, from planning to implementation: the offer, the cost structure, and the user experience.

- **The offer:** the entry-level portfolio(s) available to investors comprise typically of ETFs or similarly simple products. If multiple portfolio options are available, greater guidance will be required to explain the suitability of those options and to make it crystal clear which customer profiles they are intended for. While much attention focuses on the initial onboarding – the selection of a suitable investment solution for the client – as the robo-advice model matures, we will see the ongoing experience becoming a greater area of focus. This development poses a key question. How will the initial – and hopefully positive – client experience be maintained over subsequent years, thus ensuring the offer remains relevant and valuable?
- **Cost structure:** there must be a transparent fee structure, clearly explaining when fees will be charged and on what basis. With multiple services, the difference between service levels and their associated costs should be evident.
- **Simplicity:** this demands total clarity of language and terms. Many brands will need to start presenting information in a much more straightforward way.

CONCLUSION – ROBO-ADVICE IS NOT GOING AWAY

Robo-advice has the potential to change the investment advice channel choices even of experienced investors. It will shift attitudes towards fees charged, including for premium investing advice. It will scrutinize and challenge the elements of professional advice that go into active management of investments. It could even, as we have seen, empower an entirely new category of investors, who make the ultimate journey from advised to "advocate/advisor".

In short, robo-advice will not leave the investment landscape remotely in the same shape that it found it. In this context of profound change, an established financial institution's final decision might be to either fully embrace robo-advice, to wait for the time being or to leave it to one side permanently. For those that choose to adopt the robo-advice model, the over-arching priority is to focus on the **proposition** – the fulfilment of a client need – rather than the clever technology. In this sense, **ensuring customers and their needs are uppermost at all times**, it represents a challenge as old as the established banks themselves.

The next generation of providers will need to ensure that robo-advice is as personal, differentiated and, paradoxically, **un-robotic** as possible. Robo-advice 2.0 may be enabled by robots. It cannot be designed by them; still less **for** them.

Regulatory

Economists' Hubris – The Case of Business Ethics in Financial Services

The Dodd-Frank Act Five Years Later: Are We More Stable?

The Volcker Rule as Structural Law: Implications for Cost-Benefit Analysis and Administrative Law

A Historical Perspective on the Different Origins of U.S. Financial Market Regulators

Economists' Hubris – The Case of Business Ethics In Financial Services

Shahin Shojai – Global Head, Capco Institute[1]

Abstract

This is the sixth article in the Economists' hubris paper series, which aims to critically examine the practical applications of academic thinking. The focus of this article is business ethics, with a specific focus on the financial services industry. The main challenges that one faces in determining whether businesses do in fact act unethically, intentionally or otherwise, are that there are no universally agreed parameters for describing ethical behavior; that ethicality seems to be in the eye of the beholder; and that since we are relying solely on external data, and do not have access to the thinking processes that lead to different business decisions, we are unable to state categorically that the management knew ex-post that a given decision would result in an unethical outcome. Given these difficulties, this article suggests that firstly, while most businesses don't necessarily set out to act unethically, when ethics and profitability collide the latter seems to win most of the time and secondly, that should companies decide to, or inadvertently, act unethically they have learned from the actions of Western governments how to manage the ramifications. The increasing influence that businesses now have over those that monitor them, including governments and the media, could potentially lead to corporations becoming less concerned about the ethical ramifications of their actions and consequently result in the concept of business ethics becoming even less viable from a practical perspective.

1 The views expressed in this article are those of the author and do not in any way represent those of Capco Institute, Capco, FIS, or any of their affiliate companies or clients.

INTRODUCTION

The recent global financial crisis, which somehow never seems to end, has brought the issue of business ethics to the fore once again. Many are asking why the banks behaved the way they did in the run up to the crisis and why they were allowed to simply pay financial penalties without having to admit any wrongdoing. More importantly, as Ben Bernanke, the Chairman of the Federal Reserve at the time of the crisis and one of the people credited with saving the banking system, recently asked [Page (2015)], why weren't there more prosecutions of the executives at these financial institutions? Executives whose actions prior to the crash he judged to be "bad business and immoral."[2] It is a fair question. Why were they behaving that way, and why were they able to avoid having to accept they behaved wrongly? Of course, the people who are asking these questions now are either too young to know or have simply forgotten how the banks behaved during the Internet boom of the late 1990s.

Those of us who followed the endeavors of Elliot Spitzer remember vividly the types of emails he was able to uncover about what investment analysts really thought about some of the stocks they were issuing buy recommendations on, which their colleagues were pushing onto clients, institutional or otherwise. The famous clarification of what Henry Blodget meant by PoS will forever be etched on the minds of those of us who knew that the Internet bubble of the late 1990s was nothing but that [Cassidy (2003)]. But, the issue is that even during those investigations most financial institutions simply paid their penalties and neither accepted nor rejected any wrongdoing. They just paid a fine and moved on.

Of course, some of the more recent issues that financial services firms have faced, such as the LIBOR-fixing scandal or money laundering, have resulted in some accepting criminal behavior, and it would be interesting to see what impact they might have with regards to U.S.-style class action suits by investors. But, by and large the so-called too-big-to-fail institutions that perpetrated these deeds have remained intact and their share prices seem to go up with every penalty paid.[3]

In response to the recent crisis, the public fury at the use of taxpayer funds to bail out a number of these institutions, the never-ending series of wrongful behaviors by the banks and the need for the governments to be seen to be doing something, a number or initiatives were undertaken. These ranged from ring-fencing investment banking activities away from the retail and commercial banking activities of banks (as suggested by The Independent Commission on Banking: The Vickers Report),[4] to the myriad of regulations, which are just too long to mention here, that were introduced by the Dodd-Frank Wall Street Reform and Consumer Protection Act,[5] to the limits on bonuses that were introduced by the European regulators and, of course, MiFID II and Solvency II. The list and the requirements of the new regulations introduced are extensive and complicated, and are beyond the scope of this article.

However, one of the responses of the U.K. regulators is of import to this article: the issue of "risk culture." Hector Sants, the CEO of the Financial Services Authority (FSA) between 2007 and 2012, made a number of speeches about the importance of culture within financial services firms and how steps needed to be taken by regulators to ensure that unacceptable cultures within firms are identified [Sants (2010a, b)]. Sants (2010a) stated that: "Historically regulators have avoided judging culture and behavior as it has been seen as too judgmental a role to play. However, given the issues we continue to see over time, I believe this one-dimensional approach has to be questioned. Every other aspect of the regulatory framework is under scrutiny and we should not shy away from debating the culture question."

Since Sants' speeches, many have started looking at the topic of risk culture and how to implement the guidelines that the FSA, now the Financial Conduct Authority (FCA), and the Prudential Regulatory Authority (PRA), as well as the Financial Stability Board (FSB)[6] have set for these firms. While the number of academic studies in this space is still quite small, with the most comprehensive so far being Power et al. (2013) and Jackson (2014), most consulting firms have published numerous reports on the topic and advised how firms should go about implementing the guidelines set by the FCA.

Clive Adamson, Director of Supervision at the FCA, stated at the CFA Society's U.K. Professionalism Conference in London that the FCA's approach to assessing culture is "to draw conclusions about culture from what we observe about a firm – in other words, joining the dots rather than assessing culture directly. This can be through a range of different measures such as how a firm responds to, and deals with, regulatory issues; what customers are actually

2 Please refer to this HARDtalk twitter link for Ben Bernanke's comments on Wall Street bankers: https://goo.gl/tWf29C

3 According to Reuters, "Twenty of the world's biggest banks have paid more than U.S.$235 billion (150 billion pounds) in fines and compensation," since 2008. Reuters: http://goo.gl/dEhl1g

4 The full text of Vicker's report is available from the website of the Library of the House of Commons via this link: http://goo.gl/OhFjfE

5 The full text of the Dodd-Frank Wall Street Reform and Consumer Protection Act can be obtained via this link from the website of the Securities and Exchange Commission: https://goo.gl/K9FGqY

6 Please refer to this FSB link for their framework for assessing risk culture: http://goo.gl/5G1gDE; also the following link for how serious risk culture failing will be addressed by the PRA: http://goo.gl/o1adZn

experiencing when they buy a product or service from front-line staff; how a firm runs its product approval process and the considerations around these; the manner in which decisions are made or escalated; the behavior of that firm on certain markets; and even the remuneration structures. We also look at how a board engages in those issues, including whether it probes high return products or business lines, and whether it understands strategies for cross-selling products, how fast growth is obtained and whether products are being sold to markets they are designed for. We are able, from all of this, to draw conclusions about the culture of a firm. This includes assessing if the perceived customer-focused culture is supported by, for example, regular discussions on conduct at board level and appropriate sales incentives plans."[7]

Needless to say, that each firm is now instituting, or trying to institute, the necessary structures so as to become compliant with the FCA's guidelines.

The U.S. regulators have also focused on this topic and the President and CEO of the New York Fed, William C. Dudley, made a speech about how important culture is to the safety of the financial institutions themselves and the industry as a whole and that should the firms fail to correct their cultures they might find that their firms might be downsized in order to maintain financial stability.[8] However, Mr. Dudley acknowledged that "regardless what supervisors want to do, a good culture cannot simply be mandated by regulation or imposed by supervision." Hence, the kinds of guidelines established by the FCA might not be instituted in the U.S.

From personal experience, however, I can confirm that senior executives at most of the major U.S. financial institutions have taken the speech by Mr. Dudley very seriously and are trying to learn from their European counterparts, specifically U.K. financial institutions, what they need to do in order to improve the risk culture of their organizations.

Irrespective of whether the FCA guidelines are instituted or not, in my opinion, and experience, it is going to be very difficult to improve risk culture within organizations that live off evaluation, packaging, and dissemination of risk, something they think they understand but recent evidence illustrates otherwise [Shojai and Feiger (2010)].[9] They might become compliant, but it doesn't necessarily follow that they will be able to, or even want to, change their culture, and they are not alone. In fact, most financial executives who have been looking at risk culture recognize this fact; hence the shift in the focus of discussions away from risk culture per se and towards finding out whether and how cultures of businesses can be improved. But, of course, improving the culture of a company, assuming it can be done in practice, doesn't necessarily translate directly into more ethical behavior. And that is the premise of this article; that financial services firms are not alone in acting in a way that many deem unethical, and that the term ethical business is nothing more than an oxymoron, irrespective of the industry you are considering.

Now, I am sure there are many academics who would argue with my take on the subject and genuinely believe that businesses can be both ethical and successful. In fact, many believe they can become even more successful by becoming ethical. I am not so sure. And I will explain in the following sections why despite their best intentions, businesses might never reach the levels of ethicality that academics would deem acceptable.

This paper, which is the sixth article in the Economists' hubris series of papers, looks at the topic of business ethics, with a specific focus on the financial services industry, and is organized as follows. In Section 2, I will explain why business ethics as a subject is so difficult to understand and explain. In Sections 3 and 4, I will explain why the behavior of governments influence how businesses behave, and why there is really no genuine mechanism to ensure that businesses do not act unethically. In Section 5, I will discuss why the pressures on today's businesses makes acting totally ethically very difficult and what lessons the Enron scandal and the bankruptcy of Lehman Brothers provided businesses in responding to crises. Section 6 concludes.

WHAT IS BUSINESS ETHICS?

While I don't for a moment claim to be very knowledgeable about the literature on the subject of business ethics, what I have found is that despite an entire publication dedicated to the subject, namely the Journal of Business Ethics, it is very hard to find articles that take the challenge of improving the ethics of a business head on. While that might have something to do with the fact that different terms are used to describe business ethics, such as such corporate

7 You can find the text of the speech on the FCA website: http://goo.gl/0GbnQ

8 You can view the text of the speech delivered at the Workshop on Reforming Culture and Behavior in the Financial Services Industry, held at the Federal Reserve Bank of New York on October 20th, 2014 via this link: http://goo.gl/qr7iLX

9 Shojai and Feiger (2010) suggest that assuming that the highly dubious mathematical and statistical models developed in universities and applied within financial services firms were accurate, even though they are not, the mere fact that financial services firms are dealing with multiple banking and trading systems in multiple locations, combined with problems that compensations cause when evaluating risk profiles of each desk, means that there is absolutely no way that financial institutions will be able to obtain a holistic view of the risks they face and hence manage them.

social responsibility (CSR), corporate citizenship, sustainability, being moral, etc., the reality is that it is very hard to describe what ethical behavior actually looks like.

Ethics is in the eye of the beholder. What might seem as unethical to one person might seem completely ethical to someone else. For example, some might view the fact that Western companies are moving manufacturing jobs to locations where labor is cheaper, such as China and India, as unethical, while others might view it as quite ethical, since it is helping bring a large number of people in these poor countries out of poverty, as it certainly has. Nonetheless, it certainly does not look very ethical from the perspective of those who have lost their jobs in the West, and there is no shortage of politicians and unions who believe this is the wrong thing to do. While this might be an extreme example, it does demonstrate that literally daily decisions of companies can result in actions that might be deemed unethical. And therein lies the problem, since we are not talking about companies not committing fraud, we are talking about companies behaving unethically.

It is not hard to determine whether a company has acted illegally or fraudulently, but it is very hard to determine whether specific actions by an organization are unethical, or otherwise. Furthermore, as Black and Anderson (2013) explain, "The question of what these ethical standards should be, how we judge them, and what we are ultimately aiming for, is central to this debate. When an aspect of the law needs to be determined, there is a mechanism for deciding what the outcome should be. But how should ethics and its grey areas be determined? Should public opinion be the point of reference? To do so could be a dangerous approach as public attitudes can change over time – ethics is not a static concept. While we may agree the norms at a high level, how they are applied in practice will be hotly contested and bitterly fought. We can already see this in the retail sector, where the line between 'mis-selling' and 'mis-buying' can be closely contested. What constitutes a 'mis-sold' product for one person, may be seen as a fair transaction for another. Clients and shareholders can also push firms to conclude transactions or pursue profits at the expense of ethics."

In support of this statement, there are many, including myself, who don't only blame the banks for the recent property market crisis. It is true that they should not have given many of the property loans that they did, but no one forced the borrowers to borrow either. They took on those loans knowing full well whether they were able to repay them or not. So, who acted unethically? The mortgage lender who knew the borrower was unable to repay their loan or the borrower who took on a loan they knew they could not repay? Would the fact that one company uses its connections to beat another company that also used its connections during a tender be considered unethical? Isn't that part of everyday business? Would the employees of the company that won truly think what they did was unethical? Or are both firms unethical for using their connections to get private information to win the business? If that is the perspective one takes, how would one go about viewing client expense accounts? Should those be banned, as they might tie the client to the company that wants their business?

In fact, don't academics use personal contacts for getting articles published where they can, or promotions? Don't they cite articles and theorems with questionable validity, as is very prevalent in social sciences, simply because it's accepted wisdom to quote them? Are they also not acting unethically? If they didn't, you would not have so many unnecessary economics articles published by the same group of academic institutions within so-called tier-1 economics journals. How many of those articles are of any practical use or genuinely scientific?[10] How many predicted the current crisis and its true causes? Don't charities, or religious organizations for that matter, act unethically when they give special treatment to major donors or powerful individuals? Do we all have the same opportunities to meet religious leaders? Somehow I doubt that. It is fully within one's rights to ask, "shouldn't these religious organizations or charities treat everyone equally"? And yet, businesses are accused of acting unethically by those same individuals and organizations who act in exactly the same manner when the opportunity arises. Lee (2010) provides an interesting comparison of the potential for dishonesty between business executives and preachers, politicians, and professors, and concludes that "in business the costs of determining honesty are smaller and the benefits greater than in the other three areas... the lower the costs and the greater the benefits of determining honesty, the more restricted are the opportunities to profit from dishonesty – and the less dishonesty will surface. Based on these arguments, my conclusion is that, as a rule, businessmen are more honest than preachers, politicians, and professors when making claims about their products."

10 One of the most fascinating comments about the scientific and useful nature of the work of economists was made by Friedrich August von Hayek during his prize lecture for the Nobel Prize in Economics in 1974, where he said that: "It seems to me that this failure of the economists to guide policy more successfully is closely connected with their propensity to imitate as closely as possible the procedures of the brilliantly successful physical sciences – an attempt which in our field may lead to outright error. It is an approach which has come to be described as the "scientistic" attitude – an attitude which, as I defined it some thirty years ago," is decidedly unscientific in the true sense of the word, since it involves a mechanical and uncritical application of habits of thought to fields different from those in which they have been formed." I want today to begin by explaining how some of the gravest errors of recent economic policy are a direct consequence of this scientistic error." You can read the entire speech on the Nobel Prize website: http://goo.gl/HDLSXd. See also Shojai and Feiger (2011) for the practical applications of award winning articles in finance.

The challenge of acting ethically, of course, becomes more difficult when international perspectives are taken into account. For example, what might seem ethical behavior to Indians might be, and probably is, very different to what U.S. or French citizens consider ethical. And ethicality changes with time, as behaviors become more or less acceptable.

Of course, the mere fact that it is hard to describe what ethical behavior on the part of management looks like has not stopped a number of academics from claiming that companies can and should act ethically and benefit from it.[11] The challenge, of course, is proving that is the case. Visser (2010) provides a summary of some of the findings in this space. He finds that, similar to other studies in economics, the "findings vary." For example, Griffin and Mahon (1997) reviewed 25 years of studies and found that a majority showed a positive link between CSR and financial performance, while Margolis and Walsh (2001) reviewed 80 studies, of which 42 show a positive relationship, 19 demonstrate no relationship, and four find a negative one. Two reports by SustainAbility (2001, 2002) also suggest mixed results. Some relationships between sustainability factors and business success factors are stronger than others, and in many cases, no relationship exists. Economist Arthur Laffer, on the other hand, in a review of Business Ethics magazine's 100 Best Corporate Citizens found "no significant positive correlation between CSR and business profitability as determined by standard measures" [Gupte (2005)]. Verschoor (1998) found that the financial performance of those corporations that, in their annual reports, commit to ethical behavior toward their stakeholders or emphasize compliance with their code of conduct (at the time of the study they accounted for 26.8% of 500 largest U.S. corporations) is significantly higher than those that didn't.

The challenge of determining whether there is a strong causality between acting ethically and improved business performance is exacerbated by the fact that different studies and organizations use different parameters to measure ethicality. The Institute of Business Ethics in the U.K., for example, uses the following parameters to determine ethicality: having a code of ethics, ratings for managing socio/ethical risks and being cited consistently in the annual list of Britain's Most Admired Companies [Webley and More (2003)]. The Ethisphere Institute uses a completely different methodology. Its corporate Ethics Quotient (EQTM) consists of five core categories: ethics and compliance program (weighting 35%), corporate citizenship and responsibility (20%), culture of ethics (20%), governance (15%) and leadership, innovation and reputation (10%).[12]

Furthermore, since all studies into business ethics have to rely on externally available information it is almost impossible to determine whether the parameters they have selected accurately encapsulate the thinking that went behind the decision that lead to the unethical behavior. This is because if the management's actions do result in an unethical outcome you need to be able to ascertain whether they were aware of it ex-post or whether it was a case of unintended consequences. Sadly, no one outside the group making the decision at the time has any idea of what the thought process was at the time the decision was made.

The long list of failed mergers proves that it is almost impossible to know what is really going on inside a company from the outside, and that is despite the bidding company spending months looking at, and talking with, the target. How are academics or analysts going to get information about how ethical, or not, a company really is from the outside? I am sure very few really thought that Enron or Bernard L. Madoff Investment Securities LLC were truly unethical companies before they both imploded. The latter even had money from a number of reputable charities, and according to Bragg (2002), Enron contributed millions of dollars to charities. Hence if you use associations with charities and the church as one of the parameters in your model, you would have missed both these firms. I am also pretty sure that Volkswagen was on many people's list of the most ethical companies, specifically because unions sit on the supervisory board of the company and so the company is deemed to work in the best interests of not only shareholders but also other stakeholders.

From a purely business perspective, Karnani (2010) makes a very important point when he says that: "the idea that companies have a responsibility to act in the public interest and will profit from doing so is fundamentally flawed." He goes on to make a very pertinent point, which is that "Very simply, in cases where private profits and public interests are aligned, the idea of corporate social responsibility is irrelevant: companies that simply do everything they can to boost profits will end up increasing social welfare. In circumstances in which profits and social welfare are in direct opposition, an appeal to corporate social responsibility will almost always be ineffective, because executives are unlikely to act voluntarily in the public interest and against shareholder interests. (...) Executives are hired to maximize profits; that is their responsibility to their

11 Denis Collins, who kindly reviewed this article, suggested that I have focused solely on the most extreme form of ethical theory, namely deontology/virtue ethics, and that I have ignored the other five important theories. The six theories of ethics (egoism, social group relativism, cultural relativism, utilitarianism, deontology and virtue) are easy to understand, and if one is honest quite commendable [Collins (2012)]. However, sadly while I am certain they are rich in connotations for academics researching this discipline, they are relatively unknown to most executives, and even the general public for that matter, and consequently rarely arise during discussions of implications of business decisions or in the internal struggles employees have over such issues.

12 Descriptions of the categories are available via this link: http://goo.gl/Q8rczi

company's shareholders. Even if executives wanted to forgo some profit to benefit society, they could expect to lose their jobs if they tried – and be replaced by managers who would restore profit as the top priority. The movement for corporate social responsibility is in direct opposition, in such cases, to the movement for better corporate governance, which demands that managers fulfill their fiduciary duty to act in the shareholders' interest or be relieved of their responsibilities. That's one reason so many companies talk a great deal about social responsibility but do nothing – a tactic known as greenwashing."

I cannot imagine many executives arguing with these facts. Of course, there are many academics who think that it is possible to do both, and when the two interact it certainly happens, but to suggest that firms should forgo profitable opportunities in order to be beneficial to society is not only naïve but also ignores the main economic objective of public corporations, which is to maximize shareholder wealth.

The shareholder wealth maximization principle is only one of the foundational concepts in finance that makes business ethics a difficult topic to tackle from the perspective of the finance discipline. The theory of agency costs [Jensen and Meckling (1976)], which suggests that managers are unethical by their nature and should be controlled since they use their private, internal, information about the firm to maximize their own wealth at the expense of the owners, is another. If we assume, as most finance academics do, that the unethicality of managers is given, then there really is not much that they can do that would surprise us. There are, of course, some academics who view the agency cost theory as oversimplistic [Clarke (2014)] and accuse it of ignoring the fact that there are many managers who can and do work in the best interest of the shareholders [Carlin and Gervais (2009)], but sadly they are a small minority. Most finance academics have to accept many of the highly contentious finance theories, such as agency costs, as given and not question them if they wish to get their papers accepted in the so-called top finance journals. This might help explain why Bernardi et al. (2008) find that "none of finance's top-40 journals or the journals listed in finance's version of Cabel's (2004) indicates an interest in ethics research."

The truth is that those who believe they have found significant relationships between ethical business and economic success are placing a lot of faith in the data's ability to scientifically quantify and determine ethicality,[13] a situation not much different to financial economics who believe that they can quantify asset prices and investment returns using so-called risk-adjusted models [Shojai and Feiger (2009)]. Furthermore, there is some evidence to suggest that in many cases economists don't even realize that their personal perspectives on issues actually impacts their findings. For example, a recent study by a team from Columbia Business School into the impact of political leanings on the areas they research and the findings of their studies found a (significant) correlation between the ideologies of authors and the numerical results in their papers. "That means that a left-leaning economist is more likely to report numerical results aligned with liberal ideology (and the same is true for right-leaning economists and conservative ideology)" [Jelveh et al. (2014)]. Moosa (2013) also finds that many academics run statistical models and then add and remove variables until they get significance in the direction they are looking for. Consequently, if you want to find that more ethical businesses experience better performance you will find data to support it.

A simple test of the difficulty in determining the ethicality of business is perhaps asking some of your colleagues, friends, or students to name five companies that they consider to be ethical and to explain why. Then genuinely try and find out if they truly are ethical or not, irrespective of what your personal description of ethical is. You will seriously struggle to find many people who can name such companies. I have been a student, lecturer, and employee in the field of finance and financial services for over 20 years and I am struggling to think of one company that would meet what I would call ethical. However, that by no means suggests that corporations are behaving unethically. It is just that like most people, I am not sure what corporations should do to achieve the label of being ethical.

Of course, my comments won't persuade those who believe businesses can be both ethical, or in fact have to be, and commercially successful, but I hope that I have at least made a strong enough case for those who have actually worked in the world of business and seen the true state of affairs within most companies. Having a code of ethics does not make you an ethical organization.

13 As Hayek (1974) stated: "Unlike the position that exists in the physical sciences, in economics and other disciplines that deal with essentially complex phenomena, the aspects of the events to be accounted for about which we can get quantitative data are necessarily limited and may not include the important ones. While in the physical sciences it is generally assumed, probably with good reason, that any important factor which determines the observed events will itself be directly observable and measurable, in the study of such complex phenomena as the market, which depend on the actions of many individuals, all the circumstances which will determine the outcome of a process, for reasons which I shall explain later, will hardly ever be fully known or measurable. And while in the physical sciences the investigator will be able to measure what, on the basis of a prima facie theory, he thinks important, in the social sciences often that is treated as important which happens to be accessible to measurement. This is sometimes carried to the point where it is demanded that our theories must be formulated in such terms that they refer only to measurable magnitudes."

LEARNING TO CONTROL THE NARRATIVE, LESSONS FROM THE POLITICAL CLASS

In my personal opinion, two events had a profound impact on the way that today's businesses operate and which support Karnani's proposition that most companies talk a good talk about being ethical but in fact don't do much about it; hence the highly questionable value of the external parameters used above to assess the ethicality of businesses.

The first was the development of the 24 hour news, and in specific its application to the first Gulf War, with the advent of the Internet also being part of that, and the second was the Enron scandal and Lehman Brothers' bankruptcy. Each had, in my opinion, their specific impact on how companies now behave and respond to crises. I will discuss the Enron scandal and Lehman Brother's bankruptcy further below.

The independent press

CNN, which was founded a decade prior to the first Gulf War, was already well-known in the U.S. and a number of Western countries, but it was the first Gulf War that helped bring its specific style of presenting news, 24-hours a day, to the masses worldwide. Their round the clock coverage of the war, with images of the so-called smart bombs (or as U.S. military calls it: precision-guided munition), made for great TV. I remember how impressed and excited I felt watching the images of these rockets destroying Iraq's different military sites. The success of CNN in attracting those interested in the news, and those who became interested as a result of CNN's work, spawned a number of other 24-hour news channels, including BBC, France 24, Russia's RT, Qatar's Al Jazeera, and Fox News.

The benefits of 24 hour news are clear to the viewers. However, those about whom the news was created, namely the politicians and governments, were not so sure how such an environment could be managed, if not controlled.

Well, it didn't take long, and the same war that made CNN was also a great training ground for governments in learning to control what the media should say and what it shouldn't say, and more importantly to create the environment in which they end up having to say what you want them to say.

The first thing that the U.S. government, or military, learned was that by embedding journalists within the military their only source of information becomes you. It's surprising that many news agencies actually claim being embedded with the military during war as something to be proud of. Fox news even has an OpEd from one of its journalists about how lucky he was to be embedded with the military and what a great journalistic experience it was [Leventhal (2013)], even though he could only stick his head out of the back of the armored vehicle once in a while. Obviously, it did not occur to him that it basically meant he only saw the war from the U.S. perspective and that he was not performing his most basic task as a journalist, namely to get perspectives from both sides of the story. It basically means they are not doing any investigative work and are just reporting what they have been told by the military.

Business executives have also learned how to use their PR departments to control what is said about them in the press. The most famous example of how pressures from advertisers can force publications to self-censor was the resignation of Peter Oborne, at the time the Chief Political Commentator of the Daily Telegraph, in response to the publication's refusal to publish a critical article he had written about why HSBC had canceled the accounts of a number of well-known British Muslims for fear of losing advertising from HSBC [Oborne (2015)]. In fact, Peter Oborne even claims in his resignation OpEd that not only was there self-censorship, but that there is a possibility that the advertisers requested that negative comments be removed. He states that: "I researched the newspaper's coverage of HSBC. I learnt that Harry Wilson, the admirable banking correspondent of the Telegraph, had published an online story about HSBC based on a report from a Hong Kong analyst who had claimed there was a 'black hole' in the HSBC accounts. This story was swiftly removed from the Telegraph website, even though there were no problems. When I asked HSBC whether the bank had complained about Wilson's article, or played any role in the decision to remove it, the bank declined to comment. (...) Then, on 4 November 2014, a number of papers reported a blow to HSBC profits as the bank set aside more than £1 billion for customer compensation and an investigation into the rigging of currency markets. This story was the city splash in the Times, Guardian and Mail, making a page lead in the Independent. I inspected the Telegraph coverage. It generated five paragraphs in total on page 5 of the business section."

But, of course, the Telegraph is not alone, most newspapers have to be careful about upsetting advertisers. The issue, however, is that sadly self-censorship is not the biggest problem we face when it comes to news media. The need for 24 hour news coverage, certainly facilitated by the Internet, means that news agencies that are already under financial strains, again brought about by the proliferation of non-print news media on the Internet, are forced to produce more news with fewer people. The result is that they are forced to rely on the help of public relations departments, be they corporate or governmental. Lewis et al. (2008), who looked into press independence in the U.K., found that "60% of press articles and 34% of broadcast stories come wholly or mainly from one of these 'pre-packaged' sources." And that "19% of newspaper

stories and 17% of broadcast stories were verifiably derived mainly or wholly from PR material, while less than half the stories we looked at appeared to be entirely independent of traceable PR." The pressure from PR departments and the need to continuously generate news has resulted in the following quote from the Health Editor of The Times, Nigel Hawkes, who says "We are 'churning' stories today, not writing them. Almost everything is recycled from another source […]. It wouldn't be possible to write so many stories otherwise. Yet even more is expected, filing to online outlets is now considered to be part of the job. Specialist writing is much easier because the work is done by agencies and/or writers of press releases. Actually knowing enough to identify stories is no longer important. The work has been deskilled, as well as being greatly amplified in volume, if not in quality" [Lewis et al. (2008)].

The close relationship between PR and the press has in fact reached a point where news organizations even advertise on behalf of corporations without making it seem like an advert. Robert Peston, who has recently left his post as the economics editor of the BBC to join ITV in the U.K., mentioned this in his Charles Wheeler lecture at the University of Westminster in London.[14] Peston stated that: "Today when I talk to my pals on newspapers, they talk of constant pressure – not to get unique and exciting stories, but to find ways of turning what is now called content, and is regarded by bosses largely as a commodity, into money. It is all about, awful word, monetising news. Which, of course, in one sense is completely necessary. There will be no jobs for any of us if there is no way to generate profit from news. But news that is a disguised advert, or has been tainted by commercial interests, is not worth the name. You might say that it is all very well for me to sit here smugly moaning about this, because I am lucky enough to work for the licence-fee funded colossus that is BBC News. But even we are not immune to a trend I fear is pernicious – because I saw an interview the other day with an executive of our commercial arm BBC Worldwide who said it was inevitable that we would be running what are known as native ads. "Native ads" is a terrible Orwellian Newspeak phrase for ads that look like impartial editorial. They could be articles written by a commercial company, or features written about a commercial company by the journalists of a news organisation but sponsored by that company. Or they may be videos either sponsored by a business or produced by the business."

Similarly, articles are published that could be perceived as being biased in support of one side of an argument without following the typical journalistic paths/guidelines. For example, a recent headline stated that "Labour won't admit it, but most people don't really care about tax avoidance." The suggestion is that if it mattered to people then they wouldn't continue buying from companies that have been accused of not paying their fair share of taxes. While that might be true, and people would probably continue buying from companies they don't necessarily deem as being ethical, the mere act of purchasing doesn't suggest they are condoning their actions. Furthermore, what data was this claim based on? But, of course, it fully supports the actions of those companies that have made sure they pay as little tax as possible via schemes that are now being judged to have been illegal.[15] Another recent article claimed that "The fiddling with temperature data is the biggest science scandal ever." This is despite the following quote on NASA's website: "Multiple studies published in peer-reviewed scientific journals show that 97 percent or more of actively publishing climate scientists agree: Climate-warming trends over the past century are very likely due to human activities. In addition, most of the leading scientific organizations worldwide have issued public statements endorsing this position."[16] Interestingly, it seems that even some of the oil companies themselves have known about the relationship between CO2 emissions and climate change for more than 30 years. Now, of course, it could be that author of the article actually believes what they are saying and has found data to support their position, but the fact that a blogger is quoted as evidence that more than 97% of scientific community are not only wrong but fabricating evidence and that none of the members of this community were allowed to respond to the commentary, which is what you would expect from a professional journalist, might suggest that causing a stir was of greater interest than scientific illumination. More importantly, it supports the position of all those companies that finance research denying climate change, such as the oil companies themselves.

The Internet has not only not helped, it has in fact made matters a lot worse. What most people fail to take into consideration is that while the business models of Internet companies allows them to provide services to you and I for free, it also means that our free services are being paid for by giant corporations. The fact that Facebook and Linkedin, for example, are free is that they live off advertising. Google is the same. You get a free search engine in return for the adverts that they place on your searches.

News organizations are no different. As more people get their news from the Internet, as is quite obviously happening, the economic model of news media is also changing. They are moving away from generating most of their revenues from actual sales of print publications, as well as, of course, the advertising within them, to

14 You can read the full text of Robert Peston's speech via this link: http://goo.gl/qFsBCg
15 The BBC News link to the illegality of the Starbucks and Fiat Chrysler tax deals: http://goo.gl/NEesYV
16 You can view the NASA quote here: http://goo.gl/4SU9hc. You can view the Huffington Post article about Exxon's knowledge of the risks of fossil fuel on climate change here: http://goo.gl/2893t7

becoming predominantly reliant on advertising based on the number of clicks each news item gets. The more clicks they get, the more they can charge advertisers, and as a result the more powerful those advertisers become. The best analogy I can think of is football teams. In the past, a good proportion of a football club's income came from ticket sales. Today, according to Deloitte, ticket sales only make up between 20% and 30% of the revenues of the top clubs. Most of their income comes from broadcast rights and sponsorships and advertising.

In fact, Peter Oborne mentions this phenomenon in his resignation OpEd [Oborne (2015)]. He says that "The arrival of Mr Seiken coincided with the arrival of the click culture. Stories seemed no longer judged by their importance, accuracy or appeal to those who actually bought the paper. The more important measure appeared to be the number of online visits. On 22 September Telegraph online ran a story about a woman with three breasts. One despairing executive told me that it was known this was false even before the story was published. I have no doubt it was published in order to generate online traffic, at which it may have succeeded." And, please remember this is the Daily Telegraph, a highly respected publication, not one of those publications or TV news networks where facts are a mere nuisance.

We have all seen how outlandish and unnecessary topics have been increasingly covered by what used to be considered quality publications in order to attract more comments from the readers. It seems the more comments a topic generates the more advertising revenues can be generated; hence the attraction of publishing offensive commentaries from people that many dislike in order to generate responses.

This race to the bottom has many participants, and its more than simply publishing untruths or nonsensical commentaries. The tone has truly become derogatory in many cases. The worst example that I have recently come across, though I am sure there are probably others, in the U.K. press, which used to be, and to a large extent still is, less hostile and more polite, is the commentary by the Chief Political Correspondent of a major British TV news network, about what the Leader of the Labour Party was wearing for his dinner with the Chinese President in October 2015. This is his description of Jeremy Corbyn: "Dressed in white tie and tails for the state banquet for the Chinese President, Jeremy Corbyn looked more like a downtrodden below-stairs butler or footman in Downtown Abbey than James Bond." First of all, why is how he looks in his suit even news, and why attack so personally someone who is such an important politician, irrespective of whether you agree with him or not. Furthermore, are these the types of comments that a so-called Chief Political Correspondent should make?

Many blame the U.S.-based Fox News for the more partisan and adversarial kind of news that many of us are witnessing today on the major U.S. news channels. Some of its programs use very aggressive tactics and they do not shy away from insulting those they invite to be interviewed. Fox News has also been blamed for replacing journalism that is based on facts with one that is based on opinions.[17] However, while no one can deny that Fox News has had a significant impact on how news is broadcast in the U.S., and to some extent in the U.K., it is not clear that other broadcasters didn't use its existence as an excuse to also change the way that they operate. While different channels support different political parties in the U.S., none are purely news channels anymore, or totally objective, and few have any patience for opposing views. It is within such a partisan environment that one is not certain whether what is being broadcast is based on genuine facts, honest independent opinions, or perspectives that have somehow been influenced by third parties.

Attempts have been made to find out whether U.S. media organizations were influencing their viewers in such a way so as to view misconceptions as facts. One of the most highly publicized studies was undertaken by the Program on International Policy Attitudes (PIPA) at the University of Maryland in conjunction with Knowledge Networks [Kull et al. (2013). The PIPA study, for example, found that among those who receive most of their news from Fox News, 67% believed that links between Iraq and al-Qaeda had been found, 33% felt that weapons of mass destruction were in fact found in Iraq, and 35% believed that the majority of people in the world support the invasion of Iraq by the U.S. In fact, the study finds that 80% of these viewers feel that at least one of these three misconceptions is a fact.

What is interesting is not that 80% of Fox News' audience believe these misconceptions, after all Fox News has made no secret of the fact that it aggressively supports the Republicans, hates the Democrats, and believes the U.S. war in Iraq was both justified and successful [Rosen (2009)]. What is interesting is the fact that no one quotes the findings for the other TV news channels, who don't perform much better. For example, while the 80% number for Fox News is quoted all over the place, no one mentions that 71% of CBS viewers, 61% of ABC viewers and 55% of NBC and CNN viewers also believe that at least one of these three misconceptions is correct.

17 A documentary about Fox News, called "Outfoxed: Rupert Murdoch's war on journalism," provides an interesting insight into how the broadcaster operates. The video is available on Amazon.com via this link: http://goo.gl/IW9Flh

In fact, around 50% of their viewers also think that links between al-Qaeda and Iraq were found. This means that if these numbers are correct more than half of the U.S. TV viewing public believe that at least one of these three misconceptions is correct. In other words, news was presented in such a way that viewers could not make a definite judgement about the invalidity of these comments, even by those TV channels that now claim to be against the war.

When these private corporate-owned broadcasters are compared with publicly-owned broadcasters the differences are startling. For example, among those who obtained most of their news from the National Public Radio (NPR) and the Public Broadcasting Service (PBS) only 23% believed at least one of the three misconceptions, only 16% believed that a link between al-Qaeda and Iraq was found and only 11% believed that WMDs were found in Iraq. It is not exactly clear why such stark differences are obtained between the audiences of public and privately-owned broadcasters, but it does show that there is a greater need for independent public broadcasters to ensure citizens are better informed about facts. Which is why despite the best efforts of some politicians and media owners Britain should do its utmost to retain its highly respected TV news broadcaster, namely the BBC, in public hands.

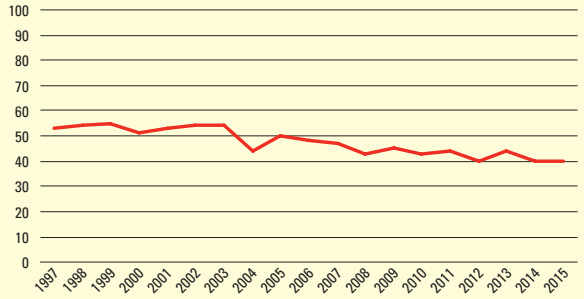
In summary, the pressures of the 24-hour news, aided by the Internet, and falling readership of print media (according to ZenithOptimedia, media consumption has fallen by 31% between 2010 and 2015 for newspapers, while it has risen by 105% for the Internet) [Ingram (2015)], have increased the negotiating position of PR and companies have learned how to use that power to influence. And, we have the Western governments to thank for that.

Why is that news?

The second thing that corporations also learned from Western governments is how to respond when a negative news item breaks. This is what I call the "why is that news" phenomenon. Western governments have mastered the art of managing bad news by asking "why is that news?"

My personal experience with this issue came during the second Gulf War. The U.N. weapons inspectors were sent to investigate Iraq's weapons of mass destruction program. But, the Iraqi government was refusing to let them enter the country because they suspected the inspectors of being CIA spies.[18] Of course, the U.S. was outraged by these allegations and accused the Iraqi government of using a weak excuse to prevent the U.N. inspectors from doing their jobs.[19] Of course, it later transpired that they were spies.[20]

It is not the fact that some of the U.N. inspectors were in fact CIA agents that is newsworthy, what is of interest is the U.S. response



Percentage of Americans who have a fair to great deal of trust and confidence in the mass media – such as newspapers, TV, and radio – when it comes to reporting the news fully, accurately, and fairly.
Source: Gallup

Figure 1 – American's trust in the mass media

to the revelation, which was to say that of course they were spies. Why wouldn't they be? Anyone who didn't realize we would send spies via the U.N. into Iraq is either very naïve or doesn't understand international politics.

The same justification was used when European leaders were outraged to find that their phones had been tapped by the U.S. National Security Agency (NSA): why is that news?[21] Most U.S. news organizations interviewed security officials and most said pretty much the same thing, that all allies spy on each other and there is no reason to be shocked by that. Trowbridge (2003) made the U.S. government perspective and response quite clear: "in response to uproar across the Atlantic, current and former U.S. intelligence officials and government leaders have argued that, when it comes to spying on allies, the U.S. isn't alone. "It's well known that our allies do spy on us." It seems what was news was that Snowden had made this information public.

18 One reference to Iraq's reaction is provided in this Sky News link: http://goo.gl/Ni4R9R

19 The response of the U.S. government to the Iraqi accusations is provided in this BBC news link: http://goo.gl/4ri4tl, or in this Washington Post link: https://goo.gl/gNsDC3

20 The original story was published by the Washington Post, and can be viewed via this link: https://goo.gl/LAI42e

21 Reactions to Edward Snowden's revelations that a number of world leaders' phones were tapped by the NSA can be found in this Guardian link: http://goo.gl/6QSH65. Interestingly, even some in the media, including the British media, have used the why is that news response. Charles Moore, the former editor of a major British newspaper, stated that: "I gather that Wikileaks worshippers have been disappointed that the citizens of Britain and the United States have not acclaimed Snowden's courage or been shocked by his revelations. Public opinion seems to have given a worldly shrug and said, "Obviously, our secret services spy on us in cyberspace; what's all the fuss?"

And, of course, the Americans were right. Less than a year after the German outrage over the NSA phone tapping, it was revealed that Bundesnachrichtendienst (BND), the German foreign intelligence agency, was collecting information on European governments, officials, and corporations for the NSA. Later, it transpired that the NSA was using the facilities created to spy on Germans as well.[22]

Obviously, the issue is not whether Edward Snowden and wikileaks are traitors or a heroes, or whether or not governments should tap each other's, and our, calls and emails; and more importantly whether such revelations are even news. The issue is that these events illustrate that there are very powerful media organizations, and in many cases high-end ones, on both sides of the pond who happily jump to the support the actions of governments and corporations, irrespective of whether they are legal or ethical.

The result of such actions have obviously not gone unnoticed. According to Gallup [Riffkin (2015)] around 60% of Americans do not trust mass media's coverage of major events (Figure 1). While that might seem bad, it is still better than how journalists are perceived in Britain, where, according to Ipsos Mori, 72% feel that they can't trust journalists to tell the truth.[23] Despite the bad numbers, journalists can take heart from the fact that in both countries they are still more trusted than their governments.

Obviously, with such media organizations around, corporations feel quite comfortable about saying one thing and doing something completely different, since they know that damages can be controlled somehow. They can also play the "why is that news" card. Unless, of course, the damage is so huge that it cannot be managed, such as the recent Volkswagen emissions scandal. Volkswagen has had to admit that it had installed devices that could detect when they were being tested and changed the performance of the car accordingly to improve results; hiding the fact that under normal driving conditions the cars were producing more than 40 times the permitted nitrogen oxide pollutants. So far, that is bad enough. What is even worse, and sadly doesn't get as much attention it seems is that according to the Financial Times: "EU officials had warned of the dangers of defeat devices two years before the Volkswagen emissions scandal broke, highlighting Europe's failure to police the car industry."[24]

In summary, we find that unless it's a major catastrophe, companies and governments can and do respond by saying "why is that news." Citizens have been trained to accept that unethical acts are undertaken in certain circumstances and the fact that they are revealed should not come as news to anyone.

The situation will certainly get worse as ownership of media organizations becomes more concentrated, as it has been over the years.[25] Lewis et al. (2008) found that 87% of the news items they studied were based on the information of a single source, and that in only 12% of cases where the claims thoroughly corroborated. There is really no fact checking, especially when the claims come from government officials, who can revoke your access to their future press conferences, or companies, who can take their advertisement dollars elsewhere.

SYSTEM OF CHECKS AND BALANCES

If as citizens we cannot rely on the media to be honest with us about what is really going on in the world, and should sort of accept that news can be, and is, distorted, then who should we look for to get any kind of resemblance of facts about the corporations that we do business with? More importantly, from the perspective of this article, how do we know whether they have indeed acted unethically, and should we really expect them to be ethical?

Of course, when most people talk about companies acting ethically these days, they are referring to how they treat their labor, be it domestically or overseas, whether they pay them such that they can have a reasonable quality of life, how they deal with their suppliers,[26] whether they cause pollution, whether they pay their fair share of taxes, and whether they pay bribes to win deals in countries where government officials who tender the contracts are less than scrupulous.

When people are considering these issues, they look to the press, non-governmental organizations (NGOs), and of course their own governments to inform them when a company acts unethically from the perspectives I mentioned above.

22 Der Spiegel: http://goo.gl/XapPuu
23 Ipsos-Mori: https://goo.gl/BvBTIL
24 Financial Times: http://goo.gl/1t8xUP
25 Based on testimony before the House Judiciary Committee examining Comcast's acquisition of NBCUniversal "in 1983, 50 companies owned 90 percent of the media consumed by Americans. By 2012, just six companies – including Fox (then part of News Corporation) and Time Warner – controlled that 90 percent," The New York Times article can be found here: http://goo.gl/mPLOzk. In the U.K., 70% of national news circulation is in the hands of just 3 companies, and BBC and ITV control 88% of the national and international news viewed on TV: http://goo.gl/J7OczD
26 This is typically referred to as fair trade and there are organizations who purport to help companies act in a manner that would be considered fair trade and individuals to buy from companies who have signed up to fair trade, such as the Fair Trade Foundation in the U.K.

Well, I think we have exhausted the idea that the media will report aggressively on whether companies have paid bribes to win deals or treated their suppliers and employees unfairly. Of course, I am not denying that such investigations do happen, and most of us have read and seen such coverage. But, they tend to be either in response to a catastrophe, such as the collapse of the Rana Plaza building in Dhake, Bangladesh,[27] where more than a 1000 people lost their lives, or ad-hoc investigative journalists who go undercover to find out how workers are treated in factories in the East or how some manufacturers pollute the sea, land, or air in the areas where they are based. The discussion on taxes are neverending, and it's quite clear that corporations have the upper hand. Needless to say that the little coverage that these issues do receive don't result in much changing. Of course, in this case the media cannot be blamed since, as was discussed above, they need to write about topics that attract readers online and it seems discussing unethical behavior by companies doesn't. Or at least it doesn't as much as the lost advertising revenues from those companies being covered.

There are also NGOs such as Transparency International that rank the degree to which the public sector of a given country are corrupt. Their annual rankings of corrupt countries is meant to highlight that corruption exists and which countries are the worst offenders. What Transparency International fails to realize is while their rankings, like all rankings, are an interesting read for a couple of days after their publication, they really don't have much of an impact on whether countries become more or less corrupt. In fact, if their ratings were to have any value or benefit at all, they would rank the companies that have paid the most bribes per year. But, obviously that is much harder to ascertain, and the last thing they need is lots of defamation cases in courts from companies that were either unfairly accused, or who knew that only under extremely specific circumstances would their actions be found and proven.

As far as businesses are concerned, Transparency International rankings are in fact quite useful, since they can guide them on how to prepare for dealings with the governments of those countries that are at the bottom of the list. More importantly, and as George Monbiot explained in his column in the Guardian [Manbiot (2015)], the parameters used are simply based on whether bribes are paid to win business and fail to take into account the kinds of corruption that all of us know is rife in the West, such as creation of tax havens, which the City of London is the largest center of, or when connections are used at the highest levels to win business. Just think about how many multi-billion dollar contracts were awarded by the U.S. governments to companies close to the inner circle of President Bush's cabinet without being put out to tender [Fifield (2013)].

And, of course, when large scale corruption on the part of a major corporation becomes public knowledge, they can find protection from their own governments if they are deemed too important to the economy; interestingly those same governments, or their regulators, that have been established to identify and prevent unethical behavior on the part of businesses. The best example of this is, of course, the Serious Fraud Office investigations into alleged bribes paid by British Aerospace (BAE) to win a multi-billion dollar contract to provide military equipment, mostly fighter jets, to Saudi Arabia, which was stopped by the then Prime Minister, Tony Blair, on the grounds of national security.[28] So, next time you feel like accusing banks of abusing their too big to fail status, think again. All companies use their connections to their advantage.

Of course, it's not only in times of crisis that corporations lean on governments for help. Business lobbies are renowned for their efforts to get governments and regulators to water down [such as they have done and continue to do to the Dodd-Frank Act, antipollution regulations, or even executive accountability in financial services firms: Picchi (2015); Hanrahan (2014); Nelsen (2015); Bowers and Treanor (2015)], or even eliminate [such as the Glass-Steagal Act: Brown (1995); Crawford (2011)] regulations that they deem to be harmful to their clients [Cave and Rowell (2014)]. As the recent decision by the FCA to not publish its report into the culture of banking in the U.K. illustrates, they can even prevent critical reports undertaken by the regulators themselves from being published [Dunkley (2015)]. There have, of course, been accusations that the so-called revolving door between financial regulators and financial services firms [Johnson and Kwak (2011); Masters (2012); Ross Sorkin (2011)] ensures that regulations are either not too restrictive or are not executed in the ways they were intended [Popper and Eavis (2014)]. It is hard to determine just how the revolving door actually impacts enforcement by regulators, but there is no doubt that it exists. A recent study by the Federal Reserve Bank of New York [Lucca et al. (2015)] found that while the revolving door does exist, it does not impact the ways in which regulators monitor financial services firms and that their findings do not find support for the so-called quid-pro-quo hypothesis, where the expectations of a future career in financial services firms might impact the strictness with which the regulators enforce regulations. It should be said that the proxy they have used, which is that the number of regulators moving to the private practice should fall during periods of high enforcement activity, is highly dubious, as all such proxies are, and simply does not provide the necessary proof that regulators are in fact being

27 BBC News: http://goo.gl/ZmKP2R
28 BBC News: http://goo.gl/dUIBny

effective in their enforcement duties. The authors themselves acknowledge the shortcomings of the proxies used, though I suspect the proxies have significantly lower explanatory powers than the authors acknowledge.

Nevertheless, despite the fact that the influence of business lobbyists on governments cannot be underestimated, it is not always easy to determine just how successful they really are in getting the changes they aim for. As Drutman (2015) states in his book, while 95% of the biggest spending lobbyists are those representing the interests of businesses it is not always clear they get what they pay for. Furthermore, it doesn't always follow that the changes they aim for are to the detriment of others. However, it would not be too farfetched to claim that similar to the growing influence of business on media, their influence, through their lobbyists, on governments, in specific in the U.S.,[29] and to a lesser extent in the U.K., is also increasing: making their ability to get away with unethical behavior even greater than in the past. Just how much more is up for debate, and beyond the scope of this paper.

The fact of the matter is that corruption is rife everywhere, in different forms, and companies and citizens are fully cognizant of this fact. They realize that in most parts of the world, even in the so-called developed and democratic West, it's the cost of doing business. Some gets reported, but by and large it is overlooked. And, of course, the citizens who buy the products of these companies are neither aware of the full scale of corruption nor naïve enough to think that this specific company is worse than any of the others.

Rating agencies and auditors

Obviously the system of checks and balances within the corporate world doesn't start or end with governments, NGOs, or the media. In most developed countries, it also relies on auditing firms, who review the financial health of organizations and help with the preparation and attestation of the annual reports and accounts, and rating agencies, who evaluate the riskiness of securities issues by large corporations, and as a result the companies themselves. What differentiates rating agencies and auditors, however, is that they have access to the kinds of internal information that are not available to anyone outside the companies they work with. And, while it is clearly not their role to assess whether companies act ethically or not, they should be able to highlight fraud, an extreme form of unethical behavior, to investors, and even possibly the general public and regulators, and could be a source through which such activities could be identified and studied.

Rating agencies

Rating agencies provide guidance on the financial status of major corporations and their obligations. They try to determine the likelihood that investors might not be able to recover their investments from the organizations they have lent to. In other words, the likelihood of failure. Their assessment of the likelihood of corporate failure, or the securities they issue, should also be able to uncover fraud, one of the means by which the management can act unethically, since it is an important factor in the business failing. Well, I don't need to tell the readers how effective they were during the run up to recent crisis in doing just that. I am sure most readers are well aware of controversy surrounding the actions of the rating agencies during the recent crisis, and none is a better illustration of that than the "Financial crisis inquiry report"[30] that the National Commission on the Causes of the Financial and Economic Crisis in the United States published in 2011. There is a fascinating point in the report where a former managing director of Moody's, Gary Witt, was asked whether: "investment banks frequently threatened to withdraw their business if they didn't get their desired rating." Witt replied, "Oh God, are you kidding? All the time. I mean, that's routine. I mean, they would threaten you all of the time... It's like, 'Well, next time, we're just going to go with Fitch and S&P.'" Another former managing director of Moody's, Jerome Fons, suggested that Moody's was complaisant when it should have been principled: "[Moody's] knew that they were being bullied into caving in to bank pressure from the investment banks and originators of these things. ...Moody's allow[ed] itself to be bullied. And, you know, they willingly played the game... They could have stood up and said, 'I'm sorry, this is not – we're not going to sign off on this. We're going to protect investors. We're going to stop – you know, we're going to try to protect our reputation. We're not going to rate these CDOs, we're not going to rate these subprime RMBS.'"

In support of the comments above, Friedman and Friedman (2010) provide the following example: "Moody's graded the securities that consisted of Countrywide Financial's mortgages – Countrywide is the largest mortgage lender in the United States. Apparently, the ratings were not high enough and Countrywide complained. One day later, Moody's raised the rating. This happened several times with securities issued by Countrywide."

While some might even accuse the rating agencies themselves for acting unethically, the main issue is not that the rating agencies are claiming to have capabilities that they obviously don't, no one does, namely that they can accurately value risk, it is that they are paid by

29 For example, the decision in support of Citizens United by the U.S. Supreme Court significantly increased the influence of businesses on politicians in the U.S. Please refer to this New York Times article for an analysis: http://goo.gl/7Tz5ha

30 The full report is available via this link: http://goo.gl/QiE0K

those same companies/institutions that they are supposed to monitor or rate (although they actually rate the specific securities issued by these institutions). They are supposedly doing their job on behalf of the investors by determining the risk of financial instruments that issuers are trying to sell in the market; hence you would expect that they would do their analysis, issue the ratings, and then sell the rating to investors who wanted to invest in those specific securities. But, that's not how the model was developed. It would be like having a police service paid for by criminals. The mere fact that these institutions still exist and that their compensation models have not changed, despite the tremendous damage that they caused by their ratings in the lead up to the crisis, illustrates how strong lobbying can protect businesses in the West. So long as these institutions are paid by those who they rate, the risks that we faced during the crisis will remain, as most observers will agree. More importantly, they cannot be relied upon to identify fraudulent behavior, and certainly not unethicality, on the part of companies they rate.

Auditing firms

The economic model of auditing firms is also the same as rating agencies; they are paid by those same companies they are auditing. In fact the degree of concentration is not too different either. The world of ratings is dominated by the three major rating agencies, Standard & Poor's (S&P), Moody's, and Fitch Group, and the world of auditing is dominated by the four major auditing firms, namely PricewaterhouseCoopers (PwC), Deloitte Touche Tohmatsu Limited (Deloitte), Ernst & Young (EY), and KPMG. According to Gerakos and Syverson (2015), the Big 4, as they are known, "handled 67% of audit engagements and collected over 94% of audit fees" of publicly traded companies in the U.S. in 2010. In the U.K., the Competition Commission's report into the audit profession found that the Big 4 "collectively audit more than 95% of the FTSE 350 companies" [Prem (2013)]. This means that we are dealing with a highly concentrated market for auditing. And efforts to reduce that concentration have not borne much fruit, including mandatory auditor rotation, since all that happens is one Big 4 audit firm is replaced for another [Fleming and Smith (2014)].

Auditors perform a very important task. They help prepare and attest the validity of the financial statements of public companies for investors. Investors rely heavily on the financial information published in companies' annual reports. They are the window through which investors look inside the companies they invest in. The obligations of the auditors, and hopefully the reliability of published accounts, increased subsequent to the introduced the Sarbanes-Oxley Act of 2002 [McConnell and Banks (2013)], which was introduced in response to the bankruptcies, mostly related to fraudulent activities, of Enron, Worldcom, and Tyco, and the collapse of the auditing firm Arthur Anderson.

PWC	2004	2014	Percentage growth
Advisory	0.4	1.1	188%
Audit	0.7	1.0	40%
Tax	0.5	0.7	49%
KPMG			
Advisory	0.5	1	105%
Audit	0.3	0.5	61%
Tax	0.3	0.4	38%
EY			
Advisory	0.3	0.9	225%
Audit	0.4	0.6	46%
Tax	0.2	0.5	150%
Deloitte			
Advisory	0.5	1.1	105%
Audit	0.4	0.7	97%
Tax	0.4	0.6	49%

Source: Agnew (2015a)

Table 1 – Breakdown of revenue growth of the Big 4 auditing firms in the U.K. (£ bln)

The Act also stated that audit firms should no longer be allowed to provide consulting services to their clients, as it was perceived that it had been instrumental in Arthur Andersen becoming willing participants in the Enron fraud. It is alleged that the fear of losing the consulting business prevented Arthur Andersen from doing its job as an independent auditor. As a result, most of the Big 4, as they were after the Arthur Andersen bankruptcy, started selling their consulting businesses. Ernst & Young sold their consulting business to Cap Gemini, PricewaterhouseCoopers (PwC) sold its consulting business to IBM, and KPMG spinned off its consulting business and called it BearingPoint.[31] The only one that didn't sell or spin-off its consulting business was Deloitte. Deloitte took a number of turns to end up back where it was. It initially announced that it would separate its consulting business and call it Braxton [Singleton (2002)], which confused many with the famous Scottish water, Buxton. Then, within 12 months of announcing its plan to spin-off its consulting business, Deloitte announced in March 2003 that it would retain its consulting business [Glater (2003)]. This obviously shocked most commentators, as well as the other Big 4 auditing firms, since they had already sold or spun off their consulting businesses.

31 For reports on the three transactions refer to the following articles: New York Times: http://goo.gl/Jh9H6J; Wall Street Journal: http://goo.gl/Th8OHs; AccountingWeb: http://goo.gl/RpzpVe.

	PwC		Deloitte[1]		EY[2]		KPMG	
	Revenues	Revenue growth	Revenues	Revenue growth	Revenues	Revenue growth	Revenues	Revenue growth
Assurance	15.2	0.3%	9.8	-3.0%	11.3	0.6%	10.03	-4.11%
Tax	8.9	1.5%	6.7	3.1%	7.5	3.6%	5.31	0.76%
Advisory	11.2	12.3%	15.3	6.3%	9.8	10.7%	9.10	0.11%

Source: Annual reports (2015 figures)
Notes: All percentage changes are calculated using dollar figures, not local currency. [1] Includes financial advisory (U.S.$ 3.1 billion in 2015, and U.S.$3.0 in 2014 when calculating percentage change), [2] includes Transaction advisory services (U.S.$ 2.5 billion in 2015, and U.S.$2.3 in 2014 when calculating percentage change).

Table 2 – Breakdown of revenues and revenue growth rates of different businesses of the Big 4 audit firms

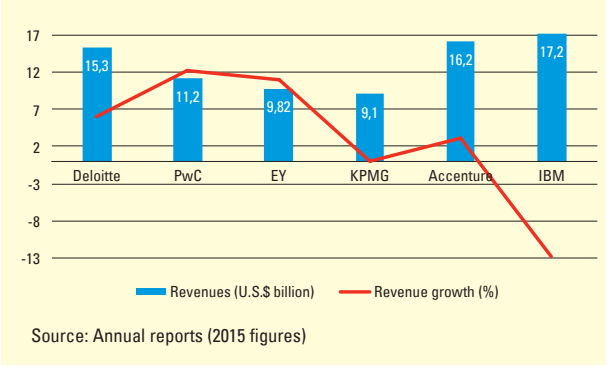
It seems Deloitte had found a way of keeping its consulting business, and it was to do so by consulting non-audit clients. James Copeland, the CEO of Deloitte at the time stated: "We fully intend to comply with those laws and regulations. (...) We really already have focused our Deloitte Consulting practice on the 75 percent of the market that we don't already audit." Needless to say, the others also reverted back to providing consulting services as well; once the non-competes with the businesses that were sold or spun off had expired.

According to research done by the Big4.com website: "By service line, audit accounts for 42% of total revenues and grew a solid 2.4% from 2013 to 2014 after a 0.2% growth in 2013. Tax services are 23% of total revenues and rose 6.2% in 2014 after rising 3.6% from 2012 to 2013. Advisory services have 35% of total revenues in 2014, with revenues accelerating by a strong 9.9% in 2014 after growing 6.8% from 2012 to 2013." The Financial Times (FT) looked at the U.K. market for the Big 4 and their findings are presented in Table 1. What is clear is that not only is the percentage of growth in the consulting business significantly higher than the audit business, it has in fact surpassed audit and tax to become the largest contributor to revenues. This means than in a few years the revenues from consulting will dwarf audit.

The growth of the consulting arms of the Big 4 in the U.K. is not only exceptional because it has surpassed the growth of audit and tax within these firms, their growth is also remarkable because their revenues in consulting is growing at a faster rate than the industry average. According to Source Information Services [Agnew (2015b)], "the big four accounting firms increased their revenues 8.9 per cent to £2.34bn in 2014, outperforming the rate of growth in the overall U.K. consulting market, which increased revenues 6.6 per cent to £6.02bn last year." If these numbers are correct, then the Big 4 have gone from pretty much having no consulting business, with the exception of Deloitte which retained its consulting business, to controlling nearly 40% of the U.K.'s consulting market in just over a decade.

From a group perspective, however, audit is still king in all but one of the Big 4 firms. Only in Deloitte, even if you exclude financial advisory, have the advisory revenues surpassed that of audit (Table 2). Nevertheless, it's quite clear that, given the growth of advisory businesses, in a few years advisory will surpass audit in all of the Big 4 firms. It has probably already done so in the major member firms, such as U.S. and Europe, of all Big 4 firms. It is in emerging markets, where advisory is still in its infancy, that audit still generates a greater share of the firms' revenues.

According to Gartner [Heng et al. (2015)], by 2014 the Big 4 controlled over 40% of the global consulting market, and they are experiencing much faster revenue growth rates than their peers. Furthermore, Deloitte is now the largest consulting firm in the world and the other Big 4 firms make up the rest of the top 4. I am not sure where Gartner derived their numbers from, and it seems for both IBM and Accenture they used fourth quarter figures rather than the full year figures. Consequently, I recreated the data from the actual published accounts of the major consulting firms. These are presented in Figure 2. While the results are different to those obtained by Gartner, it is quite clear that the Big 4 firms are fast catching up on their largest competitors, though KPMG does seem to be experiencing a dramatic slowdown in growth across all businesses. Furthermore, should the current growth rates continue they will become significantly larger than the other two giants of consulting, namely IBM and Accenture, and they could control 40% of the global consulting market within the next few years. It is also clear that the Big 4's brands and the connections they have built through their audit arms have been invaluable, otherwise such exceptional growth would have been impossible.



Source: Annual reports (2015 figures)

Figure 2 – The growing power of the Big 4 in the world of consulting

The success of the consulting arms of the Big 4 has once again raised the issue of whether the consulting partners might wish to divest themselves from the audit firm so that they don't have to share the fees they generate with their less profitable colleagues. Those who are old enough to remember, that is exactly what happened to Andersen Consulting. The consulting partners got tired of having to share their fees with the less profitable and less exciting audit arm and voted to separate the businesses, which resulted in two companies, Arthur Andersen and Accenture.

Of course, the spun off consulting arm doesn't always experience success. The Chapter 11 filing by BearingPoint, the former consulting arm of KPMG, in 2009 is proof of that.[32] Not all separated consulting arms become Accenture. And even Accenture seems to be struggling to keep up with the growth of the consulting arms of the Big 4. That is why the likelihood of the consulting partners taking the risk of losing the connections and the brand name of the auditing businesses is a lot less than many fear.

The reason the involvement of the Big 4 in consulting matters is that many are concerned that, like Arthur Andersen, the revenues generated from the consulting business, or their potential, might influence the actions of the audit firms, or at the very least damage the quality of people going into auditing, because these firms are focusing all their efforts, best people, and investments into consulting.

Academic investigations into the impact of fees from non-audit services (NAS) is, similar to all economic studies, inconclusive. Some studies find that NAS do not impact the independence of audit companies or the quality of their audits [Ashbaugh et al. (2003), DeFond et al. (2002); Schneider et al. (2006), Lim and Tan (2008), and Habib (2012)], while other, more recent studies, find that they do [Causholli et al. (2015)]. Quick and Warming-Rasmussen (2015) find that the mere act of providing NAS does not in of itself result in a negative perception of the independence of audit from the perspectives of German individual investors. However, when the ratio of NAS fees is high, it does result in a negative perception. Their findings corroborate those of Krishnan et al. (2005), Francis and Ke (2006), and Krishnamurthy et al. (2006). Goldwasser (2002) and Coffee (2006) suggest that the expectation of future NAS fees might influence some firms to send partners who are better in relationship building than audit quality, and might even impair their objectivity. Causholli et al. (2014) find "strong evidence that the anticipated future provision of NAS does represent a source of impaired independence in the current year." They also find that "clients with little or no potential for sales of new NAS would tend to be assigned to technical partners. To the degree that relationship partners possess less technical accounting and auditing skills, those clients assigned to them would receive a relatively lower quality audit."

It is obviously very difficult to scientifically determine the extent to which NAS fees impact audit quality and independence, since similar to ratings that have been inappropriately issued, you only find out the true state of affairs when there is a crisis. So long as there is not a crisis no one will know for sure how accurate or not the analysis of the independent ratings agencies or auditors truly are. For example, if the property market in the U.S. had not collapsed we would still think that the CDO ratings were accurate.

So, the question is how do we find out if NAS fees influence the independence of audit firms? Well, the honest answer is we can't. At least, not scientifically, since we have no access to the private discussions between the clients and their auditors. And, the major auditing firms are not shy in using their influence to ensure regulators don't get too critical or undertake the kind of scrutiny that might make people question their independence [Levinson (2015)]. Two qualitative parameters could possibly be considered to determine whether auditors are truly independent [Kaplan (2004)], and whether NAS might have an impact on their independence, but they cannot be tested quantitatively.

The first is to see whether there have been situations of bankruptcies, or major restatements, where the auditors did not warn the markets and regulators beforehand, even though they were aware of the problems, or were in fact found to have assisted the client in hiding its true state of affairs. In recent years there have been a number of such accusations against the major global auditing organizations, despite the establishment of auditing oversight boards,

32 CNN: http://goo.gl/Wwtf3W

such as the Public Company Accounting Oversight Board (PCAOB) in the U.S.[33] Of course, it's hard to determine whether the actions of the auditors would have been any different had there been no NAS fees involved, or expectations thereof, but their actions have raised questions about their impartiality when large fees are involved.

Needless to say, similar to the major banks, the only penalties paid by those organizations that were found to have bent the rules have been purely financial, and none faced anything even remotely similar to the problems that caused Arthur Andersen to collapse. Perhaps, the auditing industry, as some have suggested, has also become too-big-to-fail, and regulators are willing to turn the other way in order to avoid having another one of these giants fail [Economist (2014)].

Another method by which you can determine independence of auditors is to see how comfortable they are in criticizing the wrongful actions of their clients and highlighting their errors or mistakes publicly. You will find that the big auditing companies, similar to other consulting or professional services firms, are careful about coming across as too critical of their clients. While this might not be such a big issue for the major consulting firms, given the importance of auditor independence and their roles as watchdogs of corporate accounting, this might be something to be concerned about. In fact, many have questioned the role of auditors in the run-up to the recent financial crisis and why it was that they failed to raise any flags about institutions that faced severe financial problems during the crisis [House of Lords (2011); Rapoport (2010); Sikka (2009)]. This is perhaps what Goldwasser (2002) and Coffee (2006) meant when they suggested that expectations of NAS fees might influence objectivity.

Sadly, what this means from a business ethics perspective is that neither ratings agencies nor auditors, two organizations with unique access to the inner workings of their clients, can be considered as useful sources of information for determining the ethicality of corporations or as potential partners to fight against it.

Whistleblowers

The most reliable means of determining whether organizations are doing something untoward is when someone working for the organization blows the whistle on their illegal, and in certain circumstances, unethical, activities. And the regulators know this, which is why there is a heavy emphasis on protecting whistleblowers in the current risk culture discussions. However, why would whistleblowers be comfortable about blowing the whistle on their companies? They have already seen how someone like Snowden, who blew the whistle on the illegal activities of his government, has been treated. That I think was the best gift that the U.S. government could have given to corporations worldwide. People are now petrified of blowing the whistle on their companies and companies are much more careful about sharing their secrets with too many people inside the firm and much more aggressive in pursuing those who blow the whistle on their activities. The five-year prison sentence given to Hervé Falciani[34] by the Swiss court for disclosing that HSBC was helping clients launder money and evade taxes, while the bank itself was cleared of any penalty and simply paid a contribution to the state, is a case in point.

However, even before the Snowden revelations, most people were already very concerned about blowing the whistle on their employers. For one thing, and unlike what people who deal in hypotheticals think (such as regulators who want to be seen to be doing something, even if they are fully cognizant of its ineffectiveness, or academics), blowing the whistle on your employer is the most certain way of destroying your career. Of course, if you ask most employers they will tell you they would be very happy to hear from a whistleblower, but the truth is that they don't. They will find the best and quickest way of getting rid of them. And, that's because whistleblowers aren't just people with ethics, they are also people accused of getting in the way of the company's or department's progress. They are not team players. And, of course, they can't be, otherwise how would they raise their hands and say to the rest of the team "this is wrong."

Put yourself in the shoes of any employer and you will see what I am referring to. Worse still, having a whistleblower sign on your forehead is worse than having a prison record. No employer will touch you. Why would they? Who wants the headache of hiring someone who will spill the beans on the company in today's world, where bad news can travel very fast? It seems that not even the U.N., an organization that does everything in its power to destroy internal whistleblowers [Bowcott (2015); Hamilton-Martin (2015); Newman (2015)], while at the same time publishing guidelines on how they should be protected by governments.[35]

That is why no matter what the regulators say, companies will neither protect whistleblowers nor employ them. It's the fastest way of destroying your future career, and most employees have learned that.

33 Please refer to this review of PCAOB by the Washington Post: http://goo.gl/Z5dCiu
34 The Guardian: http://goo.gl/pdQuhT
35 United Nations: http://goo.gl/g4iDsa

THE REALITIES OF THE BUSINESS WORLD AND THE LESSONS LEARNED

As mentioned above, there are a number of academics, and an entire discipline in management, that believe businesses can be made to act more ethically, with the description of what is actually ethical still up for debate, and that this can be achieved by having ethical guidelines and regulations, which also includes whistleblower protection.

In my own personal experience as a former academic and someone who has been involved in a number of so-called independent studies sponsored by non-academic organizations, academics who undertake these types of studies tend to fall into two camps.

The first are those who undertake so-called independent studies paid for by industry bodies or organizations. From my own experience, I know that in most cases the press releases have been written long before the study has actually begun. And I am not referring here to those who have sold their souls to the devil and actually work for those think-tanks that are financed by specific political or business organizations and just publish documents in support of their agendas, irrespective of any of it having any basis in reality. I am referring to those who are actually employed by academic institutions. I was personally left in shock when a study I was sponsored to do was banned from publication when the findings didn't match those of the sponsor, who had already written the press release. After repeated refusals to change the findings, the report was finally destroyed. Needless to say, that was the last time I personally undertook a sponsored study myself.

I don't need to tell the readers how many such studies, with highly questionable conflicts of interest, are conducted and published each year [Friedberg et al. (1999); Lexchin (2012); Lexchin et al. (2003); Krimsky (2003); Resnik (2007); Sismondo (2008); Stelfox et al. (1998)]. I am also that sure there are many that you have come across that go against conventional wisdom, and that have received a lot of press attention, that you thought to yourself, they must be kidding.[36] Chaudhry et al. (2002) found that when the same fictitious study is sent to two groups of British Medical Journal readers and only one group is informed that it is privately funded, the group that has not been informed scored the paper significantly higher on all four metrics (importance, relevance, validity, and believability) than those who were informed, which highlights the importance of highlighting the sources of funding for studies.

Then there is the other group of academics, who are totally ethical but sadly have little idea of how the real business world actually operates. They believe that you can analyze and make determinations about what is really going on inside an organization by looking at externally available data.[37] They actually believe in the power of externally available data. These academics have rarely, if ever, worked in a truly competitive environment, where the raw thrill of competition and working hard with your colleagues to beat the competition makes it almost impossible for you to question the ethics of it. Like those soldiers in war, who have no time to think about the morality of the war they are fighting in, while trying to protect their comrades.

Social scientists, unless they have actually worked in private enterprise, which is very rare, have never experienced this. They have very little interaction with those they think they are competing with. Their main contributions are publications and there is little interaction with peers/competitors in other organizations while doing so. That is a very different dynamic to when you are going to the office every day, working with your colleagues to beat the competition, real competition; and none is more competitive and intense than the financial services industry in this regard. You want your team to win, you want your firm to win and you certainly want to impress your bosses, and in many cases that means choosing not to, or not having the time to, question their decisions.[38] It is your company against the rest. This is really where game theoretic models work at their best [Wilson (1987)], because they can help explain why individuals choose to join the most competitive teams and why they all work together to beat the competition, be it in trading equities or foreign exchange, advising clients in M&A transactions or insuring airlines. And, in the midst of all this effort it's very hard to sit back and think ethically all the time, especially since in many situations it would have been impossible to determine that the outcome might be deemed unethical ex-post. One of the criticisms leveled at most

36 A highly publicized recent example of such studies was the undercover sting by Greenpeace on two respected academics who agreed to not only write so-called independent reports denying the dangers of climate change, but also find ways to make them seem peer reviewed and avoiding the name of the sponsors being known [Goldenberg (2015)]. Olinger (2015) mentions a number of other interesting sponsored studies, including one funded by the Coca Cola Company.

37 Most companies taking over other businesses only find out what is really going on inside the target long after they have fully taken over the business. The acquisitions of Compaq [Loomis (2011)] and Autonomy [Garside (2015)] by Hewlett-Packard, with the former supposedly being the most well planned acquisition at the time, with the integration teams from both companies working for six months prior to the acquisition to make sure all was taken care of and ready for the day of the acquisition, should provide ample evidence that looking at businesses from the outside gives very little clues as to what is really going on inside. Even financial economists are aware of this so-called informational asymmetry, but it still hasn't prevented them from publishing articles on mergers and acquisitions and feeling like they have understood the dynamics of these transactions. For a critical assessment of academic analysis of mergers and acquisitions please refer to Shojai (2009).

executives is that they get so involved in their work that they simply lose objectivity. That lack of objectivity doesn't only result in bad business decisions, it can also lead to results that end up being unethical.

One of the most fascinating comments I heard about this topic was from the former CEO of one the world's largest financial services firms that faced difficulties during the recent crisis. He said that the suggestion that the executive committee got together and drew up plans to rip off clients and investors was utter nonsense. Every effort was made, with the best of intentions, to work in the best interest of clients and shareholders, but circumstances just took control of the events out of their hands. And that is absolutely right. The focus on beating the competition sometimes blinds businesses, and the people that work for them, to mistakes that cost them in the end. However, that doesn't always mean that they are intentionally trying to act unethically or immorally.

Ironically, while most of these academics acknowledge the existence of peer pressure on young people, they somehow assume that it disappears when you join a company. It doesn't.

If one were to raise a criticism it would be that under the current system, when a crisis does occur both the damages paid by the individuals involved and the number of people who are penalized is very small. As we saw during the recent financial crisis, bonuses are personalized and the losses are privatized or socialized, either paid by shareholders or in extreme cases taxpayers; very rarely paid by those who caused the losses.[39] Even if you look at some of the major ecological disasters caused by major corporations, which can put the financial crisis into context, such as the Exxon Valdez, Union Carbide's Bhopal disaster, or the BP's Deepwater Horizon explosion, you will see that in all cases the shareholders ended up paying the fines and at most the top executives lost their jobs, such as in the case of BP. None of the top executives from the parent companies went to jail or were asked to contribute to the penalties paid.[40]

Now, of course, there are some who believe that criminal penalties might dissuade some from taking excessive risks or acting fraudulently; and they could be right. However, experience shows that acting unethically is not necessarily a group exercise and there will always be people who, irrespective of the expected penalties, are willing to sail too close to the wind. Given the inability of financial services firms, or any organization for that matter, to effectively monitor the risk of the vast enterprises that they oversee [Shojai and Feiger (2010)] there will be always be people who get through the net and cause huge damage to their organizations and even industry. Examples of rogue traders who have caused unimaginable damage to the banks they worked for is proof of that. Consequently, it would be unfair to call an entire business, or industry, unethical simply because a handful of people have behaved unethically. In many of these cases, the workers, and even management, were unaware that it was taking place.

This is not to suggest that unethical behavior doesn't take place, since it certainly does. The point is that we should also take these factors into account when assessing whether firms are acting unethically. Another important fact that needs to be taken into account is the impact that the investment communities' carrots and sticks have on the behavior of management and employees of companies, especially within financial services where bonuses can in many cases dwarf salaries. There are huge compensations to be gained from meeting the targets set and serious consequences if they are missed. As Kay (2012) suggests, the situation has become exacerbated by quarterly reports, which place further pressures on the management to beat even shorter-term profitability targets.

When one looks at how the profitability of the Fortune 500 companies[41] has changed over the years it becomes clear just how great that pressure really has been, and continues to be. For example, if we compare the profitability of these companies between 1955, when the first ranking was published by Fortune, and 2015 we find that these organizations have increased their revenues and profits by multiples of 91 and 115, respectively. To put that in context, during this period, the Consumer Price Index (CPI) grew by a multiple of around 9. In fact, profitability has risen by more than twice

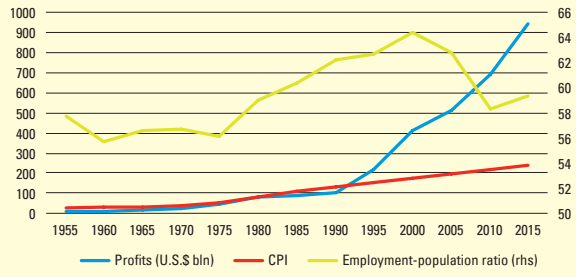
38 Silverman et al. (2014), for example, highlight how the arrogance of Joe Cassano, former president of AIG's financial products unit, had a huge role to play in the downfall of the insurance giant. The financial crises of the past 20 or so years have demonstrated how easily those who are simply making money from riding a market bubble can become extremely confident of their knowledge of the markets and intimate their colleagues, and even their regulators, from questioning their decisions, let alone challenging them. That task become significantly more difficult for those who report to them.

39 Some even accuse the U.S. government of ensuring that certain banks, in specific Goldman Sachs (whose former CEOs have been U.S. treasury secretaries a number of times, including at the time of the bail-outs), don't even experience any losses when those they contract with face difficulties during financial crises. When the U.S. government bailed out AIG with taxpayer money, which many believe was merely backdoor bailout of Goldman Sachs (as the bank received U.S.$ 12.9 billion), its objective was to prevent the insurer from defaulting on its obligations to the banks that it had sold CDSs to. And, it made sure no one experienced any losses and all were paid 100 cents on the dollar. Many believe that the government could have forced the banks to take haircuts, as had been the case when Merrill Lynch took an 86% haircut on the CDSs it had bought from Security Capital Assurance (SCA) of Bermuda just a few months earlier [Sender et al. (2010)].

40 BBC News on local management going to jail for the Bhopal disaster: http://goo.gl/6225; The Guardian on employees who will be prosecuted for the BP's Deepwater Horizon crisis: http://goo.gl/GZfNp2; The New York Times about the conviction of the Captain of Exxon Valdez being overturned and penalties paid by Exxon: http://goo.gl/qapncD

that of the S&P 500 index and three times the U.S. GDP. Figure 3, below, which compares the profitability of the Fortune 500 companies against the CPI, illustrates that profitability among these firms started to really take off from the middle- to late-1990s, around the peak of the Internet boom. The new technologies didn't only result in greater efficiencies at home, they also made it much easier to relocate certain functions to cheaper offshore locations through what became known as business service provider (BSP) boom. This assertion that outsourcing or offshoring [Irving et al. (2003)] resulted in greater profitability is somewhat supported by the falling employment to population ratio in the U.S., which fell as a result of the bursting of the Internet bubble in early 2000s, but remained lower for the subsequent 15 years. Of course, by meeting their targets the top executives have also been adequately compensated for their efforts. According to the Economic Policy Institute, between 1965 and 2014, the average annual earnings of the CEOs of major U.S. corporations increased from U.S.$832,000 to U.S.$16,316,000, resulting in the average CEO-to-worker compensation ratio to rise from 20-to-1 to almost 300-to-1 during this period [Davis and Mishel (2014)].

One could certainly raise a moral question about whether companies can afford to accept lower profits, or to at least not expect profitability to continue growing at such a severe pace, by employing more workers in their home markets or paying better salaries to those hired overseas. When you consider that the Fortune 100 companies went from generating, on average, less than U.S.$ 60 million in profits in 1955 to over U.S.$ 500 million in 1980, to around U.S.$ 2.5 billion in 2000, and just under U.S.$ 6 billion in 2015, you can ask a genuinely ethical question about how much profit can be sacrificed for the betterment of the society. Add to that the fact that around U.S.$ 2.1 trillion dollars are being kept outside the U.S. by the 500 largest U.S. companies to avoid having to pay taxes on them[42] and you can start



Sources: Fortune magazine for profitability figures and the Bureau of Labor Statistics for CPI data and Employee-population ratios

Figure 3 – The tremendous growth in the profitability of Fortune 100 companies (compared to CPI)

to appreciate just how much flexibility corporations could have to keep more people employed or pay more to those already on their payroll. But, of course, that does not happen, and companies that do try to do so will be harshly punished by the markets; markets that have got accustomed to large profitability growth rates.

Having said that, the fact that companies don't cut executive compensation to what would be more morally acceptable, or don't accept lower profitability to improve the lives of their workers, does not necessarily make them unethical. Unless, of course, all companies that act in the best interests of their shareholders are unethical. Furthermore, one cannot say for sure that these companies achieved such profits through unethical means, since none of us have access to the true state of affairs within these organizations. As mentioned before, the mere fact that companies move jobs overseas doesn't make them unethical, since they have been effective in lifting many people in poorer countries out of poverty. More importantly, experience has shown that both the management and their shareholders will do everything in their powers to fight regulations or guidelines that might impede their gains, something that both regulators and governments have learned over the years. Consequently, despite their best intentions, academics will struggle to find solutions that would force businesses to act in ways that they would deem ethical, assuming they can all agree on what it is, unless it suits businesses to do so.

The lessons of the Enron crisis and Lehman bankruptcy

Similar to the responses to the allegations that U.N. weapons inspectors were actually employees of the CIA opened the door to the use of "why is that news?", the Enron collapse also had ramifications beyond what many thought possible.

After the initial period of shock and the introduction of new regulations, companies started looking at how such an event should never happen again. And, by that, I don't mean to ensuring that they never act unethically again. I am referring to preventing the businesses from being shut down in case they do. They were trying to learn from Western governments how they should react to the revelations of bad news; the kind of bad news that could result in another Enron-type implosion. Of course, governments in less democratic countries, like Russia and China, had silenced critics and bad news for decades, but you couldn't exactly apply their strategies in the West.

41 These data is obtained from Fortune 500 rankings, which are based on revenues, available from the archives of Fortune magazine: http://goo.gl/GY882N

42 Reuters: http://goo.gl/Nxy32z

The response to bad news could certainly incorporate the "why is that news?" strategy, and many have used it. When a company or financial institution has been caught doing something illegal, the first response seems to be why is that news, everyone knows everyone else is doing it, and it seems in most cases they were right. Most banks were manipulating LIBOR,[43] most were mis-selling payment protection to those who didn't need it,[44] most were selling complex derivatives strategies to mum and pop businesses,[45] and so on.

An even more important lesson that organizations learned was to make sure that if they are ever caught doing something unethical, or even illegal, that the maximum penalty would be fines and nothing more [Warren (2016)]. And we have seen that happen a number of times as well. There have been a few episodes in recent years that have not been too different to what took place at Enron, but neither the auditor nor the client were closed down. In some cases, they were just forcibly sold to another company. But, more importantly, no one went to prison.

With regards to managing the public relations damage, companies have learned a lot. However, unlike what your reputational management textbooks suggest, which is to admit your mistakes immediately and take the necessary hit there and then, these companies have learned to actually say nothing. They literally close access to journalists, similar to how governments respond. They put all their efforts into dealing with the regulators and waiting for the news to move onto another topic. Given, as was discussed earlier, they have also learned to control the narrative, they know that their PR teams know how to manage the crisis and ensure it's not dwelled on too much.

According to Lewis et al. (2008), commercial enterprises have surpassed the U.K. government in terms of the share of PR generated news that news organizations present in the U.K.; 38% of press and 32% of broadcast, as compared to government's 21% for the press and 39% for broadcast media. Quite certainly it is not too different from that in the U.S. This ensures that the journalists don't dwell too much on bad news about these companies, as we have seen numerous times.

In addition, companies and governments have also learned how to use the comments sections of online news pages to justify the actions that have been criticized for in articles; what is known as astroturfing. Ironically, it seems that term was coined in 1985 by the then-U.S. Senator Lloyd Bentsen when he said, "a fellow from Texas can tell the difference between grass roots and AstroTurf... this is generated mail." The senator was describing a "mountain of cards and letters" sent to his office to promote insurance industry interests [Kolivos and Kuperman (2012)]. And I am sure most of us have read comments that just seem outrageous, but they do dampen the impact of the other negative comments from the readers about the accusations.[46]

Finally, institutions learned from the Lehman bankruptcy that you should do your utmost to become as systemically large as possible. Then, no matter what you do, you will be protected. Despite the huge damage that the Lehman's bankruptcy caused the global financial markets, and even the global economy, most institutions have learned that had it been a much larger institution it would not have been sacrificed in the way that it was.

If Ross Sorkin's (2010) accounts of the events that took place at the New York Federal Reserve in late 2008 are indeed correct, Lehman Brothers was sacrificed to save Merrill Lynch, by selling, or creating the environment to sell, the latter, which was much bigger, to Bank of America, rather than the former despite previous negotiations between Lehman Brothers and Bank of America.

You don't need to be a social scientist to work out how fast most institutions have grown since the global financial crisis. According to Big4.com, the Big 4 increased their employee numbers from 600,000 in 2010, to 756,000 in 2014. The data given by all four puts the figure at almost 820,000 for 2015. This means that they are now way too-big-to-fail. Likewise, for most of the major banking institutions; they have also become too-big-to-fail.

Hence, the lessons that these institutions have learned from the Enron and Lehman Brothers collapse is that if you make yourself too big to be closed down, you will not be. Just think of the recent problems that both BP and Volkswagen have faced. BP for negligence and Volkswagen for fraud. But, neither business was, or will be, closed down. They will pay their fines and just move on. All that happened to the CEO of Volkswagen, at least for now, is that he has lost his job. Both companies, however, will remain in business for years to come. Add to that, the lessons they have learned in influencing how media cover such crises and for how long, and you will see that businesses have certainly learned how to react to crisis in a way that they survive without too much damage.

43 Reuters: http://goo.gl/GChr8q
44 The Guardian: http://goo.gl/WO0BTD
45 Daily Telegraph: http://goo.gl/4jPFh
46 ComputerWorld: http://goo.gl/D8DI58

CONCLUSION

This article has presented arguments in support of the proposition that business ethics is indeed an oxymoron, and suggests that while businesses don't necessarily set out to act unethically, when ethics and profitability collide the latter tends to win most of the time.

It also highlights the fact that it is almost impossible to describe what being ethical as a business actually means, as ethics is in the eye of the beholder. What seems unethical to one group of individuals might be deemed completely ethical from the perspective of another. That despite what most academics think, it is not possible to rank businesses based on their ethicality and that unethical behavior is only determinable after the effect. It is only possible to identify unethical behavior when something goes horribly wrong and you can point to specific unethical actions that lead to the crisis.

This article also describes why a combination of falling journalistic standards, demand for and availability of 24-hour news, revenues increasingly coming from clicks on articles, and the growing power of PR has resulted in an environment where with the exception of truly catastrophic circumstances businesses can control the narrative.

Businesses have also learned from the experiences of Enron and Lehman Brothers that to protect themselves from closure they need to become as large as possible, ensure that their industry is highly concentrated, and to wait for the bad news to just pass. Given that their control of the media is growing, the speed with which bad news passes has increased.

From a purely academic perspective, the studies that find associations between business ethics and profitability are overlooking the simple fact that businesses can never be understood or analyzed from the outside in by solely relying on external data. Numerous failed acquisitions are proof of that. Academics need to accept that there are certain subjects that are simply impossible to obtain adequate information and data on to make meaningful determinations of the environment and provide prescriptive guidance on how to improve it. Business ethics is in my opinion one of those subjects. To try to understand the dynamics of so many people with different ethical beliefs all focused on making their businesses a success is a task too far and academics have to accept that. It is much more honest to accept that than accuse those who simply state that facts of not having a clear understanding of the circumstances.

My aim with this article is to explain to current students of management, and future managers of businesses, that while their aim should always be to be as ethical as possible that they should accept the world as it is and focus their efforts on making their careers as successful as possible without being unrealistically influenced by their professors who propose solutions that are neither realistic nor practical.

I am certain that those who advocate, or teach and research, business ethics as a discipline to be taught at business schools would take issue with my perspectives and genuinely believe that steps can be taken to make businesses act more ethically. I am not so sure, and I have made my case in this article. Furthermore, while some acknowledge the challenge is a big one they believe that doing something is better than doing nothing. My response is that the damage caused by such perspectives is significantly greater than many perceive, and could result in circumstances that are much less beneficial than actually doing nothing. These perspectives are no different to those who advocate that we should teach finance students about asset pricing models, or other theoretical finance topics, that have been completely discredited [Colander et al. (2009); Blommestein (2009)] and have no relationship to how the financial services industry actually operates, simply so that we are seen to be teaching them something about asset pricing rather than to not teach them anything at all. In my opinion, if we were to discard literally all of the models that we teach students of finance today we might have a better chance of developing models that can actually be used in business than we do today. The need for articles to be peer reviewed requires academics to accept the foundations of finance as gospel, resulting in a neverending series of articles that are awarded academic rewards [Shojai and Feiger (2011)] but are of no practical benefit to the students who wish to apply them in their work place. Worse, they are also doing a huge disservice to the future employers of those students by requiring them to retrain their new recruits in the practical aspects of what they were taught at business schools; a process many liken to unlearning and relearning the business discipline.

Interestingly, while the academic community is quite comfortable in leveling accusations against financial institutions and regulators for not heeding the lessons of the latest financial crisis and taking steps to avoid repeating the same mistakes, they are overlooking the fact that they have also missed a great opportunity to question the foundations upon which many of their theorems are based and to reevaluate the contributions they are making towards the discipline and the society at large. They are once again working away in their ivory towers, publishing article after article with little to no relevance to the realities of the world of business and making little or no effort to work closer with those practitioners who are supposed to put their ideas into practice. Sadly, it seems that neither the academics nor the financial community have learned much from the recent crisis and we are back to business as usual.

REFERENCES

- Agnew, H., 2015a, "Professional services: Accounting for change" Financial Times, August 27, http://goo.gl/sSKz59
- Agnew, H., 2015b, "Big four auditors extend reach into consultancy," Financial Times, March 8, http://goo.gl/6ehkkn
- Ashbaugh, H., R. LaFond, and B. W. Mayhew, 2003, "Do nonaudit services compromise auditor independence? Further evidence," The Accounting Review 78:3, 611-639
- Bernardi, R., M. Melton, S. Roberts, and D. Bean, 2008, "Fostering ethics research: an analysis of the accounting, finance and marketing disciplines," Journal of Business Ethics 82(1), 157-170
- Black, J., and K. Anderson, 2013, "Creating an ethical framework for the financial services industry," working paper, London School of Economics and Political Science
- Blommestein, H., 2009, "The financial crisis as a symbol of the failure of academic finance? (A methodological digression)," Journal of Financial Transformation 27, 3-8
- Bowcott, O., 2015, "UN accused of 'gross failure' over alleged sexual abuse by French troops," The Guardian, December 17, http://goo.gl/1l5jbj
- Bowers, S., and J. Treanor, 2015, "Lords could derail Osborne's plans to soften rules on banker accountability," The Guardian, December 13, http://goo.gl/965pMB
- Bragg, R., 2002, "Enron's collapse: workers; workers feel pain of layoffs and added sting of betrayal," The New York Times, January 20, http://goo.gl/Zf861d
- Brown, J. R., 1995, "The "great fall": the consequences of repealing the Glass-Steagall Act," Stanford Journal of Law, Business, and Finance 2:1, 129-146
- Carlin, B., and S. Gervais, 2009, "Work ethic, employment contracts, and firm value," Journal of Finance 64:2, 785-821
- Cassidy, J., 2003, "The investigation: How Eliot Spitzer humbled Wall Street," The New Yorker, April 7, http://goo.gl/U0PKZT
- Causholli, M., D. J. Chambers, and J. L. Payne, 2014, "Future non-audit service fees and audit quality," Contemporary Accounting Research 31, 681-712
- Causholli, M., D. J. Chambers, and J. L. Payne, 2015, "Does selling non-audit services impair auditor independence? New research says, 'yes'," Current Issues in Auditing 9:2, 1-6
- Cave, T., and A. Rowell, 2014, "The truth about lobbying: 10 ways big business controls government," The Guardian, March 12, http://goo.gl/9Oq2UG
- Chaudhry, S., S. Schroter, R. Smith, and J. Morris, 2002, "Does declaration of competing interests affect readers' perceptions? A randomised trial," British Medical Journal 325, 1391-1392
- Clarke, T., 2014, "The impact of financialisation on international corporate governance: the role of agency theory and maximising shareholder value," Law and Financial market Review 8:1, 39-51
- Coffee, J. C., 2006, Gatekeepers: the professions and corporate governance, New York, NY: Oxford University Press
- Collins, D., 2012, Business ethics: how to design and manage ethical organizations, Hoboken, NJ: John Wiley & Sons
- Colander, D., H. Föllmer, A. Haas, Mi. Goldberg, K. Juselius, A. Kirman, T. Lux, and B. Sloth, 2009, "The financial crisis and the systemic failure of academic economics," Kiel Institute for the World Economy, working paper no., 1489
- Crawford, C., 2011, "The repeal of the Glass- Steagall Act and the current financial crisis," Journal of Business & Economics Research 9:1, 127-134
- Davis, A., and L. Mishal, 2014, "CEO pay continues to rise as typical workers are paid less," Economic Policy Institute, June 12, http://goo.gl/HZSLgM
- DeFond, M. L., K. Raghunandan, and K. R. Subramanyam, 2002, "Do non-audit service fees impair auditor independence? Evidence from going concern audit opinions," Journal of Accounting Research 40, 1247-1274
- Drutman, L., 2015, The business of America is lobbying: how corporations became politicized and politics became more corporate (studies in postwar American political development), Oxford University Press
- Dunkley, E., 2015, "UK draws line under 'banker bashing' after scrapping assessment," Financial Times, December 30, http://goo.gl/4cixqI
- Economist, 2014, "The dozy watchdogs," The Economist, December13, http://goo.gl/wMK50m
- Fifield, A., 2013, "Contractors reap $138bn from Iraq war," Financial Times, March 18, http://goo.gl/6BYKQA
- Fleming, S., and A. Smith, 2014, "Big Four keep their grip on company audits," Financial Times, February 19, http://goo.gl/sXuC6s
- Francis, J. R., and B. Ke, 2006, "Disclosure of fees paid to auditors and the market valuation of earnings surprises," Review of Accounting Studies 11, 495-523
- Friedberg, M., B. Saffran, T. Stinson, W. Nelson, and C. Bennett, 1999, "Evaluation of conflict of interest in new drugs use in oncology," Journal of the American Medical Association 282, 1453-57
- Friedman, H. H., and L. W. Friedman, 2010, "Lessons from the global financial meltdown of 2008," Journal of Financial Transformation 28, 47-56
- Garside, J., 2015, "Hewlett-Packard sues Autonomy founder for damages," The Guardian, March 31, http://goo.gl/MZMQhx
- Gerakos, J., and C. Syverson, 2015, "Competition in the audit market: policy implications," Journal of Accounting Research 53:4, 725-775
- Glater, J. D., 2003, "Deloitte, in reversal, will retain its consulting operation," The New York Times, March 29, http://goo.gl/6N2s64
- Goldenberg, S., 2015, "Greenpeace exposes sceptics hired to cast doubt on climate science," The Guardian, December 8, http://goo.gl/j8wf2b
- Goldwasser, D. L., 2002, "The accounting profession's regulatory dilemma," CPA Journal 72:5, 8-13
- Griffin, J. J. and J. F. Mahon, 1997, "The corporate social performance and corporate financial performance debate: twenty-five years of incomparable research," Business & Society 36:1, 5-31
- Gupte, P., 2005, "Arthur Laffer: corporate social responsibility detrimental to stockholders," The Sun New York, January 19
- Habib, A., 2012, "Non-audit service fees and financial reporting quality: a meta-analysis," Abacus 48:2, 121-157
- Hamilton-Martin, R., 2015, "Ostracised, sacked … and even arrested: the fate of whistleblowers at the UN," The Guardian, September 14, http://goo.gl/BCu7w7
- Hanrahan, M., 2014, "Elizabeth Warren slams Citigroup's influence on 'Cromnibus' Bill, blasts bid to water down Dodd-Frank," International Business Times December 13, http://goo.gl/8JWgJS
- Heng, J., D. Blackmore, and J. Short, 2015, "Market share analysis: consulting services, worldwide, 2014," Gartner
- House of Lords, 2011, "'Complacency' of auditors contributed to financial crisis," House of Lords Economic Affairs Committee, http://goo.gl/v8QtM8
- Ingram, M., 2015, "Print readership is still plummeting, and paywalls aren't really helping" Fortune, June 1, http://goo.gl/4AJ7zc
- Irving, B., S. Shojai, and S. Gupta, 2003, "Discovering the endgame in the offshore debate," Journal of Financial Transformation 8, 103-112
- Jackson, P., 2014, Risk culture and effective risk governance, Risk Books
- Jelveh, Z., B. Kogut and S. Naidu, 2014, "Economists aren't as nonpartisan as we think," Political Economy, FiveThirtyEight, December 8, http://goo.gl/nstC5r
- Jensen, M., and W. Meckling, 1976, "Theory of the firm: managerial behavior, agency costs and ownership structure," Journal of Financial Economics 3:4, 305-360
- Johnson, S., and J. Kwak, 2011, 13 Bankers: The Wall Street takeover and the next financial meltdown, Vintage Books
- Kaplan, R. L., 2004, "The mother of all conflicts: auditors and their clients," Journal of Corporation Law 29:2, 363-383
- Karnani, A., 2010, "The case against corporate social responsibility," The Wall Street Journal, August 23, available at: http://goo.gl/ckJZ9t
- Kay, J. A., 2011, "Kay review of UK equity markets and long-term decision making," Department for Business, Innovation & Skills
- Krimsky, S., 2003, Science in the private interest, Lanham, MD: Rowman and Littlefield
- Krishnamurthy, S., J. Zhou, and N. Zhou, 2006, "Auditor reputation, auditor independence, and the stock-market impact of Andersen's indictment on its client firms." Contemporary Accounting Research 23, 465-490
- Krishnan, J., H. Sami, and Y. Zhang, 2005, "Does the provision of nonaudit services affect investor perceptions of auditor independence?" Auditing: A Journal of Practice & Theory 24, 111-135
- Kolivos, E., and A. Kuperman, 2012, ""Web of lies – legal implications of astroturfing." Keeping good Companies 64:1, 38-41

- Kull, S., C. Ramsay, S. Subias, E. Lewis, and P. Warf, 2003, "Misperceptions, the media and the Iraq war," PIPA/Knowledge Networks, University of Maryland, http://goo.gl/Scsn3P
- Lee, D. R., 2010, "Why businessmen are more honest than preachers, politicians, and professors," The Independent Review 14:3, 435-444
- Leventhal, R., 2013, "Embedded in Iraq as the war began – the most challenging assignment of my career," Fox News, March 19, http://goo.gl/pX96IH
- Levinson, C., 2015, "Accounting industry and SEC hobble America's audit watchdog," Reuters, December 16, http://goo.gl/1SWdL9
- Lewis, J., A. Williams, B. Franklin, J. Thomas and N. Mosdell, 2008, "The quality and independence of British journalism: tracking the changes over 20 years," report prepared by the Cardiff School of Journalism, Media and Cultural Studies of Cardiff University
- Lexchin, J., 2012, "Those who have the gold make the evidence: how the pharmaceutical industry biases the outcomes of clinical trials of medications," Science and Engineering Ethics 18:2, 247-261
- Lexchin, J., L. A. Bero, B. Djulbegovic, O. Clark, 2003, "Pharmaceutical industry sponsorship and research outcome and quality: systematic review," British Medical Journal 326, 1167-70
- Lim, C., and H. Tan, 2008, "Non-audit service fees and audit quality: the impact of auditor specialization," Journal of Accounting Research 46:1, 199-246
- Loomis, C. J., 2011, "Why Carly's big bet is failing," Fortune, August 21, http://goo.gl/KWyWxu
- Lucca, D., A. Seru, and F. Trebbi, "Worker flows in banking regulation," Journal of Monetary Economics 65, 17-32
- Masters, B., 2012, "Enter the revolving regulators," Financial Times, April 23, http://goo.gl/9tvcOj
- Monbiot, G., 2015, "Let's not fool ourselves. We may not bribe, but corruption is rife in Britain," The Guardian, March 18, http://goo.gl/nb7FVC
- Margolis, J. D., and J. P. Walsh, J.P., 2001, People and profits? The search for a link between a company's social and financial performance," Mahwah, N.J.: Lawrence Erlbaum Associates
- McConnell, D. K., and G. Y. Banks, 2003, "How Sarbanes-Oxley will change the audit process," Journal of Accountancy, September 1
- Moosa, I., 2013, "The failure of financial econometrics: "stir-fry" regressions as a con job," Journal of Financial Transformation 34, 43-49
- Nelsen, A., 2015, "UK lobbying for even weaker EU air pollution laws, leaked papers show," The Guardian, August 12, http://goo.gl/DcdGZ2
- Newman, 2015, "UN under global fire for persecution of whistleblowers," New American, October 8, http://goo.gl/XtrRnA
- Oborne, P., 2015, "Why I have resigned from the Telegraph," Open Democracy UK, 17 February, https://goo.gl/6UDbCX
- Olinger, D., 2015, "CU nutrition expert accepts $550,000 from Coca-Cola for obesity campaign," The Denver Post, December 27, http://goo.gl/GGcRBB
- Page, S., 2015, "Ben Bernanke: More execs should have gone to jail for causing Great Recession," USA Today, November 13, http://goo.gl/ECygN2
- Picchi, A., 2015, "The watering down of Wall Street reform," CBS news, January 14, http://goo.gl/eMM1L2
- Popper, N., and P. Eavis, 2014, "Secret Goldman Sachs tapes put pressure on New York Fed," The New York Times, October 2, http://goo.gl/856nZ9
- Power, M., S. Ashby, and T. Palermo, 2013, "Risk culture in financial organisations," research report, London School of Economics and Political Science
- Prem, S., 2013, "The uncompetitive culture of auditing's big four remains undented," The Guardian, February 23, http://goo.gl/TaLZA8
- Quick, R., and B. Warming-Rasmussen, 2015, "An experimental analysis of the effects of non-audit services on auditor independence in appearance in the European Union: evidence from Germany," Journal of International Financial Management & Accounting 26:2, 150-187
- Rapoport, M., 2010, "Role of auditors in crisis gets look," The Wall Street Journal, December 23, http://goo.gl/9JQlcy
- Resnik, D., 2007, The Price of truth, New York: Oxford University Press
- Riffkin, R., 2015, "Americans' trust in media remains at historical low," Gallup, September 28, http://goo.gl/6RwzuM
- Rosen, J., 2009, "Iraq war 6 years later: success but no 'victory'" Fox News, March 19, http://goo.gl/9LKbMP
- Ross Sorkin, A., 2010, Too big to fail: the inside story of how Wall Street and Washington fought to save the financial system-and themselves, Penguin Books
- Ross Sorkin, A., 2011, "Revolving door at S.E.C. is hurdle to crisis cleanup," The New York Times, August 1, http://goo.gl/dCnSV
- Sants, H., 2010a, "Do regulators have a role to play in judging culture and ethics?" Speech to Chartered Institute of Securities and Investments Conference, London, UK.
- Sants, H., 2010b, "Can culture be regulated?" Speech to Mansion House Conference on values and trust, London, U.K.
- Schneider, A., B. K. Church., and K. M. Ely, 2006, "Non-audit services and auditor independence: a review of the literature," Journal of Accounting Literature 25, 169-211
- Sender, H., G. Tett, and F. Guerrera, 2010, "Finance: The hindered haircut," Financial Times, January 26, http://goo.gl/eIJdYK
- Shojai, S., 2009, "Economists' hubris – the case of mergers and acquisitions," Journal of Financial Transformation 26, 4-12
- Shojai, S., and G. Feiger, 2009, "Economists' hubris: the case of asset pricing," Journal of Financial Transformation 27, 9-13
- Shojai, S., and G. Feiger, 2010, "Economists' hubris – the case of risk management," Journal of Financial Transformation 28, 25-35
- Shojai, S., and G. Feiger, 2011, "Economists' hubris – the case of award winning finance literature," Journal of Financial Transformation 31, 9-17
- Sikka, P., 2009, "Financial crisis and the silence of the auditors," Accounting, Organizations and Society 34:6-7, 868-873
- Silverman, S. B., R. E. Johnson, N. McConnell, and A. Carr, 2014, "Arrogance: a formula for leadership failure," Society for Industrial and Organizational Psychology, July 12, http://goo.gl/yoc3u
- Singleton, S., 2002, "Name game: Deloitte Consulting becomes Braxton," TechRepublic, August 6, http://goo.gl/6SfBKa
- Sismondo, S., 2008, "Pharmaceutical company funding and its consequences: a qualitative systematic review," Contemporary Clinical Trials 29, 109-113
- Stelfox, H., G. Chua, K. O'Rourke, and A. Detsky, 1998, "Conflict of interest in the debate over calcium channel antagonists," New England Journal of Medicine 338, 101-106
- SustainAbility, 2001, "Buried treasure: uncovering the business case for corporate sustainability," SustainAbility and UNEP
- SustainAbility, 2002, "Developing value: the business case for sustainability in emerging markets," SustainAbility, IRC and Ethos Institute
- Trowbridge, A., 2013, "NSA spying: Ally anger justified?" CBS News, July 3, http://goo.gl/FkOAHU
- Visser, W., 2010, "The age of responsibility: CSR 2.0 and the new DNA of business," Journal of Business Systems, Governance and Ethics 5:3, 7-22
- Verschoor, C. C., 1998, "A study of the link between a corporation's financial performance and its commitment to ethics," Journal of Business Ethics 17:13, 1509-1516
- von Hayek, F. A., 1974, "The pretence of knowledge," Lecture to the memory of Alfred Nobel, December 11, http://goo.gl/HDLSXd
- Warren, E. A., 2016, "Rigged justice: 2016; how weak enforcement lets corporate offenders off easy," United States Senate, http://goo.gl/elewlk
- Webley, S., and E. More, "Does business ethics pay? Ethics and financial performance," Institute of Business Ethics, http://goo.gl/LGExxB
- Wilson, R. B., 1987, "Game-theoretic analyses of trading processes," in Bewley, T. (ed.), Advances in economic theory: Fifth World Congress, Cambridge UK: Cambridge University Press

The Dodd-Frank Act Five Years Later: Are We More Stable?[1]

Todd J. Zywicki – George Mason University Foundation Professor of Law and Executive Director, Law & Economics Center, George Mason University School of Law

Abstract

In response to the global financial crisis, in 2010 Congress enacted the Dodd-Frank Financial Reform Act, which was ostensibly designed to "end" the problem of too-big-to-fail banks and otherwise reform and modernize the American financial system. I, and others, have elsewhere considered the impact that Dodd-Frank has had on the financial services industry, banking industry, and consumers. This article focuses on a larger long-term influence of Dodd-Frank and the financial crisis: the impact on the rule of law and freedom. Although Dodd-Frank and the regulations enacted under it could, in theory, be repealed or amended in the future, it will be far more difficult to reverse the impact of Dodd-Frank and the financial crisis on the rule of law, constitutional government and individual freedom and protection from arbitrary government.

1 This article is based on testimony presented to the U.S. House of Representatives, Committee on Financial Services (Sept. 17, 2015).

INTRODUCTION

In response to the global financial crisis of 2008, in 2010 Congress enacted the Dodd-Frank Financial Reform Act, ostensibly to prevent future financial crises, eliminate the problem of too-big-to-fail (TBTF) financial institutions, and increase consumer protections for financial services consumers.

As Dodd-Frank celebrates its fifth birthday, it remains highly controversial and many question whether it has on net been harmful or beneficial to the stability and efficiency of the American financial services sector, the larger economy, and consumers. Less examined, however, is the long-run impact of Dodd-Frank on individual freedom, the rule of law, and constitutional government. In theory, the adverse effects of Dodd-Frank on the economy and financial system can be ameliorated by future legislation to repeal or amend the law. Efforts to reverse the long-run impact on the rule of law and individual freedom, however, will be more difficult. And, in the long run, these impacts may be more important than the direct economic effects of the law.

Freedom and an effective financial services system go together. Freedom to gain access to capital to start and grow a business, freedom to buy a home and provide for your family's financial security, freedom to choose those whom you entrust with your hard-earned money provide the means for pursuing the American dream.[2]

This article reviews the long-run impact of Dodd-Frank on individual freedom and the rule of law, providing a cautionary tale for the future as well as signals for concrete reforms that Congress and a new President should consider going forward.

THE IMPACT OF DODD-FRANK ON FREEDOM: THE REGULATORY BURDEN

In the new world of Dodd-Frank, the success of a financial institution is no longer determined by its ability to be among the best providers within a highly competitive market. Instead, it is determined by which institutions can best wind their way through the labyrithian halls of Congress and the Federal Reserve Board.

According to one widely-cited estimate, Dodd-Frank requires 398 new rulemakings by federal agencies[3] and as of July 2014 (when one-quarter of the rulemakings were still left to be completed) Dodd-Frank was estimated to have imposed U.S.$21.8 billion and 60.7 million paperwork hours in compliance costs.[4] Projecting forward, it is estimated by one economist that over the next 10 years the full compliance costs of Dodd-Frank will result in U.S.$895 billion in reduced GDP or U.S.$3,346 per working-age person.[5] Furthermore, these compliance cost estimates do not include all of the costs and burdens of complying with the various guidances, informal actions, and other measures that federal regulators impose on financial institutions and their customers.

But to only consider the economic costs of Dodd-Frank means that another more intangible cost is ignored, namely that Americans are less free as a result of Dodd-Frank and what it has spawned. In particular, the financial crisis and the legislation and regulation that has followed in its wake have weakened the rule of law, centralized vast amounts of authority in the hands of unaccountable political bureaucracies, unleashed arbitrary regulatory discretion, and empowered interest groups beyond any time in American history. Moreover, not only did the unleashing of political discretion help to create and worsen the last crisis, by entrenching rather than limiting political discretion, Dodd-Frank and the regulatory norms it embodies, has created moral hazard that is laying the foundation for the next financial crisis.

THE FINANCIAL CRISIS AND THE DECLINE OF THE RULE OF LAW

The recent financial crisis reveals four lessons that highlight the importance of upholding the rule of law during crises in order to preserve individual freedom. First, adherence to the rule of law during the crisis is crucial to allow the economy to restore coordination after a period of economic dislocation. Second, adherence to the rule of law during the crisis is necessary to restrain opportunism by politicians and special interests tempted to use the opportunity presented by the crisis to piggyback satisfaction of their own narrow – and often unrelated – interests. Third, once discretion and political favoritism are unleashed during the crisis, history tells us that the dissipation of the crisis does not bring with it a restoration of the

2 See Durkin, T. A., G. Elliehausen, M. E. Staten, and T. J. Zywicki, 2014, Consumer credit and the American economy, Oxford University Press

3 Polk, D., 2015, Dodd-Frank Progress Report, http://www.davispolk.com/Dodd-Frank-Rulemaking-Progress-Report/. Romano, R., 2014, Dodd-Frank's regulatory morass, Penn Program on Regulation RegBlog, November 10, available in http://www.regblog.org/2014/11/10/romano-dodd-frank-consequences/.

4 Winkler, A., B. Gitis, and S. Batkins, 2014, Dodd-Frank at 4: more regulation, more regulators, and a sluggish housing market, July 15, available in http://americanactionforum.org/research/dodd-frank-at-4-more-regulation-more-regulators-and-a-sluggish-housing-mark.

5 Holz-Eakin, D., 2015, "The growth consequences of Dodd-Frank," American Action Forum, May 6, available in http://americanactionforum.org/research/the-growth-consequences-of-dodd-frank.

rule of law. Instead there is a sort of "ratchet effect," by which the power seized during the crisis is entrenched in the post-crisis regulatory regime. Finally, once discretion and the government's power to pick winners and losers arbitrarily is entrenched, this institutional framework creates moral hazard for policians and special interests that creates the conditions for the next crisis, which will probably be met by similar means.

The world of Dodd-Frank exemplifies this progression. As a result of Dodd-Frank's heavy and vague regulatory regime, the law is not just hampering the economy but adversely impacting the ability of Americans to gain access to capital to pursue their dreams in life. Access to capital is the lifeblood of the ability to plan for one's financial future, buy a home, or open a bank account. Thus, not only is Dodd-Frank having an adverse economic impact, the freedom to pursue one's dreams in life are being crushed under the thicket of costly and arbitrary rules and a regulatory system so complex that only well-lawyered multi-billion dollar banks can survive. On issues ranging from which financial insitutions are considered TBTF to the loan terms of your new car, a handful of unelected Washington bureaucrats are prying into household and small business finances to make those decisions for you.

Why the rule of law matters during a financial crisis

To make a loan, a bank must be able to do two things.[6] It must be able to price the risk of the loan accurately in light of its risk of loss, such as by adjusting the interest rate, downpayment, or other terms of the loan. If the lender cannot price the risk of loss accurately, then the lender must reduce its risk exposure, either by limiting those to whom it lends (such as refusing to lend to higher-risk borrowers) or by lending less to the same people (such as by reducing available credit lines).

Economic uncertainty interferes with the ability of lenders and borrowers to accurately assess the full risk and cost of making loans and conducting commerical activity. As a result, economists have uniformly found that adherence to the rule of law is an essential condition for economic prosperity, democratic governance, and civil liberties.[7] Moreover, the rule of law serves as a barrier to government corruption and rent-seeking by powerful special interest groups. By ensuring equal and transparent treatment of everyone, the rule of law constrains the discretion to arbitrarily pick winners and losers that provides the engine and incentives for political corruption.[8]

Adherence to the rule of law is especially important during periods of economic dislocation, such as during the financial crisis. During such times, billions of decentralized individual decision-makers need to reestablish coordination of their affairs, to make decisions to work, invest, hire, and the like. When other elements of the economic system are in greater flux, adherence to the bedrock predictability of the rule of law takes on special institutional significance.

Instead, the federal government responded erratically and unpredictably during the financial crisis, thereby exacerbating uncertainty and confusion, such as by deciding to bail out Bear Stearns but not Lehman Brothers and attaching different and arbitrary conditions to each subsequent bailout. In so doing, the government's departure from rule of law values worsened the financial crisis and continues to hamper the economy's return to economic stability. As David Skeel has shown, one reason for the catastrophic nature of Lehman Brothers' failure was that the firm – counting on a government bailout – rejected a merger offer as insufficiently generous. [9] Indeed, as several prominent scholars have observed, it likely was not Lehman's failure that spooked the markets, but rather Treasury Secretary Hank Paulson's panicked response to Lehman's failure.[10] As noted by Richard Kovacevich, CEO of Wells Fargo during the financial crisis, prior to TARP and a month after the Lehman bankruptcy, "markets had declined but were still behaving reasonably well, except for those financial institutions that were having liquidity issues."[11] It was only when TARP was announced – and critically, when the government strong-armed all big banks into taking bailout money, even those that didn't want it – that "isolated liquidity issues turned into a tsunami impacting all banks and all industries." In short, the TARP created the very panic that bailout apologists contend that the TARP supposedly stemmed.[12]

Political opportunism and the rule of law

Adherence to the rule of law is especially important during periods of crisis because that is the time when potential for political opportunism by politicians and interest groups is most dangerous. The actual operation of the government's response to the financial crisis shows the reality of how politicians and special interests use

6 See Zywicki, T. J., 2012, "Economic uncertainty, the courts, and the rule of law, 35 Harvard Journal of Law and Public Policy, 195

7 See Zywicki, T. J., 2003, "The rule of law, freedom, and prosperity," 10 Supreme Court Economic Review, 1

8 Id.

9 Skeel, D. A., 2011, The new financial deal: understanding the Dodd-Frank Act and its (unintended) consequences, Wiley

10 Wallison, P. J., 2013, Bad history, worse policy: how a false narrative about the financial crisis led to the Dodd-Frank Act, AEI Press; Taylor, J. B., 2009, Getting off track: how government action and interventions caused, prolonged, and worsened the financial crisis, Hoover Press.

11 Kovacevich, R. J., 2014, "The financial crisis: why the conventional wisdom has it all wrong," 34(1) Cato Journal

12 Zywicki, T. J., 2015, "The rule of law during times of economic crisis," August 26, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2651893.

power and political connections unrestrained by the rule of law for their benefit.

Consider the infamous TARP program, which was authorized to provide a temporary bailout for illiquid banks that needed short-term help, but not insolvent banks. The task of distinguishing between illiquid and insolvent banks, however, was not an easy one and required great discretion by those making those decisions. Several economists have subsequently studied how bailout funds were allocated and they have uniformly reached the same conclusion: that bailout funds were directed to banks with "political clout, not those most in need of liquidity."[13] Banks that lobbied the most, contributed the most money to political campaigns, or had former banking regulators or Treasury Department officials on their boards of directors were significantly more likely to receive bailout funds than less-politically connected banks, even where those other banks ostensibly met the TARP's requirements more closely.[14]

Similarly, as I have discussed elsewhere, the entire taxpayer loss in the illegal diversion of TARP funds to General Motors and Chrysler is attributable to preferential treatment provided in those bankruptcy proceedings to the United Auto Workers and various other politically-powerful labor unions that had nothing to do with furthering the financial recovery of those companies.[15] Moreover, the government's intervention in the auto bailouts provided a field day for political opportunism. Politicians used the strings supplied by taxpayers' largesse to influence ordinary business decisions ranging from preventing the closure of particular obsolete manufacturing facilities that happened to be located in a particular politician's electoral district, to the identity of suppliers of raw materials, to providing secret financial incentives for Fiat to manufacture "green" cars after the government ordered Chrysler to be given away for free to the Italian automaker.[16] Although American automakers have returned to profitability since they were bailed out, this has been despite the government's influence, as low gasoline prices have driven a boom in sales of pickup trucks and other larger vehicles, not the small cars urged by government central planners during the bailout process.[17]

The case study of the auto bailouts also provides a particularly illuminating illustration of why upholding the rule of law matters to both short-term and long-term freedom and prosperity. The primary losers from the government's intervention in the Chrysler bankruptcy case were holders of Chrysler's secured corporate bonds, including the Indiana state teachers and police retirement funds. While secured creditors typically would be paid in full before unsecured creditors, in that case secured creditors received only 29 cents on the dollar while UAW's underfunded health-care VEBA plans received over 40 cents on the dollar.[18]

But the full cost of the government's intervention was not just the direct costs to investors such as Indiana's taxpayers and public employees, there was also an indirect cost to the economy from this egregious violation of the rule of law. As I wrote at the time, "By stepping over the bright line between the rule of law and the arbitrary behavior of men, President Obama may have created a thousand new failing businesses. That is, businesses that might have received financing before but that now will not, since lenders face the potential of future government confiscation. In other words, Mr. Obama may have helped save the jobs of thousands of union workers whose dues, in part, engineered his election. But what about the untold number of job losses in the future caused by trampling the sanctity of contracts today?"[19]

Unfortunately my prediction has been proven correct: subsequent economic analysis of the long-term effects of plundering Chrysler's secured creditors found that in the wake of the government's action, firms in heavily-unionized industries saw decreased bond prices and increased bond yields, "consistent with the government's intervention in the Chrysler bankruptcy increasing lenders' assessment of the risk of lending to firms with a strong labor presence, leading to a signficiant increase in borrowing costs for those firms."[20] By destabilizing contracts to benefit a powerful special interest, the government created a cloud of political risk over financial markets and the economy.

13 Couch, J. F., M. D. Foster, K. Malone, and D. L. Black, 2011, "An analysis of the financial services bailout vote," 31 Cato Journal 119, online http://object.cato.org/sites/cato.org/files/serials/files/cato-journal/2011/1/cj31n1-8.pdf.

14 For a summary of these studies, see Zywicki, T. J., 2016, "Rent-seeking, crony capitalism, and the crony constitution," Supreme Court Economic Review (Forthcoming), available in http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2651587 (Aug. 26, 2015).

15 See Zywicki, T. J., 2014, "The corporatist legacy of the auto bailouts," Law & Liberty Blog, January 13, available at http://www.libertylawsite.org/2014/01/13/the-corporatist-legacy-of-the-auto-bailouts/.

16 Zywicki, T. J., 2011, "The auto bailouts and the rule of law," 7 National Affairs, available at http://www.nationalaffairs.com/publications/detail/the-auto-bailout-and-the-rule-of-law.

17 Zywicki, T. J., 2014, "The corporatist legacy of the auto bailouts," Law & Liberty Blog, January 13, available at http://www.libertylawsite.org/2014/01/13/the-corporatist-legacy-of-the-auto-bailouts/.

18 See Zywicki, supra note 16. This also ignores the still-unexplained decision of bailout operatives to terminate the pension plans of Delphi's white collar employees as part of that company's bankruptcy case. See Special Inspector General for the Troubled Asset Relief Program, Treasury's Role in the Decision for GM To Provide Pension Payments to Delphi Employees (Aug. 15, 2013), available in https://www.sigtarp.gov/Audit%20Reports/SIGTARP_Delphi_Report.pdf.

19 See Zywicki, T. J., 2009, "Chrysler and the rule of law," Wall Street Journal, May 13

20 Blaylock, B., A. Edwards, and J. Stanfield, "The role of government in the labor-creditor relationship: evidence from the Chrysler bankruptcy," 50(3) Journal of Financial and Quantitative Analysis 325, 327

The end of the crisis does not bring about the restoration of the rule of law

Still another cost of deviations from the rule of law during a financial crisis in the name of claimed "emergency" is that the abatement of the crisis does not bring about a subsequent restoration of the rule of law. Instead, as we have seen, the post-crisis period produced a codification and consolidation of government discretion, making it a long-term element of the economy and society. Although having the superficial appearance of a statute, Dodd-Frank's 2,300 pages of legislation largely enshrines much of the arbitrariness and lawlessness that characterized the government's activities during the crisis. For example, it gives the government virtually unreviewable authority to seize what it deems to be failing financial institutions and to deem certain institutions but not others to be "systematically risky" – although it nowhere defines the criteria that qualify a firm as "systemically risky" and provides limited judicial review of the government's actions.

THE IMPACT OF THE DECLINE OF THE RULE OF LAW ON PERSONAL FREEDOM

Three striking examples of the post-crisis regulatory environment illustrate the erosion of the rule of law in action: the adverse effect of Dodd-Frank on small banks, the execution of Operation Choke Point, which limited access to financial services for politically disfavored industries, and the activities of the Consumer Financial Protection Bureau (CFPB).

Disappearing small banks

One well-documented effect of Dodd-Frank has been to promote consolidation of the banking industry by driving out smaller community banks that comparatively lack the resources to comply with Dodd-Frank's crushing and ham-fisted regulatory burden. For example, a recent study by scholars at the Kennedy School of Government found that in the period since Dodd-Frank was enacted, the asset bases of smaller banks have shrunk twice as fast after Dodd-Frank's enactment compared to before, a result that they attribute to the high regulatory costs imposed by Dodd-Frank.[21] In addition, a detailed Mercatus Center study of the impact of Dodd-Frank on smaller banks has found that the law has imposed huge compliance costs on small banks and that they have been less able to bear those costs than large banks.[22]

By replacing fair and free marketplace competition for consumer loyalty with competition to best engage in regulatory arbitrage, Dodd-Frank is restricting consumer freedom of choice and innovation. This impact is most noticeable with respect to home mortgages. Community banks historically have provided more than half of the residential mortgages in the U.S. According to the Mercatus Center study, 64 percent of small banks reported that they were making changes to their mortgage offerings because of Dodd-Frank and 15 percent said that they had either exited or were considering exiting residential mortgage markets entirely.[23] Nearly 60 percent of small banks reported that the CFPB or the qualified mortgage rule had a "significant negative impact" on their mortgage operations. Nearly 60 percent said that the CFPB has had a significant negative effect on bank earnings and more than 60 percent said that changes in mortgage regulations had had a significant negative effect on bank earnings.

Moreover, by imposing a one-size-fits-all mechanical underwriting system for mortgages, the Qualified Mortgage rule has deprived community banks of a significant competitive advantage against megabanks: their intimate familiarity with their customers and their ability to engage in relationship lending with their customers. One illustration of the value of the traditional relationship-lending model for residential mortgages is that the default rate for residential mortgages made by community banks (with less than U.S.$1 billion in assets) was 3.47 percent in 2013 compared to a default rate of 10.42 percent for banks with more than U.S.$1 billion in assets.[24] Thus, this regulatory-induced decline in the market share of small banks is not only hurting consumers, it is making the banking system less stable and less effective. Consumers face a market with fewer choices, less innovation and less competition than before.

The ripple effects of the displacement of smaller banks by large banks are not limited to the direct impact on the banking system but carry over to other markets as well, including agricultural and small business loans. Community banks historically have provided the majority of agricultural and small-business financing in the U.S.[25] As community banks have been driven out of the market by

21 Lux, M., and R. Greene, 2015, "The state and fate of community banking," M-RCBG Associate Working Paper No.37 (February 2015) online at http://www.valuewalk.com/wp-content/uploads/2015/02/Final_State_and_Fate_Lux_Greene.pdf.

22 Pierce, H., I. Robinson, and T. Stratmann, 2014, "How are small banks fairing under Dodd-Frank?" George Mason University Mercatus Center Working Paper No. 14-05 (February 2014) online at http://mercatus.org/publication/how-are-small-banks-faring-under-dodd-frank

23 Id.

24 Peirce, H., 2013, Senior Research Fellow, The Mercatus Center at George Mason University, testimony before the House Committee on Oversight and Government Reform, 113th Congress, 1st session. July 18, http://smbiz.house.gov/uploadedfiles/12-3-2013_peirce_burdensonsmallbanks_testimony_112613.pdf.

25 See Lux and Greene, supra note 21, at 1 (noting that community banks provide 77 percent of agricultural and over half of small business loans).

regulatory costs, small business credit has contracted as well, dampening entrepreneurship and economic growth. As noted by one analysis, large firms have performed well since the financial crisis and subsequent recovery, but small firms have suffered low rates of formation, employment growth, and wage growth.[26] Indeed, the number of small firms in the economy actually declined over the period since the crisis, as more small firms disappeared than were created, the first time that this has happened since data became available in the 1970s.[26] A primary explanation for this drop in small business formation and growth is Dodd-Frank and increased financial regulation since the financial crisis, which has fallen especially hard on smaller banks relative to larger banks.[28] Overall, a recent analysis of FDIC (Federal Deposit Insurance Corporation) data found that while bank loans to small businesses had declined by 16% since 2008, loans to large businesses had increased by 37% over that same period.[29] As one commenter described the situation, large banks "have effectively abandoned the small business market."[30] Another analysis concluded that small business loans are down about 20% since the financial crisis while loans to larger businesses have increased by about 4% over the same period.[31] It appears that some of the unmet demand from the reduction in community bank lending is being served by non-bank lenders that charge higher rates than traditional small business bank loans and which, ironically, are much less-regulated that the traditional banks that they have replaced.[32]

According to Wells Fargo Quarterly survey of small business owners, in the third quarter of 2015, just 33% of small business owners surveyed stated that it would be "very easy" or "easy" to obtain credit if they needed it and 22% said that it would be "somewhat difficult" or "very difficult."[33] Only 19% said it would be "very easy" to obtain credit when they needed it; even more remarkable, that is the highest level for those saying credit is "very easy" since the recession hit and Dodd-Frank was enacted, as for most of that period the rate has been in the low-teens. By contrast, during the period from the 1Q2004-4Q2007, an average 51% of small business owners said that it was "very easy" or "somewhat easy" to obtain credit if they needed it, and about 12% said it would be difficult. In addition, among those who said that it was easy to obtain credit in the 2004-07 period, 2/3 of those reported it was "very easy" compared to "somewhat easy," whereas only about half of those who said that it would be easy in the post-Dodd-Frank pool reported that it would be "very easy."

As smaller banks have been disappearing and exiting certain markets, large banks have grown still larger and Dodd-Frank has increased their insulation from competitive pressures. In fact, large banks have admitted as much. For example, JP Morgan Chase CEO Jamie Dimon observed that the aggregate costs of complying with all of the rules, regulations, and capital costs associated with Dodd-Frank has built a "bigger moat" to protect his bank from competition from smaller rivals.[34] Similarly, Goldman Sachs CEO Lloyd Blankfein announced in 2010 that the bank would be "among the biggest beneficiaries" of Dodd-Frank as its regulatory costs and regulatory-created profit opportunities would be particularly advantageous to large banks that could bear those costs more easily than smaller competitors.[35]

Moreover, because many of Dodd-Frank's most expensive rules kick-in once a bank reaches U.S$10 billion in assets, that figure acts as a sort of tripwire – either banks try to remain below that threshold, or if they do cross it, then they accelerate their merger activities to try to gain the size and economies of scale necessary to cope with heightened regulatory costs. Thus, the market is becoming increasingly bifurcated between large banks and very small banks, as medium-sized banks grow larger.[36] On the other hand, only one new bank has been formed since the financial crisis and small banks continue to merge or otherwise disappear as a result of their own regulatory costs. This phenomenon of the disappearance of small banks and the lack of creation of new ones led economists from the Dallas Federal Reserve bank to ask whether small banks are "too small to succeed" in light of the huge growth in regulatory cost and complexity imposed in the period since the financial crisis.[37] They too note the important role played by community banks in small business lending and agricultural markets and the adverse effects on small-business formation and growth as a result of this trend toward the disappearance of small banks.

---

26 Goldman Sachs, 2015, "The two-speed economy," 2, April
27 Id.
28 Goldman Sachs, 2014, "Who pays for bank regulation," June
29 Simon, R., 2015, "Big banks cut back on loans to small business," Wall Street Journal, November 26
30 Id.
31 See Mills, K. G., and B. McCarthy, 2014, "The state of small business lending: credit access during the recovery and how technology may change the game," Harvard Business School Working Paper 15-004
32 Id.
33 Gallup/Wells Fargo Small Business Survey 24-25, https://wellsfargoworks.com/File/Index/btDsu4gv9UqK07hpFKVbgw (Oct. 3, 2015) (responses to Question 11).
34 Rouan, R., 2013, "Dimon says Dodd-Frank puts 'bigger moat' around JPMorgan Chase," Columbus Business First, February 5, available in http://www.bizjournals.com/ columbus/blog/2013/02/dimon-says-dodd-frank-puts-bigger.html.
35 Carney, T. P., 2015, "Goldman and JPMorgan sit safely behind the walls of Dodd-Frank," Washington Examiner, February 12, available in http://www.washingtonexaminer .com/goldman-and-jpmorgan-sit-safely-behind-the-walls-of-dodd-frank/article/ 2560179.
36 Picker, L., and M. Monks, 2013, "Small banks feel the urge to merge," Bloomberg Business, October 3, available at http://www.bloomberg.com/bw/articles/2013-10-03/dodd-frank-fills-small-banks-with-the-urge-to-merge.
37 See Ash, P., C. Koch, and T. F. Siems, "Too small to succeed? – Community banks in a new regulatory environment," Dallas Fed Financial Insights, Vol. 4, Issue 4

Targeting businesses by operation choke point and the CFPB

In the post-Dodd-Frank era, the vast, ill-defined sway that regulators exercise over banks has enabled them to not only pick winners and losers in the financial system but to also use their clout to force banks to do their bidding outside of the formal regulatory process. Indeed, in some instances government regulators have essentially deputized banks as arms of the federal government, directing banks to attack private parties engaged in legal activities – without evidence of wrongdoing or the public scrutiny that a direct government action would bring. Consider two examples that demonstrate the point: Operation Choke Point and the Consumer Financial Protection Bureau's initiative against auto dealers for purported disparate impact in lending rates.[38]

Operation Choke Point

Consider first the shadowy initiative known as Operation Choke Point, which seems to have been spearheaded by the Department of Justice and FDIC. Under Operation Choke Point, government regulators targeted a myriad of legal, but politically unpopular industries, such as firearms dealers, coin dealers, pornography, sellers of "racist materials," home-based charities, and most intensely, payday lending.[39] The FDIC, of course, had no jurisdiction over these industries and absent any demonstrable wrongdoing, the DOJ could not outlaw them either. Yet these limitations did not stop them.

Instead, the FDIC instructed regulated banks to cease providing banking services to these particular industries, with special attention paid to payday lenders, to "choke off the air" needed for these firms and industries to function.[40] Without the ability to clear checks and process electronic payments, payday lenders and other targeted firms simply could not exist and conduct business. Notably, the government's instructions were issued without any evidence that any of the industries on the affected list had done anything illegal, with no due process to the adversely affected firms, and, indeed, with a complete lack of transparency, including a reluctance to even admit except under pressure that the initiative even existed. Equally notable was the selective nature of the government's list of controversial industries that created "reputation risk" for banks, which included industries such as firearms sales but ignored other controversial industries such as abortion clinics. In one particularly colorful example of the lawless nature of the program, a senior official in the Division of Depositor and Consumer Protection instructed that any communications by FDIC Chairman Martin Gruenberg "always mention pornography when discussing payday lenders and other industries, in an effort to convey a 'good picture regarding the unsavory nature of the businesses at issue.'"[41] Aggressive oversight by Congress eventually persuaded FDIC to withdraw its list of target industries and to formally claim that it was terminating Operation Choke Point,[42] but news reports indicate that it might still be continuing and that its implementation has simply shifted to the CFPB.[43]

Despite these formal actions, there are reports that suggest that Operation Choke Point or some variant thereof, continues to operate within the financial services sector.[44] For example, it has been reported by one bank that the Treasury Department forced it to categorically discontinue providing money transfer services to Somalia. According to Oxfam International, the result of this prohibition on remittances may be the starvation of three million Somalis who depend on remittances from the West.

CFPB and alleged discrimination by auto dealers

A second example is the effort of the CFPB to enforce fair lending laws on auto dealers for the loans that they issue. Fair lending laws that prohibit discrimination in making loans apply to auto dealers. It is equally clear, however, that Dodd-Frank prohibits the CFPB from exercising jurisdiction over loans made by auto dealers, leaving that responsibility by implication to other federal agencies such as the Federal Trade Commission and DOJ.[45]

Lacking the authority to reach the auto dealers, the CFPB came up with a creative solution – it decided to hold the financial institutions

38 The following discussion draws from Zywicki, supra note 14.

39 The entire list of targeted industries was promulgated informally by the FDIC in U.S. House, Committee on Oversight and Government Reform, 2014, "The Department of Justice's "Operation Chokepoint": illegally choking off legitimate businesses?" Staff Report 113th Congress at 11, May 29, available online at http://oversight.house.gov/wp-content/uploads/2014/05/Staff-Report-Operation-Choke-Point1.pdf.

40 See, e.g., Letter from M. Anthony Lowe, Director, FDIC Chicago Regional Office to Board of Directors of [Redacted] Bank (Feb. 15, 2013), available at http://oversight.house.gov/wp-content/uploads/2014/10/Regional-Director-Letter.pdf (stating that providing banking services to payday lending companies "carries a high degree of risk to the institution, including third-party, reputational, compliance, and legal risk" and that as a result "activities related to payday lending are unacceptable for an insured depository institution").

41 U.S. House, Committee on Oversight and Government Reform, 2014, "Federal Deposit Insurance Corporation's involvement in "Operation Choke Point," Staff Report 113th Congress at 1, December 8, available at http://oversight.house.gov/wp-content/uploads/2014/12/Staff-Report-FDIC-and-Operation-Choke-Point-12-8-2014.pdf.

42 See Hoover, K., 2014, "FDIC removes Operation Choke Point's 'hit list,' clarifies guidance to banks," The Business Journals, July 29, available at http://www.bizjournals.com/bizjournals/washingtonbureau/2014/07/fdic-removes-operation-choke-points-hit-list.html.

43 See Witkowski, R., 2015, "CFPB launches its own Choke Point-Style operation," American Banker, April 8, available at http://www.americanbanker.com/news/law-regulation/cfpb-launches-its-own-choke-point-style-operation-1073659-1.html.

44 See Sant, G., and B. Williams, 2015, "The choking continues after "Choke Point"," American Banker, October 19

45 Dodd-Frank Wall Street Reform and Consumer Protection Act § 1029(a), Pub.L 111-203, H.R. 4173 (2010).

(the indirect lenders) responsible for any alleged discriminatory lending patterns by the auto dealers themselves. Indirect lenders bear this responsibility even though they have no interaction with the borrower, information about the borrower's race, or any reason to believe that the dealers are engaged in discriminatory lending patterns. Moreover, the indirect lenders would be held responsible according to the theory of "disparate impact," making the indirect lenders responsible for any statistical anomalies that seemed to exist, regardless of the lack of any evidence of intentional discrimination.

A prime illustration of the modern approach to the modern regulatory approach was the CFPB's decision to target Ally Financial for its first high-profile settlement for alleged discrimination in auto dealer markups.[46] According to internal documents examined by the House Financial Services Committee, the CFPB identified Ally as its first target not because Ally had acted in a particularly improper fashion, but because Ally was particularly vulnerable to being strong-armed into a settlement. This was for three reasons. First, as a result of the continued legacy of the auto bailouts, the federal government still held a 73.8% stake in Ally at that time (and still held 63.4% at the time the case was actually settled). Second, Ally had an application pending in front of the Federal Reserve to become a financial holding company, approval of which was necessary to continue its insurance and used-car remarketing operations. Third, the FDIC was conducting a Community Reinvestment Act review of Ally and settlement of the CFPB investigation was a precondition to receive a satisfactory CRA rating, which in turn was necessary for approval of Ally's status change to become a financial holding company. Faced with these obstacles, Ally eventually capitulated and finally paid U.S.$98 million for restitution and civil penalties.

On the other hand, because the CFPB never identified particular victims of discrimination but relied on statistical aggregates, it had no way of identifying the race of the supposed victims or to identify those to whom restitution should be paid. Instead, the CFPB relied on a statistical technique known as Bayesian Improved Surname Geocoding, which has been demonstrated to be statistically invalid for these purposes.[47] Indeed, according to documents secured by the House of Representatives Financial Services Committee, the CFPB itself was aware of the flaws in the methodology and the CFPB's proposed use, yet nevertheless persevered, using it as a basis to establish liability. The result has been to issue "restitution" checks to many people who have provided no evidence that they were the subject of racial discrimination – including at least one identified beneficiary who is not even a minority.[48]

The examples demonstrate the hazards of the absence of the rule of law in the modern financial regulatory system as the federal government has essentially weaponized America's financial institutions to carry out policies that it couldn't otherwise accomplish. Moreover, much of the policymaking is done in back rooms with no other formal protections or transparency. For example, Operation Choke Point was a secretive government program the very existence of which proved difficult to confirm, much less its details and implementation (it is not even clear today whether the program continues and if so, which agency is executing it). The CFPB's attack on indirect auto lenders was issued through a five page "Guidance" document that provided no information about the basis for the CFPB's charge of discrimination or, originally, any methodology for determining liability, no opportunity for public comment or other due process protections and no assessment of the impact on consumers.[49] In fact, according to a recent report in the Wall Street Journal, by narrowing the range over which dealers and consumers can bargain, the overall effect of the CFPB's micro-managing of the auto finance market has resulted in higher interest rates on car loans for consumers.[50] Meanwhile, those entities that are politically disfavored, such as payday lenders and firearms dealers, are crushed with no due process and no opportunity to defend themselves in any transparent regulatory proceeding.

Other examples of regulatory overreach under Dodd-Frank

The arbitrary exercise of regulatory authority has real-world consequences for consumers and the economy. For example, the complexity and risk under the Qualified Mortagages rule when combined with the threat of "put back" liability for loans based on trivial technical violations has led several leading mortgage lenders to exit the market for borrowers with lower credit scores.[51] As John Sumpf, the chief executive of Wells Fargo stated, "If you guys want

46 Report of Republican Staff of the Committee on Financial Services, U.S. House of Representatives, 2015, "Unsafe at any bureaucracy: CFPB junk science and indirect auto lending," November 24

47 See Baines, A. P., and M. J. Courchane, 2014, "Fair lending: implications for the indirect auto finance market," American Financial Services Association, November 19, available at http://www.crai.com/sites/default/files/publications/Fair-Lending-Implications-for-the-Indirect-Auto-Finance-Market.pdf.

48 Wall Street Journal, 2015, "Do two half victims make a whole case?" Opinion, April 13, available at http://www.wsj.com/articles/do-two-half-victims-make-a-whole-case-1428966741

49 Consumer Financial Protection Bureau, 2013, "Indirect auto lending and compliance with the Equal Credit Opportunity Act," CFPB Bulletin 2013-02, March 21, available at http://files.consumerfinance.gov/f/201303_cfpb_march_-Auto-Finance-Bulletin.pdf.

50 See Andriotis, A., and G. Nagesh, 2015, "Crackdown on racial bias could boost drivers' costs for auto loans," Wall Street Journal, August 31, available at http://www.wsj.com/articles/crackdown-on-racial-bias-could-boost-drivers-costs-for-auto-loans-1441038864. The CFPB ignores other important elements of the inquiry, especially that unlike many other credit transactions a car loan from an auto dealer is not a stand-alone transaction but is linked to the purchase of a car. For example, auto dealers offer promotional financing deals on particular car models in order to move inventory (rather than cutting the sticker price), which can result in spurious implications of differential pricing overall.

51 See Hall, C., "Wells chief warns on mortgage lending," August 26, available at http://www.ft.com/intl/cms/s/0/cdfe20f8-2a2d-11e4-a068-00144feabdc0.html#axzz3lZKUva3B.

to stick with the programme of 'putting back' any time, any way, whatever, that's fine, we're just not going to make those loans and there's going to be a whole bunch of Americans that are underserved in the mortgage market."[52] Similarly, Federal Reserve Chairwoman Janet Yellen has observed, "Banks, at this point, are reluctant to lend to borrowers with lower FICO scores. They mention in meetings with us consistently their concerns about put-back risk, and I think they are – it is difficult for any homeowner who doesn't have pristine credit these days to get a mortgage."[53]

Government power unconstrained by the rule of law also has direct implications for consumers by cultivating an environment of bureaucratic hubris at the expense of the rest of us. Consider the CFPB's extraordinary data mining program of American families' financial accounts. According to a report by the Government Accountability Office, the CFPB collects information on 10.7 million individual consumer credit reports on a monthly and quarterly basis, more than 500 million credit card accounts on a monthly basis, and 29 million active mortgages and 173 million total mortgages on a monthly basis.[54] Moreover, because this data-mining program was not initiated according to any sort of formal notice and comment rulemaking procedure, it is not subject to cost-benefit analysis or any other evaluation as to whether such extensive snooping is necessary to further any legitimate regulatory purpose. In fact, George Mason University economist Thomas Stratmann has estimated that the number of credit card accounts for which the CFPB wants to collect consumer information on is some 70,000 times greater than is necessary for the agency to execute its regulatory mission.[55] Indeed, the Bureau itself has refused to permit consumers from opting-out of the program, admitting that if consumers were permitted to withdraw consent to the program the government would be unable to obtain the data.[56]

But the costs of CFPB's demand for information do not fall solely on the banks that must provide it. While the CFPB claims that this data is anonymous, every bit of information increases the risk to consumers of identity theft and other misuse of their information. In fact, testifying before this committee last year, CFPB director Richard Cordray admitted that the information the CFPB collects is not 100 percent secure and could be hacked.[57] Moreover, according to a recent article in Science, using only three months of anonymous credit card data, the researchers were able to reidentify 90 percent of individuals, with women being more readily reidentifiable than men.[58]

While the unnecessary acquisition and retention of troves of Americans' information is troubling enough in itself, it is especially worrisome in light of repeated rebukes of the CFPB's faulty data security systems.[59] Following massive data security breaches and compromising of personal information by the Internal Revenue Service and Office of Personnel Management, it is inexplicable that the CFPB continues to insist on vacuuming up excessive amounts of consumer data without considering the privacy threat to consumers. Leaving aside the risk of creating a massive trove of financial data for private hackers to target, Americans also have a fundamental interest in not having their purchases tracked by the federal government and an expectation that the government should not demand any more personal financial data than is necessary to advance its legitimate regulatory purposes.

MORAL HAZARD AND THE RULE OF LAW

The erosion of the rule of law creates a problem for the future: because of the government's demonstrated unwillingness to abide by the rule of law – and the courts' unwillingness to force it to do so in the midst of a financial crisis[60] – the government is unable to credibly commit itself to not use its authority to intervene in the economy, to bail out large banks and to exercise its authority in a political fashion.

Thus, at the same time that smaller banks are being ground under Dodd-Frank's regulatory wheel, there is a general consensus that the Act has failed to address the most fundamental regulatory problem highlighted by the financial crisis: financial institutions that are considered TBTF are backed by an implicit government guarantee.

52 Id.

53 Id.

54 Government Accountability Office, Consumer Financial Protection Bureau: Some Privacy and Security Procedures for Data Collections should Continue Being Enhanced, 2014

55 See Letter of Professor Thomas Stratmann to Congressman Scott Garrett, January 23, 2014), available in http://mercatus.org/sites/default/files/StratmannCFPBStatisticMethods.pdf.

56 See Pollock, R., 2014, "Federal consumer bureau data-mining hundreds of millions of consumer credit card accounts, mortgages," Washington Examiner, January 29, available at http://www.washingtonexaminer.com/consumer-bureau-data-mining-hundreds-of-millions-of-consumer-credit-card-accounts-mortgages/article/2543039.

57 Id.

58 de Montjoye, Y-A., L. Radaelli, V. K. Singh, and A. Pentland, 2015, "Unique in the shopping mall: on the reidentifiability of credit card metadata," 347 Science No. 6221 536-39

59 See Government Accountability Office, Consumer Financial Protection Bureau, 2014, "Some privacy and security procedures for data Collections should continue being enhanced," September; Board of Governors of the Federal Reserve System, Consumer Financial Protection Bureau, Office of Inspector General, 2014, "Security control review of the CFPB's cloud computing–based general support system," 2014-IT-C-010, Washington, D.C.: July 17

60 See Zywicki, supra note 6.

Instead of resolving or mitigating that problem, Dodd-Frank has entrenched the TBTF problem. A report by the Government Accountability Office concluded that while Dodd-Frank may have reduced the size of the so-called "TBTF subsidy" for large banks it did not eliminate it, indicating that large banks still retain an implicit government guarantee.[61] A study by the International Monetary Fund concluded that the subsidy to TBTF banks in the U.S. amounts to some U.S.$70 billion per year in lower capital costs and that in turn the existence of an implicit government guarantee promotes the moral hazard problem of greater risk-taking by large banks.[62]

Despite the elaborate procedures concocted in Dodd-Frank for the resolution of financial distress by banks, the fundamental problem is that these procedures simply are not considered credible by market actors. No one seriously believes that a future President and future Congress will feel themselves bound to abide by Dodd-Frank's requirements when it comes to the resolution of distress by financial firms. This disbelief reflects the erosion of the rule of law and, in this sense, the expectation that large banks will be bailed out effectively becomes a self-fulfilling prophecy – just as Treasury Secretary Paulson's primary justification for bailing out banks being that the markets "expected it."[63]

More generally, in the post-Dodd-Frank world, the combination of vast, unaccountable political power combined with the increased clout of powerful special interests to use the regulatory process has – unsurprisingly – led to an explosion of lobbying activity by financial services firms to avoid the imposition of the crushing burden of heavy and arbitrary government action. In other cases, lobbying reflects rent-seeking activity and efforts by some firms to influence the political and regulatory process to gain a competitive advantage over rivals. In addition, the power of politicians to pick winners and losers arbitrarily has created greater opportunities for rent-extraction by politicians who can threaten to impose new regulations unless bought off by lobbying efforts and campaign contributions.[64]

Little wonder that the financial services industry spends tens of millions of dollars every year on lobbying expenditures to seek special treatment under the law or to protect themselves from arbitrary regulation. In a world where government officials hold the power to hand out billions of dollars of regulatory prizes and punishments with no accountability and no need to justify their actions according to any coherent principle – other than political expediency – powerful special interests are going to try to influence that process to their advantage.[65] The virtue of the rule of law is to restrain the discretionary power of the government to draw these sorts of arbitrary distinctions that permit some interests to benefit politically at the expense of others.[66]

CONCLUSION: DODD-FRANK AND THE DECLINE OF THE RULE OF LAW

In this world of lawlessness and arbitrary regulatory authority clout is king. What does that mean for the rest of us? It is not often appreciated, but it is the average American or small business that benefits the most from upholding the rule of law. Big financial firms can survivie – indeed, even thrive – in a world devoid of settled rules and transparent governance. They can afford to hire the lawyers and lobbyists to wend their way through the arcane political and regulatory processes.

But everyone else – small businesses and ordinary families trying to get ahead in life – do not have access to expensive, well-connected lawyers and lobbyists. When we have to pay more for a car loan or cannot obtain a credit card, mortgage, or small business loan to make our families' lives better, we cannot find a high-priced lobbyist to grease the skids for us. When our government spies on our credit card accounts without our conent and seeks to "choke off" banking services for legal businesses, we are less free. Dodd-Frank has interjected the tentacles of the federal regulatory state into every aspect of our financial system, and as a result we are less free to obtain the means to make our lives better.

61 Government Accountability Office, 2014, "Large bank holding companies: expectations of government support," July

62 International Monetary Fund, 2014, "Big banks benefit from government subsidies," Global Financial Stability Report, available at http://www.imf.org/external/pubs/ft/survey/so/2014/POL033114A.htm. Other studies reach different conclusions on the continued existence of the TBTF subsidy. For a summary of the literature as well as a caveat on the conclusions that can be drawn, see Zywicki, supra note 12.

63 See Zywicki, T. J., 2013, "The next financial crisis: what will the market 'expect'?" Law & Liberty Blog, May 19, available at http://www.libertylawsite.org/book-review/the-next-financial-crisis-what-will-the-markets-expect/.

64 See Zywicki, supra note 14 (citing example of threats to impose new comprehensive regulations on hedge funds); see also Carney, T. P., 2010, "Schumer's racket: lobbyists and hedge funds," Washington Examiner, May 26, available at http://www.washingtonexaminer.com/article/13668.

65 See Tullock, G., 1967, The Welfare Costs, of Tariffs, Monopoly, and Theft, Western Economic Journal 5 :3, 224-232.

66 See Zywicki, supra note 12.

The Volcker Rule as Structural Law: Implications for Cost-Benefit Analysis and Administrative Law*

John C. Coates – John F. Cogan Jr. Professor of Law and Economics, Harvard Law School

Abstract

The Volcker rule, a key part of Congress's response to the financial crisis, is best understood as a "structural law," a traditional Anglo-American technique for governance of hybrid public-private institutions such as banks and central banks. The tradition extends much farther back in time than the Glass-Steagall Act, to which the Volcker Rule has been unfavorably (but unfairly) compared. The goals of the Volcker Rule are complex and ambitious, and not limited to reducing risk directly, but include reshaping banks' organizational cultures. Another body of structural laws, part of the core of administrative law, attempts to restrain and discipline regulatory agencies, through process requirements such as cost-benefit analysis (CBA). Could the Volcker rule be the subject of reliable, precise, quantified CBA? Given the nature of the Volcker rule as structural law, its ambitions, and the current capacities of CBA, the answer is clearly "no," as it would require regulators to anticipate, in advance of data, private market behavior in response to novel activity constraints. If administrative law is to improve regulatory implementation of structural laws such as the Volcker Rule, better fitting and more nuanced tools than CBA are needed.

* Reprinted with permission from the Capital Markets Law Journal, Oxford University Press, Volume 10 Issue 4. I thank Jody Freeman, Jeff Gordon, Stephen Kane, Todd Rakoff, Meg Tahyar, and an anonymous referee for useful comments and suggestions. Remaining errors are mine.

The best known section of Congress's response to the financial crisis – the "Volcker Rule," section 619 of the Dodd-Frank Act[1] – is a "structural law," with implications for efforts to use cost-benefit analysis (CBA) to enhance regulatory accountability as the rule and others like it are implemented. After briefly characterizing structural laws, this article places the Volcker Rule in historical context, as part of a long tradition of Anglo-American attempts to use structural laws as a technique for governance generally, and of hybrid public-private institutions such as banks and central banks in particular. The article then outlines how another set of laws and institutions, developed later and reflected in administrative law, have been used to constrain regulatory agencies, including those overseeing capital markets, by imposing special procedures and analytical requirements before rules can be changed, such as "CBA," to enhance the policy-neutral accountability of the agencies, but also as a non-neutral political tool of the banks themselves. Finally, the article asks whether, as others have argued, structural laws such as the Volcker Rule should be subject to legally mandated, quantified CBA. Unlike some commentators[2] the article gives an answer, no, that is both consistent with U.S. legal traditions, and based on common sense, given the nature of the Volcker Rule as structural law and the current capacities of CBA. The analysis here, it is hoped, casts light both on the Volcker Rule and on the potential value (and risks) of legal mandates for CBA in administrative law generally.

A BRIEF HISTORY OF STRUCTURAL LAWS IN ANGLO-AMERICAN FINANCIAL HISTORY

Structural laws generally

Different laws function differently. One type of law, a structural law, attempts to create a "structure" that will organize, constrain, and channel activity. Structural laws create and provide for the governance of organizations (e.g., a regulatory agency, such as the Federal Reserve Board; a quasi-public corporation, such as the Federal Deposit Insurance Corporation; or quasi-private corporations, such as systemically important financial institutions), institutions (e.g., a system of connections, such as a road, computer[3] or payment system; or a market, such as a stock exchange) and physical objects (e.g., the blue mailboxes used by the U.S. Postal Service, buildings, safe deposit boxes, nodes of the internet).[4] Structures can be built by affirmative government action (as with a highway system or the Fedwire payment system) or through laws aimed at private or partly private persons (as with regulations of financial markets such as the New York Stock Exchange).

Not all laws are structural. Many are direct commands aimed at private individuals or entities, such as requirements to pay income taxes. Other laws consist of mandates, to make specified disclosures or to maintain specified capital levels, for example. Others are bans aimed at behavior that is socially undesirable, such as theft and fraud.

The distinction between ordinary laws and structural laws has less to do with their form, than with their goals. Rather than banning an undesirable behavior, as with an ordinary law, structural laws may require transparency, which will lead those covered to alter their behavior, or they may ban otherwise unobjectionable behavior, in order to increase some desirable behavior, or in order to simplify supervision of behavior that can create social risks. Structural laws, in other words, are indirect, and have their effects "ex-ante," in advance of some decision by those affected. As a result, they can be more self-executing than other laws, in the sense that once created, they require lower levels of public enforcement effort. As a result, even if they create large initial compliance and adjustment costs, structural laws are often more efficient at achieving public goals than laws that function as simple commands enforced solely through the fact or threat of criminal prosecution or civil fine. Structural laws, particularly those affecting organizational governance or behavior, often affect remote actors without any self-conscious change in behavior or even affirmative awareness by those affected, can be less likely to generate evasion. Structural laws affecting

1 Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub L No 111-203, s 619, 124 Stat 1376, 1620 (2010) (codified at 12 USC s 1851 (2012)). Section 619 is colloquially called the "Volcker Rule" because former Federal Reserve Board Chairman Paul Volcker was a prominent backer of the law. It is not technically a "rule" in the standard legal usage in the U.S., but part of a statute, now implemented through rules and regulations adopted by designated federal agencies, as discussed below. Also as discussed below, the Volcker Rule applies to "banking entities," rather than to "banks" or "bank holding companies," the customary nouns for regulation in the U.S. In this article, I generally use "bank" in a non-technical sense, to include both technical "banks" under U.S. law, as well as bank holding companies and other entities affiliated with banks, except where indicated otherwise.

2 For example, Committee on Capital Markets Regulation, Committee Issues Statement on Applying the Volcker Rule (18 February 2014) ("we reiterate our concern over the lack of cost-benefit analysis in the Volcker Rule... For a regulation as significant as the Volcker Rule, conducting cost-benefit analysis in accordance with best practices should be an agency priority, even where not required by law.").

3 For example, Lessig, L., 1999, Code and other laws of cyberspace, Basic Books (architecture as a regulator of cyberspace); Reidenberg, J. R., 1998, "Lex informatica: the formulation of information policy rules through technology," 76 Texas Law Review 553, 554 ("Technological capabilities and system design choices impose rules on participants.").

4 I take the mailbox example from Edward K Cheng"s illuminating article, Cheng, E. K., 2005, "Structural laws and the puzzle of regulating behavior," 100 Northwestern University Law Review 655, 662 (noting how the use of uniform steel mailboxes helped greatly reduce the costs of enforcing laws against tampering or stealing mail); see also Katyal, N. K., 2002, "Architecture as crime control," 111 Yale Law Journal 1039, 1042 (analogizing "code" as architecture to actual building architecture as means for reducing crime).

organizational governance or behavior can be more effective than either direct commands or disclosure requirements intended to inform third parties, which depend on those third parties obtaining the information, processing it and acting on it in rational, or at least systematically predictable ways.

Structural laws in finance

Structural laws have long been a core component of Anglo-American legal history. Fundamental structures for government itself – the separation of powers, federalism, for example – are embedded in the U.S. Constitution. In finance, structural laws have also been common and traditional from the outset of modern banking.

Reflecting the mixed public-private character of large banks and systemically important financial institutions (SIFIs), structural laws have been used to serve several purposes, at times complementary and at times competing: to restrain the power of banks, to limit their profitability when privately owned, to protect banks from competition, to reduce systemic risk, and to shelter central banks from political pressures in their management of the money supply.[5]

At its creation, the Bank of England was structurally limited to financial activities by the terms of its charter, reinforced by custom, and so was barred from engaging in nonfinancial activities, such as trade, or, as the Industrial Revolution progressed, manufacturing.[6] It remained a privately owned organization long after it had taken on the public obligations modernly associated with a central bank.[7] But throughout its private existence, it was constrained by structural laws, partly enforced by the terms and conditions of periodic bailouts caused by poorly managed financial panics.[8] In 1844,[9] 1946[10] and 1998,[11] Parliament passed structural laws that radically reshaped the Bank's basic functions, ownership and governance.

In the U.S., too, both the First and Second Banks of the USA had their activities carefully limited by the terms of their charters, and indeed, similar constraints were imposed on all banks and corporations in the early American period.[12] This tradition carried past the Civil War, and constrained (and still constrains) the national banks created by the National Banking Act of 1863,[13] to engage in banking business alone, and also to constrain their geographic reach, typically to a single city, county or state.[14] These privately held banks were seen as necessary for public functions – the creation of a currency and payment system – but were also viewed with suspicion. Likewise, when the Federal Reserve Banks were created in 1913,[15] they were similarly constrained for similar reasons, and their governance a highly negotiated political compromise between regional, sectoral, and partisan interests. The Federal Reserve System more generally has had its governance and powers carefully negotiated and renegotiated through structural laws during its entire existence.[16]

Structural banking laws in the twentieth century

Over the course of the twentieth century, a new feature in the legal landscape affecting finance was the emergence of regulatory agencies.[17] Early structural laws were relatively simple, and were contained in bank charters or statutes. Two things changed this, and led to more detailed and complex structural laws in the finance sector. First, the laws needed to cover the behavior of a greater number of banks and institutions, as a result of growth in the economy and the financial sector, accompanied by a commitment to private enterprise – in both the Jacksonian era and in the Gilded Age – and the continued resistance to a single, dominant central bank. These developments increased the number and significance of banks, generating a need for structural constraints that could be imposed other than through a focused chartering decision. The activities of the banks, trust companies and other financial firms became more complex, too, as the country's markets matured. Bond underwriting, stock issues, mergers and acquisitions, options trading, speculation in the commodities markets, were all features of the late nineteenth and early twentieth century financial system.

---

5 Reduction of systemic risk is now the widely acknowledged primary goal of such laws. On limiting power and limiting profit, see Hammond, B., 1957, Banks and politics in America from the revolution to the civil war, Princeton University Press (recounting battles between national and state bank promoters and among bank and non-bank political interests generally); Jackson, A., 1832, "Veto message," 10 July, http://tinyurl.com/9hmony accessed 29 August 2015 ("The present value of the monopoly [to be granted to the Second Bank of United States] is $17,000,000, and this the act proposes to sell for [$3,000,000]..."); Roe, M., 1994, Strong managers, weak owners, Princeton University Press (noting ways that U.S. financial laws preserved autonomy for corporate managers). On sheltering monetary policy from politics, see Blinder, A., 2004, The quiet revolution: central banking goes modern, Yale University Press; Lastra, R. M., 2006, Legal foundations of international monetary stability, Oxford University Press; Miller, G. P., 1998, "An interest-group theory of central bank independence," 27 Journal of Legal Studies 433; see also Bagehot, W., 1871, Lombard Street, Henry S. King and Co., II.65 ("A trade [such as central banking] peculiarly requiring consistency and special attainment would be managed by a shifting and untrained ruler ... [At least in England,] the practical result ... would ... be bad ... for Government ... to choose" the governors of the central bank).

6 Bank of England Act 1694 (5 & 6 W&M c 20 s 28); Clapham, J. H., 1970, The Bank of England: a history, Cambridge University Press; Carruthers, B. G., 1996, City of capital: politics and markets in the English financial revolution, Princeton University Press; Desan, C., 2014, Making money: coin, currency and the coming of capitalism, Oxford University Press.

7 Bagehot (n 5) II.27 describes the Bank"s public function in plain terms ("great public duty").

8 Ibid II.10.

9 Bank Charter Act 1844 (7 & 8 Vict c 32) (monopolizing note-issuing powers in the Bank, limiting note issuances to reserves in gold plus up to 14 million in government debt).

10 Bank of England Act 1946 (9 & 10 Geo 6 c 27) (nationalizing the ownership of the Bank of England).

11 The Bank of England Act 1998 (Commencement) Order 1998 (1998 No 1120 (c 25)).

12 Hammond (n 5).

13 12 U.S. Code ch 2ff.

14 12 U.S. Code s 24 (Seventh) (limiting national banks to activities "incidental to the business of banking").

15 Act of 23 December 1913, 38 Stat, L 251.

16 Conti-Brown, P., 2015, "The institutions of Federal Reserve independence," 32 Yale Journal of Regulation, 2, 257

17 The first bank regulator, the Office of the Comptroller of the Currency (OCC), supervised national banks, and was created in 1863, but the remaining federal banking agencies were not created until the twentieth century.

Secondly, the goals of structural laws became more ambitious, beyond the kind of constraint on activities reflected in the earlier period. This ambition reflected greater recognition of the importance of a functioning currency, which became acute during the Civil War, that required significantly greater public finance than previous wars. The resulting national system of banks, interconnected to each other and to English banks and trading companies, was capable of transferring capital from and to different parts of the economy. This capacity accompanied and supported the emergence and growth of canal and shipping companies, resource extraction and distribution companies such as Standard Oil, and the great regional and then transcontinental railroads. This system also generated a series of increasingly serious financial crises, however, that led to the formation of the Federal Reserve, and an increasing effort to regulate the financial system as a whole.

With the emergence of regulatory agencies, structural laws took on a different, more complex agenda. Agencies were delegated increasing amounts of authority and discretion to achieve such tasks as governance, risk management, and conflict management. Reserve requirements, loan underwriting standards, and capital requirements were modernized, and other modern features of financial law emerged, first through supervisory guidance and enforcement, then through formal rules and regulations.

These developments were concentrated in the wake of crises and failures – most significantly after the Crash of 1929 and the Great Depression, but they also reflected a combination of interest group politics – as banks began to move into other financial sectors – and the back-and-forth of private efforts to evade existing laws followed by regulatory efforts to combat the evasion. The most famous structural laws enacted in this period were contained in the Banking Act of 1933, which established federal deposit insurance and incorporated the Glass-Steagall Act. That law established a structural regime separating investment and merchant banking (i.e. equity underwriting and equity investment) from commercial banking (i.e. deposit-taking and lending).[18] Shortly later, the Investment Company Act imposed a stringent set of structures on any company that wanted primarily to engage in the business of investing or holding securities and to raise capital from the public.[19]

The same approach was taken, and indeed tightened, in the Bank Holding Company Act of 1956, after banks began to use holding companies to evade the structural restraints of the National Banking Act to move into non-banking financial activities such as insurance and to operate across state lines.[20] Similar structural laws were later imposed on state banks as one of the costs of federal deposit insurance.[21]

Survival of structural laws through the era of "deregulation"

Contrary to popular impression, many important structural laws constraining finance survived the 1970s, 1980s and 1990s, commonly said to be a period of "deregulation" – in truth, "re-regulation," since few banking activities were fully deregulated. Instead, technological and market changes forced traditional limits on product competition to be loosened and price controls on interests rates and geographic market segmentation, all designed to limit competition in the name of safety and soundness, to be eliminated. As these structural limits were being lifted, little thought was given to how reforms in the structure of the financial sector should lead to reforms in the philosophy and structure of supervision and regulation. Most limits on the activities of both banks (national and state) and holding companies remained largely intact, as did the limits on investment companies, even as the limits on "investment banking" in Glass-Steagall Act began to erode. The limits on investment banking were largely eliminated, but not in dramatic fashion, as often suggested. Instead, it occurred over a lengthy period of time, exemption by exemption, exception by exception.[22] Banks (and their lawyers) and regulators negotiated and renegotiated the precise contours of the structural limits imposed by Glass-Steagall, a process that was functionally complete by the early 1990s.[23] Those

18 Perkins, E. J., 1971, "The divorce of commercial and investment banking: a history," 88 Banking Law Journal 483.

19 Coates, J. C., 2009 "Reforming the taxation and regulation of mutual funds: a comparative legal and economic analysis," 1 Journal of Legal Analysis 591.

20 Public Law 511, 84th Congress, ch 240, 2d Session, HR 6227: An Act to Define Bank Holding Companies, Control their Future Expansion, and Require Divestment of their Nonbanking Interests.

21 Federal Deposit Insurance Act s 24 (generally limiting insured bank activities).

22 Different exemptions and regulatory interpretations were exploited by national banks to offer discount brokerage services in a subsidiary (1974), sponsor pooled investment funds equivalent to closed-end funds (1971), offer variable annuities (including equity-like returns) (1993), become members of securities exchanges (1986), advise investment companies (1987), lend securities (1986), manage collective investment retirement accounts (1986) and privately place commercial paper (1989). See Carpenter, D. H., and M. M. Murphy, 2010, "Permissible securities activities of commercial banks under the Glass-Steagall Act and the Gramm-Leach-Bliley Act," Congressional Research Service Report, April 12. Available at:www.crs.gov accessed 29 August 2015. Bank holding companies and non-bank subsidiaries became even more aggressive in pursuing exemptions and interpretations of this kind. Ibid.

23 As a result of these renegotiations, JP Morgan—despite being a commercial, deposit-taking bank—had emerged as a major investment bank by the mid-1990s, competing anew for the same business that it had been forced to divest in the Great Depression (forming Morgan Stanley), and led the underwriting of numerous securities offerings, including underwritings for which the author was counsel to the bank. Chernow, R., 2001, The House of Morgan: an American banking dynasty and the rise of modern finance, Grove Press, tells some of this history well.

negotiations were necessary because, as with the Volcker Rule, the Glass-Steagall Act contained a number of vague terms and phrases, such as "control," "dealing," "affiliated," and "engaged principally."[24] It required follow-up legislation in 1935 to clarify and resolve inconsistencies contained in the initial statute.[25]

The Gramm-Leach-Bliley Act of 1999,[26] the product of the bold attempt at "corporate nullification" of the Glass-Steagall Act and Bank Holding Company Act by Citigroup,[27] had a more dramatic effect. Even it, however, only partially relaxed structural constraints on U.S. banks, and its primary effect was to allow bank holding companies to move into the insurance business, and not to repeal or reverse the Glass-Steagall Act,[28] which (as noted above) had already largely been renegotiated as a major constraint on the ability of large commercial banks to move back into investment banking through affiliates. More importantly, however, is what the Gramm-Leach-Bliley Act did not do. It did not eliminate the basic structural law constraining banks to financial activities.[29] In the period leading up to the financial crisis bank holding companies were, and today still are, limited by this core structural constraint – they did and must still confine their activities to "financial" activities, and are not permitted to engage in manufacturing, trade, or commerce more generally. This core constraint is carefully circumscribed and in some instances, elaborately specified ways – for example, for temporary periods after foreclosure of assets used to secure debts,[30] or pursuant to the capital-limited ability to make merchant banking investments in non-financial portfolio companies.[31] Each of these exemptions is rounded out with lengthy regulations and interpretations.

The Volcker Rule as structural law

In sum, throughout Anglo-American history, structural laws were routinely used to confine systemically important activities (deposit-taking, money markets, and payment systems) to a limited set of entities. These laws, with all the same kinds of ambiguities and line-drawing difficulties that any structural law will create, have thus been operative long before the Volcker Rule was conceived. Paul Volcker would have known the outlines of that history, and it may be presumed to be part of the reason he proposed the rule that carries his name. As one of the most reputable central bankers in U.S. history, Volcker would have known about the long-standing structural limits on banks in the U.S. At a high level of generality, the Volcker Rule is of a kind with many long-standing structural laws in the financial sector, some of which endured, some of which did not.

What, then, briefly is the Volcker Rule, and how does it compare to its predecessor structural laws? The Volcker Rule is an attempt to reduce the risk and improve the governance of U.S. "banking entities" – essentially deposit-taking banks and companies that control such banks – by channeling them into the most basic and traditional core functions of banking – financial intermediation and lending – and away from two types of speculation – "trading" for the account of the bank and indirect investments through unregulated collective investment funds.[32] More specifically, it bans banking entities from engaging in "proprietary trading" or holding "ownership interests" in hedge or private equity funds, subject to a number of exceptions.[33] These definitions were to be further specified by the relevant banking agencies pursuant to delegate rulemaking authority, which (as is conventional) allows for further derogations and interpretations over time. Specific regulations implementing the Volcker Rule were approved (after many delays) in

24 12 USC s 24 (Seventh) (1997).

25 Banking Act of 1935 (23 August 1935) ch 614, 49 Stat 684. This law was 37 pages long, but in fairness, only about a third of it was devoted to amending and clarifying the Glass-Steagall Act, as it also substantially reorganized the Federal Reserve Board structure.

26 Pub L 106-102, 113 Stat 1338, enacted 12 November 1999.

27 On Citigroup"s bold move to buy control of The Travelers, despite being then limited to core banking activities, see PBS, Frontline, "The long demise of Glass-Steagall." Available at: http://tinyurl.com/owk6j accessed 29 August 2015. The phrase "corporate nullification" is from an astute recent article primarily about the high technology sector, but applies to this earlier effort in the financial sector. Pasquale, F., and S. Vaidhyanathan, 2015, "Uber and the lawlessness of "sharing economy" corporates," The Guardian, 28 July.

28 cf Chatterjee, R. R., 2011, "Dictionaries fail: The Volcker Rule"s reliance on definitions renders it ineffective and a new solution is needed to adequately regulate proprietary trading," 8 BYU International Law & Management Review, 8:1, 33-62 ("The 1999 passage of the Gramm-Leach-Bliley Act, also known as the Financial Services Modernization Act of 1999, effectively reversed the changes made by the Glass-Steagall Act.").

29 Unfortunately, it also did not reform or modernize the resolution regime for holding companies or non-bank subsidiaries of financial holding companies, even though it created legal incentives for banking organizations to move more financial activities, liabilities and risks out of banks into those entities. This was one of the biggest regulatory weaknesses of the U.S. financial system during the crisis. Whether it has been fixed by the Dodd-Frank Act remains the subject of active debate. See, for example, Roe, M. J., and S. Adams, 2015, "Restructuring financial firms in bankruptcy: selling Lehman"s derivatives portfolio," 32 Yale Journal of Regulation 2.

30 12 USC s 1843(c)(2).

31 12 USC s 1843(c)(2).

32 For the definition of "banking entity," see ibid s 1851((h)(1). Banks that limit their deposit-taking activities to trust-related activities are generally exempt. ibid. The relevant regulatory agencies are not given authority in the statute to exempt entities from this definition. Certain non-bank financial companies supervised by the Federal Reserve Board are also covered, as are foreign banks treated as bank holding companies under s 8 of the International Banking Act of 1978. See ibid s 1851(a)(2) and (h)(1).

33 "Proprietary trading" is defined as "Hedge fund" and "private equity fund" are statutorily defined as any fund that would be an investment company under as defined in the Investment Company Act of 1940, 15 USC 80a-1ff, but for s 3(c)(1) or 3(c)(7) of that Act, 15 USC 80a-3 (c)(1) and (7), which in general terms exempt "private" funds – i.e., those that are not marketed to the public. The terms also include "similar funds" as determined by regulation by the federal banking agencies, the SEC and the Commodity Futures Trading Commission (CFTC). The Volcker Rule explicitly permits many types of socially useful or deemed not to be speculative trading, such as trading in U.S. government and municipal securities, market-making and underwriting of corporate debt and equity based on the reasonably expected near-term demand of clients.

December 2013, were finalized on 1 April 2014 and were largely (if not wholly) effective as of 1 July 2015.[34]

On the surface, the Volcker Rule may not appear to be a structural law. It appears to consist of a simple command – do not engage in the specified activities. However, it is clear that the goal of Mr. Volcker and other supporters of the rule was not to suppress the activities so banned, which remains legal for non-banking entities. Rather, the goal was to increase the reliability and safety of large banks' more traditional activities. In so doing, the rule is meant to work a change in the organizational culture of banks, and so indirectly to reduce the interconnectedness of banks from other, riskier components of the capital markets. By reducing the need to rely on a bonus-culture conventional on trading floors, the goal was also to dampen the incentives of individual bankers to take risk, and to reduce the power of traders within banks. By reducing or at least flattening the growth in compensation flowing to the part of the financial sector underwritten in numerous ways by the tax-paying public, the rule had the less obvious likely effect of reducing the power and influence of banks generally, and to reduce moral hazard implicit in such government support. Whether these goals will be realized by the Volcker Rule, and at what cost, remains to be seen, as discussed below. But as so understood, the Volcker Rule is indeed a quintessential structural law, directly analogous to structural banking laws dating back to the 17th century, as reviewed above.

Others may not have understood the fact and influence of these structural laws on Anglo-American financial history. Academic economists in the Obama administration, who only grudgingly included the Volcker Rule in the Dodd-Frank Act,[35] may have been under the misapprehension that the sole structural law of consequence in U.S. financial history was the Glass-Steagall Act, which had failed to prevent investment banks (Lehman, Bear Stearns) from being sufficiently interconnected with commercial banks to threaten the financial system as a whole in the crisis. Alternatively, they may have been concerned that any nation-specific structural law would be doomed to fail under the forces of international competition or rent-seeking stakeholders. Outside the administration, the market-oriented ideologues whose voices are the loudest in their critiques of the Volcker Rule may still fail to appreciate the roles that structural laws have always played in creating money, reducing systemic risk and accomplishing political settlements that simultaneously enable and constrain the financial sector. Fantasies of anarchic "golden ages" in finance are common in certain circles.[36]

To be sure, the Volcker Rule is an innovation. The structures it seeks to impose on the financial markets are distinct from those imposed by prior laws. Its expected effects cannot be understood in isolation from other, equally innovative legal reforms contained in the Dodd-Frank Act. As one element of that law, the Volcker Rule represents a novel effort to require banks to be "more focused on the business of banking, so they are better able to serve as safe places for families to deposit their savings and to extend credit to consumers and businesses."[37] As an innovation, the Volcker Rule cannot be evaluated based solely on the structural precedents described above, about which there remains much that is unknown, in any event. However, the history of structural laws must be part of any fair-minded assessment of the Volcker Rule's goals, promise and likely costs and benefits.

The Volcker Rule versus the Glass-Steagall Act

Many seem to think the Volcker Rule represents not an innovation but a "watered down" version of the Glass-Steagall Act.[38] This claim is not entirely wrong. The efforts in both laws to curtail speculative and presumptively risky behavior are similar, and the Volcker Rule will contain a larger number of exceptions and industry accommodations from the outset than did the Glass-Steagall Act, consistent with it being more "watery" than its predecessor.

But the claim is not accurate either. The specific activities targeted by the Volcker Rule are not a subset of those targeted by the Glass-Steagall Act, but overlap with them. As a result, the likely consequences for banks and their affiliates generally also differ. For example, proprietary trading of corporate "investment securities"[39] would have been permitted for non-bank affiliates of banks under

34 The banking agencies and the SEC issued a joint final rule. Prohibitions and Restrictions on Proprietary Trading and Certain Interests in, and Relationships with, Hedge Funds and Private Equity Funds, Department of the Treasury and others (2013). Available at: www.sec.gov/rules/final/2013/bhca-1.pdf accessed 29 August 2015 [hereinafter Joint Volcker Rule Release] (to be codified at 12 CFR pt 44 (OCC); 12 CFR pt 248 (Fed Reserve); 12 CFR pt 351 (FDIC); 17 CFR pt 255 (SEC)). The CFTC issued a final rule separately. Prohibitions and Restrictions on Proprietary Trading and Certain Interests in, and Relationships with, Hedge Funds and Private Equity Funds, CFTC. Available at: www.cftc.gov/ucm/groups/public/@newsroom/documents/file/federalregister121013. pdf accessed 29 August 2015 (codified at 17 CFR pt 75). Extensions were granted by Board of Governors of The Federal Reserve System, Order Approving Extension of Conformance Period Under s 13 of the Bank Holding Company Act (18 December 2014).

35 Krawiec K. D., and G. Liu, 2015, "The Volcker Rule: a brief political history," 10 CMLJ 507.

36 cf Desan (n 6), critiquing such fantasies about the history of money in modern economies.

37 U.S. Department of Treasury, 2015, "Dodd-Frank at five years," July. Available at: http://tinyurl.com/oh8xqg6 accessed 29 August 2015.

38 E.g., Sekar, A., 2015, "The Glass-Steagall Act explained." Available at: http://tinyurl.com/p3ujnbz accessed 29 August 2015.

39 Banks themselves could invest in, but not underwrite, "investment securities" under the Glass-Steagall Act. They included most short-term, marketable debt securities, including those issued by private corporations. Under 12 USC s 24 (Seventh), "investment securities" was defined to mean "marketable obligations, evidencing indebtedness of any person, copartnership, association, or corporation in the form of bonds, notes and/or debentures commonly known as investment securities under such further definition of the term "investment securities" as may by regulation be prescribed by the Comptroller of the Currency." That definition is broader than the government securities exception in the Volcker Rule. More importantly, consistent with the Glass-Stegall Act, as interpreted over time, bank affiliates could engage in proprietary trading prior to the Volcker Rule.

the Glass-Steagall Act, but will not generally be permitted under the Volcker Rule. A general securities underwriting or dealing business, by contrast, was prohibited as the Glass-Steagall Act was initially implemented, but came to be gradually permitted during the 1970s, 1980s, and 1990s, as described above, and would survive the Volcker Rule largely intact. Equity investments for the proprietary account of a bank or holding company were generally banned under the Glass-Steagall Act, and generally continue to be so banned for banks, and short-term equity investments – that is, "proprietary trading" in stocks – would be banned under the Volcker Rule for all covered banking entities, whereas longer term merchant banking investments for non-bank subsidiaries of financial holding companies would continue to be permitted under the Volcker Rule, but only to the extent permitted by the Bank Holding Company Act, as modified by the Gramm-Leach-Bliley Act, and subject to the Volcker Rule ban on investments in private investment funds.

The casino-like speculative culture of banks was the focus of the concerns expressed by Paul Volcker at a roundtable held at Harvard University during the crisis, which the author attended, and presumably the same focus of the rule writers in the drafting process leading to the Dodd-Frank Act. It is reflected in Mr Volcker's comments on the proposed Volcker Rule regulations: "The need to restrict proprietary trading is not only, or perhaps most importantly, a matter of the immediate market risks involved. It is the seemingly inevitable implication for the culture of the commercial banking institutions involved, manifested in the huge incentives to take risk inherent in the compensation practices for the traders. Can one group of employees be so richly rewarded, the traders, for essentially speculative, impersonal, short-term trading activities while professional commercial bankers providing essential commercial banking services to customers, and properly imbued with fiduciary values, be confined to a much more modest structure of compensation?"[40]

Changing "culture" may strike some readers as a soft, vague, or secondary goal for policy. But financial economists have over the past 20 years come to recognize the way that organizations, including how they screen for and manage employees, and the values and incentives they create, can generate first-order effects in the capital markets. Organizational culture of this kind is not easily understood through the lens of myopic cynicism, or even with the useful – but for that very reason limited – standard working assumptions of neoclassical economic models.[41] If one thinks that bonus-driven speculative trading, even a (relatively) safe financial asset such as corporate debt, can create a casino-like atmosphere inside a bank, and undermine its governance,[42] then the Glass-Steagall Act's approach would fail to address that problem, whereas the Volcker Rule would. If one thinks that equity underwriting or customer-driven market making is excessively speculative and risky, then the Volcker Rule will not address that problem, whereas the Glass-Steagall Act (if strictly enforced) would. Neither is clearly a watered down approach for the other.

A full discussion of the many nuances of the line between conduct permitted and prohibited for covered entities under the Volcker Rule, and how if at all regulators will address unintended consequences of the rule and the implementing regulations, is beyond the scope of this article.[43] Indeed, one standard complaint about the rule – made not primarily by lawyers who profit from its complexity, but by bank managers, business journalists, and Paul Volcker himself – is that the implementing regulations are too complex and long, requiring 71 pages, and 892 pages more in the accompanying releases.[44]

---

40 Volcker, P. A. 2012, "Commentary on the restrictions on proprietary trading by insured depositary institutions," attached to Letter from Paul A Volcker to financial regulatory agencies, 13 February, http://tinyurl.com/qanzy6d accessed 29 August 2015 (emphases added).

41 cf Summers, L. H., and W. R. Easterly, 1992, "Culture is not to blame," Financial Times, 15 April ("The primacy of economic incentives over [national] culture is good news for courageous reformers."); with Hermalin, B. E., 2000, "Economics and corporate culture," in: Cooper, C. L., S. Cartwright and P. C. Early (eds.), The international handbook of organizational culture and climate ("with a few exceptions . . . economists have ignored the issue of corporate culture in their studies of firms and other organizations"); Cohan,W. D., 2015, "Can bankers behave?" The Atlantic, May. Available at: http://tinyurl.com/pto4bu6 accessed 29 August 2015 (former Lazard banker attributing change in culture at Morgan Stanley to effects of Volcker Rule); Cohn, A., E. Fehr and M. A. Marechal, 2014, "Business culture and dishonesty in the banking industry," 516 Nature 86 ("Employees of a large, international bank behave, on average, honestly in a control condition. However, when their professional identity as bank employees is rendered salient, a significant proportion of them become dishonest."); and Guiso, L., P. Sapienza and L. Zingales, 2015, "The value of corporate culture," 117 Journal of Financial Economics 60 ("With few notable exceptions, the finance literature has ignored the role corporate culture can play").

42 See Lewis, M., 1989, Liar"s poker, Hodder Paperbacks (detailing culture of bond traders and its effects on Salomon Brothers).

43 For example, application of the Volcker Rule to non-U.S. entities and activities is itself a complex topic. See, for example, Massari, J., 2015, "Foreign bank cross-border securities trading under the Volcker Rule: exploring the trading outside the United States exemption"s unintended consequences," 10 Capital Markets Law Journal 523; Yoshiya, S., 2015, "The Volcker Rule: regulatory challenges and unintended consequences for banks in Asia," 10 Capital Markets Law Journal 542. Likewise, the Volcker Rule will require ongoing collaboration between the regulatory agencies that have traditionally overseen banks and those that have traditionally overseen financial markets. Dombalagian, O. H., 2015, "The Volcker Rule and regulatory complementarity," 10 Capital Markets Law Journal 469.

44 For example, Sloan, A., 2013, "The Volcker Rule: a triumph of complexity over common sense," Washington Post, December 19 (criticizing rule as too long and complex); Culp, S., 2013, "Final Volcker Rule leaves facing compliance hurdles," Forbes, December 17; Wallison, P. J., 2013, "Why the Volcker Rule will harm the U.S. economy," The American, December 13 (criticizing rule as too long and complex); Bobelian, M., 2013, "Will the Volcker Rule work?" Forbes, December 11 (noting length and complexity); Armstrong, R., 2011, "Paul Volcker says Volcker Rule too complicated," Reuters, November 9 (quoting Volcker as criticizing complexity of rule and attributing it to bank industry lobbyists).

But in both spirit and level of detail the Volcker is not different in kind from the structural laws described above. As for the spirit, both the precedents and the Volcker Rule work in three similar, structural ways. First, they banned some set of activities for designated entities, with the goal of encouraging the remaining activities, while reducing their risk. Secondly, for those same entities, they created or preserved government subsidies. The subsidies were both explicit (for example, federal deposit insurance, access to the Fed's payments system, ability to borrow from the Fed's discount window) and implicit (too big to fail), as well as barriers to competition from non-banks (for example, requirements of bank charter and regulation for deposit-taking institutions). But because those subsidies were now flowing to entities limited in their power and activities, the subsidies would be more likely to have the public-regarding benefits they were intended to have, and less likely to cross-subsidize risky and less publicly valuable activities, or enrich private citizens at taxpayer expense. Thirdly, the laws imposed special and often detailed, lengthy, and complex regulatory requirements, such as capital requirements[45] and bank supervision, some of which functioned to reinforce the structural nature of the laws; changing the nature of the banks' activities indirectly, rather than through simple command and control obligations.

As for the level of detail, some readers may demur. They will say, as some have, that the Glass-Steagall Act was nice and simple and short (merely 37 pages!), while the Volcker Rule regulations are long (over 900 pages!).[46] The comparison is silly. The Volcker Rule in the Dodd-Frank Act itself is short (only 11 pages!), shorter than either the Glass-Steagall Act or the Federal Reserve Act. The Glass-Steagall Act regulations,[47] interpretations and case law were sufficiently long, complex, and often inconsistent that banking law texts prior to Glass-Steagall Act's repeal commonly devoted many pages to what still amounted to a highly abbreviated summary of the laws governing securities activities by banks.[48] As noted above, the Glass-Steagall Act also changed significantly in operation over time – reflecting complexity generated by steady pressure from banks to push that law's boundaries wherever profit made it attractive to do so.

In sum, the Volcker Rule is a structural law designed to protect the banking system that is similar in kind to many prior laws shaping the financial sector. That fact may cast some light on how to evaluate it and predict its effects. But before turning to that task, let us first review a distinct set of structural laws that are important to any understanding of how the Volcker Rule will be implemented in practice: the many statutes and regulatory processes that in the U.S. are the general domain of administrative law.

ADMINISTRATIVE LAW AS STRUCTURAL LAW

In possible tension with the traditional use of structural laws to constrain banks and capital markets is a newer set of structural laws, designed to constrain the very government agents responsible for implementing complex modern financial regulations. Administrative law – the body of statutes and court doctrines channeling and controlling the use of law-making power by government officials – grew in importance in the twentieth century. It now occupies a role practically co-equal with the substance of financial regulation in any understanding of how such regulation affects capital markets in practice. Most recently, legal requirements for CBA have come to the fore as part of the administrative law arsenal, posing the question of whether structural laws such as the Volcker Rule can be usefully subject to such analysis.

Ambiguity in structural constitutional laws

As noted at the outset of the prior section, the U.S. Constitution contains an important set of structural laws that constrain the most basic functions of those responsible for making, enforcing and interpreting law. Based on political philosophical commitments to divided and accountable government, these structural laws separate the "legislative powers" the "executive power" and the "judicial power"[49] into

45 Capital requirements sound simple – simpler than dividing proprietary trading from customer-driven market making. But even a quick glance and the huge number of pages devoted to each of the major capital rule reform initiatives sponsored by the Basel Committee should be enough to dispel the idea that capital rules are short and simple in practice. See Basel Committee on Banking Supervision, Bank for International Settlements (http://www.bis.org/bcbs) accessed 29 August 2015.

46 Culp (n 43).

47 In 1998, prior to the repeal of the Glass-Steagall Act, the formal regulations in Subpart C of the Federal Reserve Board"s Regulation Y (http://tinyurl.com/nuylocz) accessed 29 August 2015, which governed non-banking activities of bank holding companies, alone took up more than 25 single-spaced narrow-margin pages. A single statement of guidance from the Federal Reserve Board in 1998 regarding securities activities of banks took up 14 pages. See Supervisory Policy Statement on Investment Securities and End-User Derivatives Activities, July 27. Available at: http://tinyurl.com/o7bq75h accessed 29 August 2015. The Federal Reserve Board"s current web page (http://tinyurl.com/ovf6no6 and accessed 29 August 2015) lists 12 supervisory policy statements on securities activities since 1990, which is an incomplete listing of the relevant guidance from the Fed alone. To that should be added comparable regulations and guidance from the OCC and the FDIC.

48 For example, Jackson, H. E., and E. L. Symons, Jr, 1999, Regulation of financial institutions, West Group, 117-41 (materials showing efforts to define legal bank activities); Herlihy, E. D., et al. 1996, "The new aggressive era in financial institutions mergers and acquisitions," in Mergers and acquisitions of financial institutions 1995: an unprecedented year of consolidation, Practising Law Institute, 48-57, 120, 154-56 (discussing aspects of law relevant to bank acquisitions of securities firms prior to repeal of Glass-Steagall Act). A search of Westlaw returns over 100 Federal court decisions running to more than 1000 pages, which interpreted the Glass-Steagall Act prior to its repeal. This does not count the pages of formal regulations proposed or adopted and informal guidance provided under that law.

49 U.S. Constitution, art I s 1; art II, s 1; art III, s 1.

three distinct branches of government, and similarly layer those powers in two levels, federal and state.[50] For the Republic's first 100 years, these ambiguous and inconsistent commitments generated disputes and conflicts, some resolved but many deferred, suppressed or ignored, only to erupt even more violently over time. For example, the U.S. Civil War can be attributed in part to the decision to avoid reconciling the entanglement of some but not all states with slavery, on the one hand, with a clear statement of national unity reflected in the supremacy of national laws over state laws, on the other hand.

Another dispute suppressed during the Constitutional ratification process was the authority and propriety of the federal government to create a central bank.[51] The initial suppression of this dispute led to controversies in the Washington administration, when Alexander Hamilton sought to enhance the country's financial capacities through a strong U.S. Treasury and the First Bank of the United States in 1791. When the First Bank's charter expired, it was not renewed, in part of because of the controversy over its legality.[52] After the Second Bank was created in 1816 and the state of Maryland sought to tax it, basic structural controversies over both separation of powers and federalism as applied to the financial sector found their way to the Supreme Court. In the landmark legal dispute of McCulloch v Maryland,[53] Justice Marshall interpreted article I's "necessary and proper" clause[54] generously for the national government, and at the same time took a narrow view of the states' residual powers where they arguably interfered with those of the government of the country as a whole.

Another set of latent conflicts created by the ambiguous and inconsistent commitments to divided government, however, did not arise in full form until the role of government generally began to expand. In the late nineteenth and early twentieth centuries, in response to the massive social and economic effects of the Industrial Revolution and the rise of corporate capitalism, the U.S. entered an "Age of reform."[55] State and federal governments alike began to enact new kinds of laws addressing shipping, industrial accidents, wages, working conditions, labor, immigration and – as noted above – money markets and banking. Beginning with the Civil Service Act and the Civil Service Commission in 1883, and then the Interstate Commerce Commission in 1887, efforts to isolate government employees and agents from cronyist and partisan pressures led to a wave of civil service protections and the creation of "independent" government agencies, often elaborately designed to achieve political compromise over the expected distribution of power they were expected to wield.[56]

The majority of the financial regulatory agencies – the Federal Reserve Banks and (later) the Board, the SEC, the FDIC, the CFTC and (most recently) the Consumer Financial Protection Bureau – are all examples of such "independent" agencies.[57] Even the Office of the Comptroller of the Currency (OCC), which (as part of the Department of the Treasury) had long functioned as a core part of the "executive" branch, overseeing national banks, was re-identified as an "independent" agency in the Dodd-Frank Act.[58] Among the typical features of independent agencies are multi-member commissions with staggered terms, and sometimes specifications of party, designed to prevent any one administration from effecting wholesale change in their policies. The banking agencies (but not the SEC or the CFTC) also have effective budget autonomy, giving them substantially more discretion than agencies that have to persuade Congress to refund them every year.[59]

As the role of government grew, along with the role of "independent" agencies, numerous battles were fought in the courts over whether laws passed by legislatures or regulation adopted by agencies were constitutional.[60] When two world wars and the Great Depression led to even more innovation and expansion of public administration of what had previously been private activity, the court

50 Ibid s 8 (enumerating powers of Congress).

51 Klarman, M., 2015, working paper

52 See Hammond (n 5).

53 17 U.S. 316 (1819). The First Bank also was involved in legal controversy, over whether it could sue in its own name, or whether its president, directors and shareholders, residing in one state, could sue on its behalf citizens of another state, in federal court, to recover stolen property. In Bank of the U.S. v Deveaux, 9 US 61 (1809), the Court held that the Constitution prevented the Court from expanding its jurisdiction beyond that established by Congress, and that the Bank"s charter"s terms implicitly denied it standing to sue itself as a "citizen" in federal court. At the same time, the Court held that because the U.S. Constitution, like all constitutions, "from its nature, deals in generals, not in detail," and should be interpreted in that light, with the result here that the bank could serve as a placeholder in a suit by its president, board and shareholder-citizens.

54 US Constitution, art I, s 8.

55 Hofstadter, R., 1955, The age of reform: from Bryan to F.D.R., Knopf, 23-93; Sanders, E., 1999, Roots of reform: farmers, workers, and the American state, University of Chicago Press, 1877-917; Trachtenberg, A., 1982, The incorporation of america: culture and society in the gilded age, Hill & Wang

56 Skowronek, S., 1982, Building a new American state: the expansion of national administrative capacities, Cambridge University Press, 1877-920; Keller, M., 2007, America"s three regimes: a new political history, Oxford University Press; Lowi, T. J., and N. K. Nicholson, 2009, Arenas of power: reflections on politics and policy, Routledge

57 See Conti-Brown (n 16).

58 Dodd-Frank Act, s 315.

59 United States General Accounting Office, 2002, "SEC operations: implications of alternative funding structures," GAO-02-864. Available at: http://www.gao.gov/new.items/d02864.pdf

60 These battles are typically encapsulated as the "Lochner era," referring to Lochner v New York 198 U.S. 45 (1905). See also Allgeyer v Louisiana 165 U.S. 578 (1897) (corporations have liberty of contract, and the due process clause of the Fourteenth Amendment prevents a state from barring a corporate "citizen" from mailing a notice describing goods it seeks to insure under a policy issued by a foreign insurance company); Reagan v Farmers" Loan and Trust Co 154 U.S. 362 (1894) (railroad corporations could not be required to charge less than the tariff proposed by the state railroad commission under due process clause if it would leave the railroad unable to pay its debts); Conn Gen Life Ins Co v Johnson 303 U.S. 77, 90 (1938) (Black, J, dissenting) (collecting cases).

battles came to a head, resulting in the end of the "Lochner era" and a great retreat by federal courts from attempting to curtail the exercise of economic regulatory power, whether through the legislatures or the agencies.[61] As part of the political settlement over this retreat, however, and increasingly over time as progressive advocates of active government found themselves disappointed with the behavior of regulatory agencies,[62] a new body of structural laws emerged. Now generally labelled "administrative law," these structural laws variously intended to constrain or improve the functioning of what has come to be called the "Fourth Branch."

Major components of administrative law

Chief among the structural components of administrative law in the U.S. is the Administrative Procedures Act (APA).[63] Coupled with a residual, if uncertain, "right of review" by courts of agency decisions and reinforced by a "presumption of reviewability,"[64] the APA has given courts (and hence, private plaintiffs) a varying, but at times important role in checking the process and at times substance of financial regulation. The APA (among other things) imposes specified procedures for agencies to follow before enacting rules. Absent clear Congressional direction, courts have held that rules are presumptively reviewable by courts for adherence to statutory commands and process regularity.

In addition, with greater controversy and less consistency, courts have subjected agency regulations to substantive "hard look" review, testing them by asking if they are "arbitrary and capricious" or otherwise fail to respect the minimal demands of rationality. In principle, courts have self-imposed limits on their own roles, by stressing the need to defer to agencies on a variety of questions, including statutory interpretation[65] and rationality of agency rules.[66] Observers of the courts have at times criticized them for exceeding or applying these limits in inconsistent ways, with the result that at times neither legislatures nor agencies but courts – neither accountable nor expert – have become the ultimate rulemakers for the capital markets.[67]

Reinforcing the role of courts in reviewing agency decisions, and increasingly in the last quarter of the twentieth century, legal mandates have emerged for the conduct, interagency review and publication of CBA.[68] CBA – or more generally, regulatory impact analysis – is a component of the process that agencies commonly follow in considering proposed rules, and (for executive agencies) a legal requirement.

CBA of financial regulation

CBA of financial regulation (CBA/FR) has emerged as an important topic in policy and legal debates,[69] due in part to the unprecedented number and importance of new regulations called for by the Dodd-Frank Act, including the Volcker Rule.[70] Interest groups seeking to delay and shape those regulations have joined a set of policy entrepreneurs and academics whose long-term project has been to spread the use of CBA generally. A related but partially distinct group of political entrepreneurs has the long-term and largely

61 Standard histories treat the Lochner era as ending in the late 1930s, with West Coast Hotel Co v Parrish 300 U.S. 379 (1937), upholding minimum wage legislation and overturning Adkins v Children"s Hospital 261 U.S. 525 (1923); United States v Carolene Products Co 304 U.S. 144 (1938) (legislative authority over economic matters plenary, entitled to presumption of constitutionality), and cf Humphrey"s Executor v United States 295 U.S. 602 (1935) (President may not remove officer of "quasi-legislative" independent agency) with Myers v U.S. 272 U.S. 52 (1926) (finding unconstitutional law requiring advice and consent of Senate for President to remove executive branch official, a postmaster); see also Nebbia v New York 291 US 502 (1934) (upholding price controls over milk); Landis, J., 1938, The administrative process, New Haven: Yale University Press, 15-46 (articulating legality and advantages of multimember, bipartisan, expert independent agencies).

62 Rodriguez, D. B., 1997, "Jaffe"s law: an essay on the intellectual underpinnings of modern administrative law theory," 72 University of Chicago-Kent Law Review 1159; Shapiro, M., 1986,"The APA: past, present, future," 72 Virginia Law Review 447; Landis, J., 1960, Chairman of the Subcommittee on Administrative Practice and Procedure, Senate Committee on the Judiciary, 86th Congress, 2d Session, Report on Regulatory Agencies to the President-Elect; Stewart, R., 1975, "The reformation of American administrative law," 88 Harvard Law Reviwew 1667.

63 5 USC ss 500-96 (1946).

64 Jaffe, L. L., 1965, "The right to judicial review," in Jaffe, L. L., (ed.) Judicial control of administrative action, Little, Brown & Co., ch 9; Krent, H., 1997, "Reviewing agency action for inconsistency with prior rules and Regulations," Chicago-Kent L Rev 1187; Abbott Laboratories v Gardner 387 U.S. 136 (1967) (presumption of court reviewability); Motor Vehicle Mfrs Ass"n v State Farm Mut Auto Ins Co 463 U.S. 29, 43-44 (1983) (standard for "hard look" review by courts of agency decisions).

65 Chevron USA, Inc v Natural Resources Defense Council, Inc 467 U.S. 837 (1984). For a recent Supreme Court case in which Chevron deference seemed not to play a significant role in limiting the court"s involvement, see Michigan v EPA 576 U.S. (2015).

66 Motor Vehicle Mfrs Ass"n v State Farm Mut Auto Ins Co (n 63); Citizens To Preserve Overton Park, Inc v Volpe 401 US 402 (1971). For a discussion of the relationship between Chevron and "hard look" review, see Stephenson, M. C., and A. Vermeule, 2009, "Chevron has only one step," 95 Virginia Law Review 597.

67 Ahdieh, R. B., 2013, "Reanalyzing cost-benefit analysis: toward a framework of function(s) and form(s)," 88 NYU Law Review 1983; Cox, J. D., and B. J. C. Baucom, 2012, "The emperor has no clothes: confronting the D.C. Circuit"s usurpation of SEC rulemaking authority," Texas Law Review 1811, 1840; Fisch, J. E., 2013, "The long road back: business roundtable and the future of SEC rulemaking," Seattle University Law Review ssrn.com/abstract¼2164423; Sunstein, C. R., and A. Vermeule, 2014, "Libertarian administrative law," 29 June, working paper, June 29. Available at: http://ssrn.com/abstract¼2460822, accessed 1 July 2014.

68 Adler, M. D., and E. Posner, 2006, New foundations of cost-benefit analysis, Harvard University Press; Sunstein, C. R., 2003, Risk and reason: safety, law and the environment, Cambridge University Press; Sunstein, C. R., 2002, The cost-benefit state: the future of regulatory protection, ABA Book Publishing. In the financial regulatory context, see Coates, J. C., 2015, "Cost-benefit analysis of financial regulation: case studies and implications," 124 Yale Law Journal 1, 913-26; Kraus, B., and C. Raso, 2013, "Rational boundaries for SEC cost-benefit analysis," 30 Yale Journal of Regulation 289, 342; Lee Y-H. A., 2015, "SEC rules, stakeholder interests, and cost-benefit analysis," 10 Capital Markets Law Journal 311.

69 See, for example, Symposium, 2014, "Developing regulatory policy in the context of deep uncertainty: legal, economic, and natural science perspectives," 43 Journal of Legal Studies (including several articles on the topic of CBA of financial regulation); "The administrative law of financial regulation," 2015, 78 Law & Contemporary Problems (including several articles on the topic of CBA of financial regulation); Colloquium, 2014, "Critiquing cost-benefit analysis of financial regulation," George Washington Law.

70 The full title of this statute is the Dodd-Frank Wall Street Reform and Consumer Protection Act (n 1) (codified as amended in scattered sections of 12 USC) (Dodd-Frank Act).

partisan project of embedding CBA/FR in judicial review of regulations under the APA.[71] White papers calling for CBA/FR have elicited academic symposia and multidisciplinary efforts to study and improve CBA/FR, while a continuing flow of bills have been introduced in Congress to require or empower the President to mandate CBA/FR. A few of these bills have received at least some bipartisan support, even as some judges on the DC Circuit continue to use CBA as a tool for intervening in regulatory contests.[72]

In the U.K., the two main financial regulatory agencies are required by statute to conduct quantified CBA/FR, unless in the opinion of the agencies the costs or benefits "cannot reasonably be estimated" or "it is not reasonably practicable to produce an estimate," in which case the agency must publish its opinion and explain it.[73] In striking contrast to the recent U.S. experience, however, courts have not repeatedly overturned rulemakings by the old Financial Services Authority (FSA) and its successors for inadequate CBA. A rare example of a court decision even referring to CBA by the FSA is R (on the application of the British Bankers Association) v FSA et al.,[74] which rejected a challenge by a banking trade group to the handling of complaints about "Payment Protection Insurance" by the FSA and the Financial Ombudsman Service, which handles consumer financial complaints.

How might these efforts play out in the context of a structural law such as the Volcker Rule? The answer to that question – analyzed in the final section of this article – may help guide future efforts to assess the costs and benefits of CBA/FR itself, and so to guide the intersection of structural laws governing banking and structural laws governing administrative agencies.

Evaluating administrative law's effects

All of these components of administrative law have the effect of constraining regulatory discretion, and the potential to improve regulatory decisions. They also all have the cost, however, of slowing down regulatory action, and potentially hiding from the public the goals and effects of regulation (or de- or re-regulation) generally under a veneer of legalistic or technocratic analysis. They also have the potential cost of putting regulations at risk of the same kind of judicial second-guessing reflected in the Lochner era,[75] or at risk of the same kinds of partisan or cronyist influences that independent agencies were designed to combat. At different times in legal history, they have functioned as tools for unhappy pro-regulatory lobbies to try to nudge agencies to be more vigorous in protecting the public, or as tools for unhappy anti-regulatory lobbies to try to slow down or blunt the effect of new regulatory efforts.[76]

A policy-minded citizen trying to evaluate the effects of administrative law – or more plausibly, the effects of one of its structural components – would need to conduct a meta-CBA – to ask if the benefits of these administrative law constraints on regulatory action outweigh their costs? At least part of that meta-analysis would require, in turn, a careful consideration of what CBA can practically achieve, in the context of specific regulations, such as the Volcker Rule.

CBA AND THE VOLCKER RULE

This last section of the article takes up the following, related questions: could the regulations needed to implement a complex, ambitious structural law such as the Volcker Rule be the subject of useful CBA? If so, would that analysis consist solely of the identification of qualitative effects of the rule, or could it usefully contain a precise and reliable quantification of those effects? Would a requirement or expectation of such analysis be expected to enhance and detract from the regulatory process for the Volcker Rule? Would such analyses create a means of constraining agency discretion and improving agency accountability, or give the agencies cover for using crude guesstimates to camouflage the likely effects of the rule? Would they only impose unnecessary and pointless delays, or give partisan or cronyistic enemies of the public-regarding goals of the law weapons to undermine its effectiveness in court or in an inter-agency process? The answers to these questions must remain somewhat speculative, but if CBA/FR is clearly ripe for implementation, the potential for CBA of a structural law like the Volcker Rule should at least be susceptible to qualitative assessment.

Administrative law requirements relevant to the Volcker Rule

To be clear, the independent financial agencies have largely been exempted from CBA requirements. The formal releases published by the financial agencies in the Federal Register contain no general CBA/FR of the Volcker Rule. Legally, the financial agencies are subject to no general CBA/FR mandate, and the statutory requirement for and authorization of the regulations implementing the Volcker Rule are part of the Bank Holding Company Act of 1956,[77]

71 Pub L 79-404, 60 Stat 237 (codified as amended in scattered sections of 5 USC).
72 Coates (n 67) 882.
73 Financial Services Act, 2012, amending inter alia s 138I (Financial Conduct Authority) and 138J (Prudential Regulation Authority) of the Financial Services and Markets Act 2000.
74 [2011] EWHC 999 (Admin).
75 See n 59 above.
76 See citations in n 56 above.
77 Bank Holding Company Act of 1956, Pub L No 84-511, 70 Stat 133 (codified as amended in scattered sections of 12 USC). The Bank Holding Company Act of 1956 (BHCA) contains a broad regulatory delegation of authority to the Federal Reserve Board to "issue such regulations and orders as may be necessary to enable it to administer and carry out the purposes" of the Act and to "prevent evasions thereof." Ibid s 5, 70 Stat 137.

which does not contain even the loose kind of requirement in the securities laws that the SEC consider "efficiency" or in the commodities laws that the CFTC consider costs and benefits.[78] Nothing in the language of section 619 of the Dodd-Frank Act itself required CBA of the regulations.[79] The formal rulemaking contained limited cost-related information in its analyses under two minor components of administrative law – the Regulatory Flexibility Act and the Paperwork Reduction Act[80] – but no information about benefits or non-compliance costs.

The OCC's CBA of the Volcker Rule

The OCC, however, did release its own CBA of the Volcker Rule.[81] It identified a number of "non-monetized" (qualitative) benefits from the rule. They included (a) improved supervision by bank regulators, due to metrics reporting required by the rule; (b) better management of risk by bank managers (for the same reason); (c) reduced conflicts of interest; (d) protecting "core banking services" and improved bank safety and soundness (reduced risk of bank failures); (e) reduced "tail risk" from trading activities and reduced risks of financial crises; (f) improved corporate governance of banks resulting from reduced stock market liquidity; and (g) reduced harms caused by excess liquidity.[82] As the OCC noted, the "benefits of the regulation can be difficult to quantify including the value of enhanced economic stability."[83]

The OCC also identified a number of costs of the rule. For a subset, the OCC provides quantified estimates: (a) compliance costs (U.S.$405 to U.S.$541 million); (b) additional capital costs for permissible investments in covered funds (U.S.$0 to U.S.$165 million); (c) the OCC's own costs of supervising compliance with the new rule (U.S.$10 million); and (d) a one-time hit to the value of assets owned by banks but restricted by the rule, resulting from reductions in demand for those assets due to the rule. For the last type of cost, the OCC drew on academic research estimating a similar haircut in corporate bond values when bonds are downgraded by credit rating agencies and insurance companies (subject to regulations limiting their ownership of junk bonds) are forced to sell such bonds, deriving a range of costs from U.S.$0 to U.S.$3.6 billion.

However, the types of costs that are likely to be the largest ongoing costs were not quantified. Foremost among these non-quantified costs is the reduced liquidity in markets where banks were significant trading participants, particularly arising from interdealer trading, which is not treated as a permissible source of "customer" demand under the rule.[84] Banks, as a result, will not be able to hold certain assets as "inventory," which will reduce liquidity in the markets for those assets and make it harder for banks to share risk with other banks when permissible customer-driven trading results in banks taking on large blocks of equities. Banks may incur higher costs to hedge or shed those risks, or face more difficulties in managing risks. The reduction in liquidity caused by the ban on inter-dealer trading will likely reduce the depth of those markets and the ability of issuers to raise capital within them.[85] Another potential cost of the rule is similar to one relevant to any structural law making conduct of an activity more difficult or expensive within a bank, including capital rules, for example. That potential cost is the migration of trading activity to non- or less-regulated "shadow" banks, which could pose systemic risks, offsetting (and possibly exceeding) the benefits of risk reduction within the banking system.

78 See 7 USC s 19(a)(1) (2010) (requiring the CFTC to "consider the costs and benefits" of its regulatory actions). This is true even though the SEC and the CFTC were also required to adopt the Volcker Rule, because their authority (and mandate) to do so is (unusually) in the BHCA, not the statutes that traditionally authorize them to act. Office of the Comptroller of the Currency, Analysis of 12 CFR Part 44, U.S. Department of the Treasury (March 2014), available at: http://www.occ.gov/topics/laws-regulations/legislation-ofinterest/volcker-analysis.pdf accessed 29 August 2015.

79 The specific section that authorizes the Volcker Rule, 12 USC s 1851 (2010), added to the BHCA by the Dodd-Frank Act, contains a similarly broad grant of authority and does not condition rulemaking on any particular finding or process, other than (1) to "consider" a statutorily mandated January 2011 study of how to implement the section conducted by the Financial Stability Oversight Council, see 12 USC s 1851(b)(1)-(2)(A); Financial Stability Oversight Council, 2011, "Study and recommendations on prohibitions on proprietary trading and certain relationships with hedge funds and private equity funds," available at: http://www.treasury.gov/initiatives/documents/volcker%20sec%20%20619%20study%20final%201%2018%2011%20rg.pdf accessed 29 August 2015; and (2) to coordinate rulemaking among the Federal Reserve Board, FDIC, OCC, SEC and CFTC so as to "assur[e], to the extent possible, that such regulations are comparable and provide for consistent application and implementation ... to avoid providing advantages or imposing disadvantages to the companies affected ... and to protect the safety and soundness of banking entities and nonbank financial companies supervised" by the Federal Reserve, 12 USC s 1851(b)(2)(B)(ii).

80 Joint Volcker Rule Release (n 33) 928-44 (conducting analysis under the Paperwork Reduction Act (PRA)); ibid 944-48 (conducting analysis under the Regulatory Flexibility Act (RFA)). The American Bankers Association (ABA) and other plaintiffs sued to enjoin enforcement of the Volcker Rule on the ground that the agencies" RFA analysis failed to consider the rule"s "significant economic impact on a substantial number of community banks." See Emergency Motion of Petitioners for Stay of Agency Action Pending Review at 15-16, Am Bankers Ass"n v Bd of Governors of the Fed Reserve Sys, No 13-1310 (DC Cir, 24 December 2013), available at: http://www.aba.com/Issues/Documents/12-24-13ABA EmergencyMotionforStayofVolckerRuleOwnership InterestProvision.pdf. The Joint Volcker Rule Release specifically addressed potential impacts by exempting banks below various specified size thresholds from reporting and compliance burdens. The ABA suit focuses on one indirect effect of the rule, which is to ban "banking entities" (including all depository institutions, small or large) from holding "ownership interests" in hedge and private equity funds (Subpart C of the Volcker Rule), including debt instruments that give holders the right to remove a collateral manager for a collateralized debt obligation – an entity that holds multiple trust-preferred or other securities, which (as the ABA in its papers admits) collapsed in value during the financial crisis. See ibid 2, 7.

81 See Office of the Comptroller of the Currency (n 77).

82 ibid 18-22. The FSOC also identified the benefit that the rule would reduce the risk that banks have effective liability for nominally off-balance sheet funds they sponsor. Financial Stability Oversight Council (n 78).

83 Office of the Comptroller of the Currency (n 77) 1.

84 ibid 15.

85 cf Cox, J. D., J. R Macey and A. L Nazareth, 2013, "A better path forward on the Volcker Rule and the Lincoln Amendment," Bipartisan Policy Center 85. Available at: http://www.sec.gov/comments/s7-41-11/s74111-648.pdf accessed 29 August 2015.

In sum, the OCC's CBA/FR did not include a quantification of the benefits, and only quantified a subset – and likely a small portion – of the costs of the Volcker Rule. The result was that the OCC confidently categorized the rule as "major" for purposes of the CRA,[86] because that categorization only requires bounding the rule's costs, but did not reach any conclusion about the rule's net costs and benefits.

Is a fully quantified CBA of the Volcker Rule feasible?

Could the agencies go beyond conceptual CBA and conduct a fully quantified CBA/FR? The short answer is no. The reason is simple, and derives from the nature of the Volcker Rule as a novel structural law. Because of its nature, there simply are no historical data on which anyone could base a reliable estimate of the benefits and costs of preventing banks from engaging in proprietary trading or investing in hedge and private equity funds. As a structural law, the Volcker Rule will be significantly constitutive of the very capital markets it regulates, making forward-looking predictions about how those markets will function under the rule inevitably speculative. Professor Jeffrey Gordon has argued this point more generally about financial regulations,[87] but regardless of whether it applies to all or even most financial regulations, it clearly applies to structural laws such as the Volcker Rule.

In addition to this core problem posed by structural laws, any effort to quantify those benefits runs straight up against numerous other difficulties. Any complete quantified CBA/FR of the Volcker Rule would require estimates of the costs and frequency of financial crises, which in turn would require macroeconomic modeling, subjective data selection and the prediction of policy responses to any emergent crisis. The difficulties with the Volcker Rule are compounded beyond those facing any regulation designed to reduce the odds and effects of a financial crisis, however, for two reasons. First, the rule has additional, separate benefits, such as the mitigation and reduction of conflicts of interest, which can only be quantified by relying on causal inferences with low-powered tools about complex institutional arrangements.

Secondly, and perhaps more important, it remains unclear how, if at all, the Volcker Rule will in fact reduce the risk or cost of financial crises. For reasons sketched in Sections 1 and 2 above, the rule's proponents (including Volcker himself) strongly believe that it will, by decreasing the role of speculation within banks, changing their organizational culture, and by limiting the ability of banks to attract and retain individuals with a risk-taking temperament.[88] But those judgments rest on personal experience and direct observation, not on publicly available historical data, nor is there any mechanical relationship between an activity (proprietary trading) and failure, as there may be with other elements of banking that are regulated, such as capital levels. Ironically, perhaps, the primary category of benefits (reduced systemic crisis risk from less speculation by banks) is inherently speculative, as with any novel structural law.

Quantifying the aggregate costs of the rule would be equally difficult. While the OCC quantified a subset of costs, it did not quantify the costs that are likely to be largest – especially the potential costs of lower liquidity. As the OCC noted, it may be possible to develop guesstimates for those costs: there are research papers estimating the cost of reduced liquidity for specific categories of assets.[89] But, as the OCC also noted, any estimates produced by relating predicted reductions in liquidity to this sparse research literature would be "difficult."[90] Among other things, a full set of cost estimates would require predicting the impact of the rule on liquidity across a range of financial markets, including anticipating entry by institutions not subject to the rule – institutions that could be expected to take advantage of any competitive opportunities opened up by the exit of banks subject to the rule. Those estimates would then have to be linked to estimates of the impact on the cost of capital from any expected reduction in the liquidity of one channel for capital raising, again taking into account possible substitution effects from other channels. Then, finally, the effects on output of any estimated capital cost increase would have to be quantified, using a macroeconomic model.

In sum, the result of any CBA of the Volcker Rule would be complex, difficult, constrained by limited data, highly contestable and sensitive to modeling assumptions. Any bottom-line "quantification" emerging from such an analysis would consist of no more than guesstimates that likely would straddle net benefits of zero by a large amount in either direction. An administrative law mandating that the banking agencies achieve the impossible – to reliably and precisely quantify the costs and benefits of the Volcker Rule – is by definition impractical, and would have more negative effects (delaying otherwise defensible, and in this case, legally mandated, regulation) without any clear offsetting benefit. Mandatory quantified CBA of the Volcker Rule flunks its own cost-benefit test.

86 See Office of the Comptroller of the Currency (n 77) 1, 23.

87 Gordon, J. N., 2015, "The empty call for cost-benefit analysis in financial regulation," 43 Journal of Legal Studies, S 351.

88 Financial News, 2012, "Paul Volcker fights for Volcker Rule," 14 February; Moyers, B., 2012, "Paul Volcker on the Volcker Rule," 5 April, http://billmoyers.com/segment/paul-volcker-on-the-volcker-rule/ accessed 29 August 2015.

89 Office of the Comptroller of the Currency (n 77) 17 (citing Hasbrouck, J., 2009, "Trading costs and returns for U.S. equities: estimating effective costs from daily data," 64 Journal of Finance 1445, 1445-77).

90 Ibid 1, 23.

CONCLUSION

In this article, the Volcker Rule has been analyzed as a "structural law," a type of law that aims to shape behavior not only or primarily through direct commands but indirectly, by shaping and channeling the institutions of banking and in so doing change their cultures. At once more ambitious and more powerful than ordinary laws, structural laws work indirectly and sweepingly. Such laws, the article argues, have a long pedigree in Anglo-American legal history. At the same time, modern structural laws require more delegation to regulatory agencies, and so run up against another set of structural laws – those comprising the bulk of administrative law.

One component of administrative law over the last several decades has increasingly been legal commands that agencies engage in CBA, and ideally quantification of the costs and benefits of important new regulations. The difficulty with such an administrative law approach, however, is that it requires agencies to do the impossible, in the case of new structural laws such as the Volcker Rule: to anticipate, in advance of relevant data, the private market behavior in response to novel structural constraints on banking activity, such as that reflected in the Volcker Rule. In other words, if administrative law's goals are to be achieved in the context of major banking laws such as the Volcker Rule, they must find some other way to do so than through requirements of CBA. Perhaps inter-agency dialogue will help, perhaps laws and budgetary tools designed to encourage regulatory experiments will help, perhaps agencies can be pressed to include sunsets and other means to evaluate and adapt regulations over time.[91] But for novel structural laws such as the Volcker Rule, CBA is not a promising way forward.

91 Coates, J. C., 2015, "Towards better cost-benefit analysis: an essay on regulatory management," 78 Law and Contemporary Problems 1-23 (making these and other suggestions for improving regulation through better CBA by financial regulatory agencies).

A Historical Perspective on the Different Origins of U.S. Financial Market Regulators[1]

Susan M. Phillips – Dean Emeritus, George Washington University School of Business[2]
Bluford H. Putnam – Chief Economist, CME Group[3]

Abstract
There are varying objectives and cultural differences among the major regulators of derivative markets in the U.S. This article seeks to shed some light on the sources of differing missions among the Federal Reserve Board (Fed), Securities and Exchange Commission (SEC), and the Commodity Futures Trading Commission (CFTC) by exploring their origins. While the CFTC was not created until 1974, it has its origins in the Cotton Futures Act of 1914/16, and its focus was on the integrity of markets. The SEC was created by the Securities Exchange Act of 1934 in response to the Great Depression with a focus on investor protections. After a series of banking panics in the late 1800s and early 1900s, the Federal Reserve Act of 1913 established the Fed to promote banking system stability. After the Great Depression and WWII, the Fed's objectives were broadened to include a focus on managing the economy to achieve full employment and price stability. Our perspective is that to understand the regulatory ecosystem in the U.S., one has to appreciate the implications of the different priorities of each regulator and, critically, whether its original focus was on market integrity, investor protections, or systemic risk.

1 Disclaimer: All examples in this article are hypothetical interpretations of situations and are used for explanation purposes only. The views in this article reflect solely those of the authors and not necessarily those of CME Group or its affiliated institutions. This paper and the information herein should not be considered investment advice or the results of actual market experience. The authors would like to thank Professor Sykes Wilford from the School of Business at The Citadel in Charleston, SC, for his assistance with this article.

2 Susan M. Phillips, retired, has considerable regulatory experience, including SEC Economic Fellow, 1976-1978, CFTC Commissioner, then Chairman, 1981-1987, and Governor, Federal Reserve Board, 1991-1998.

3 Bluford H. Putnam, is a former central bank economist, who started his career at the Federal Reserve Bank of New York (1976-1977).

INTRODUCTION

The regulation of derivative markets in the U.S. focuses on protecting individual investors from fraud and criminal activity, assuring the integrity of markets and safeguarding the economy against systematic risk emanating from the financial sector. These three critical objectives are not embedded in one regulatory authority, but are instead distributed across several institutions with very different origins and priorities, based in no small way on the historical context that led to their creation. That is, to appreciate the sources of different regulatory philosophies among the major regulators of derivative markets in the U.S., one has to examine why each institution was brought into existence and how that shaped its specific regulatory style and priorities.

Toward that end, this paper first succinctly summarizes the origins, mission, and policy focus of the three major institutions regulating financial derivatives in the U.S.; namely, the Commodity Futures Trading Commission (CFTC), the Securities and Exchange Commission (SEC), and the Federal Reserve Board (Fed). We also include an analysis of how the Dodd–Frank Wall Street Reform and Consumer Protection Act of 2010 (Dodd-Frank Act) added powers and responsibilities to each of these regulatory institutions.

With the historical context as our foundation, we provide a set of observations about how their different birth stories and missions have affected the regulatory ecosystem in the U.S. Our analysis is intended to shed light on why the CFTC, SEC, and the Fed may approach apparently similar challenges with different philosophical approaches.[4] In addition, we provide critical perspectives on such issues as transparency, self-regulatory organizations, too big to fail (TBTF), capital adequacy, and the unintended consequences of macro-prudential regulation on the effectiveness of monetary policy.

HISTORICAL BACKGROUND

CFTC

Establishment

While not created in its present form until 1974, the CFTC had its origins in the Cotton Futures Act 1914/16.[5] The Cotton Futures Act was specifically focused on the issue of the terms and standards for the physical delivery process when a futures contract is held to maturity. The delivery process is viewed as having the potential for fraud and manipulation, which is why futures and options have long been regulated, first by exchanges, then by governments. In addition, federal pre-emptive regulation has allowed a distinction from state gambling regulations, preventing states from attempting to regulate futures and options exchanges under local gambling laws. Since its establishment in 1974, the CFTC has been given broad authority over named commodities "…and all services, rights and interests, … all other goods and articles except onions and motion picture box office receipts."[6]

Mission

Quoting from the CFTC's official website, "the CFTC's mission is to protect market participants and the public from fraud, manipulation, abusive practices and systemic risk related to derivatives – both futures and options – and to foster transparent, competitive and financially sound markets."[7]

Market Integrity

The actions, rules, and regulations initiated by the CFTC have a clear focus on how markets work and ensuring the integrity of the market place. Trading must be on exchanges (designated contract markets), although there have been some exceptions granted since 2001. Market professionals must be registered. Margin requirements are enforced.[8] Capital requirements are set to assure that exchange traded contracts will be honored. There are a variety of anti-manipulation initiatives, including speculative limits, delivery oversight, and daily settlement. Large trader reports are provided to exchanges and the CFTC to assist in market regulation, but not generally publicly disclosed except in aggregated form. The role of central clearing houses is primary to how futures and options exchanges function, and the CFTC has relied in part on clearinghouse oversight as well as embraced self-regulatory organizations (SROs) such as the National Futures Association (NFA) and the exchanges themselves.

Dodd-Frank

The Dodd-Frank Act gave the CFTC more authority to supervise and regulate over-the-counter (OTC) markets in swap transactions, and also in particular, swap dealers. Among many other things, clearing

4 Kindly note that we are not covering all aspects of the three agencies' responsibilities; rather, we will focus on market regulation responsibilities

5 The Cotton Futures Act was originally passed in 1914, but it was deemed by the courts as revenue raising legislation, which constitutionally must originate in the U.S. House of Representatives. Since the 1914 version originated in the Senate, it was declared void, and the 1916 version was then passed in the proper sequence from House to Senate.

6 See 7 U.S.C. § 13-1; CEA § 9-1

7 CFTC, www.cftc.gov/About/MissionResponsibilities/index.htm

8 Margin requirements are established by CFTC regulations and delegated to exchanges with oversight by the CFTC. By law, the Fed was given powers related to margin requirements, however, it chose to delegate its role in setting margins to the CFTC and the SEC

and trade execution for standardized derivative products, including certain swap agreements, were mandated to move to exchanges or swap execution facilities and be centrally settled in clearing houses. In keeping with the CFTC's tradition of focusing on the integrity of markets, the additional powers given to the CFTC in the Dodd-Frank legislation were generally aimed at strengthening the infrastructure of derivative markets to ensure their integrity.

SEC

Establishment

The SEC was created by the Securities Exchange Act of 1934 (as a result of the stock market crash of 1929 that preceded the Great Depression) and charged with enforcing the Securities Act of 1933. The focus was aimed directly at providing stronger investor protections. In the years and decades that followed, the SEC was also given responsibility for enforcing a number of other investor protection acts passed by the U.S. Congress, including, among others, the Trust Indenture Act of 1939, the Investment Company Act of 1940, the Investment Advisers Act of 1940, and the Sarbanes–Oxley Act of 2002.

Mission

"The mission of the US Securities and Exchange Commission is to protect investors, maintain fair, orderly and efficient markets, and facilitate capital formation."[9]

Investor protections

Major characteristics of the SEC's approach to market regulation include transparency and disclosure (e.g., financial data by firms, stock ownership by management, market transaction data, etc.). Insider trading rules play an important role to level the trading field so that insiders cannot benefit by having an informational advantage over the general public. As with the CFTC, there are requirements for the registration of securities market professionals – brokers and dealers.

Unlike the CFTC, which views exchange-traded derivative markets as focused on risk management and is neutral on the direction of markets, the SEC has specific restrictions on short selling of stocks. Remember that part of the SEC's mission is to encourage capital formation, and it has accepted the view that in certain circumstances short-selling may cause harm to the capital formation process.

As with the CFTC, the SEC has embraced reliance on SROs to implement and enforce regulations (e.g., FINRA – Financial Institutions Regulatory Authority, as well as the securities exchanges).

Dodd-Frank

The Dodd-Frank Act gave the SEC more powers related to robust record-keeping and real-time reporting regimes including audit trails. Provisions of the Act also focused on giving the SEC anti-disruptive trading initiatives and increased securities exchange oversight to be implemented as a result of the "Flash Crash" in May of 2010 and the Wall Street bailouts associated with the financial panic of 2008 and the subsequent Great Recession. In keeping with the SEC's focus on investor protections, the Act included new governance, capital and reporting requirements for individual firms. The role of the credit rating agencies in the lead-up to the 2008 financial crisis came under severe criticism and the SEC gained powers in this realm as well to better protect investors.

FED

Establishment

After a series of banking panics in the late 1800s and early 1900s, the Federal Reserve Act of 1913 established the Fed to promote banking system stability.

Mission

The Federal Reserve Act of 1913 was all about the safety and soundness of the banking and financial system (i.e., systematic risk) and created an institution with powers of lender of last resort. The dual objectives of encouraging full employment and maintaining price stability were added after the Great Depression, an episode in which, by many counts and assessments, the Fed failed to use its lender-of-last-resort powers to limit the damage from the stock market crash of 1929 and potentially to avoid the downward spiral into deflation and the Great Depression.[10]

Stabilizing the banking system, then managing the economy

The Fed is the central bank of the U.S.[11] It was founded by Congress in 1913 to provide the nation with a safer, more flexible, and more stable monetary and financial system. Over the years, its role

9 SEC, www.sec.gov/About/WhatWeDo.shtml

10 For example, see Bernanke, B. S., 1983, "Non-monetary effects of the financial crisis in the propagation of the Great Depression," NBER Working Paper No. 1054. Also, Bernanke, B. S., 2000, Essays on the great depression, Princeton University Press

11 The U.S. had been without a central bank since 1836 when the charter of the U.S. Bank was allowed to expire. In 1832, Congress passed an act to extend the charter of the U.S. Bank beyond its expiration date, and President Jackson vetoed the charter extension. The role of the central bank became a major issue in the 1832 Presidential election, and when President Andrew Jackson won a second term, the issue was settled and the charter was allowed to expire.

in banking and the economy has expanded.[12] Today, the Fed's duties fall into several general areas: (1) implementing the nation's monetary policy by influencing the monetary and credit conditions in the economy in pursuit of maximum employment, stable prices, and moderate long-term rates; (2) supervising and regulating banking institutions to ensure the safety and soundness of the nation's banking and financial system to contain systemic risk that may arise in financial markets; (3) providing financial services to depository institutions, the U.S. government, and foreign official institutions, including playing a major role in operating the nation's payments system. Until the financial panic of 2008, the primary tools of the Fed included bank reserve requirements, discount window (elastic currency, lender of last resort), and open market operations (T-bills). With the advent of the financial panic of 2008 and the Great Recession, the Fed expanded its toolkit, expanding its balance sheet and engaging in transactions involving a wider range of securities and derivatives (e.g., increased direct purchases of U.S. Treasury securities and mortgage backed securities, as well as creating and lending to special purpose vehicles holding a variety of credit and derivative exposures).

Dodd-Frank

The Dodd-Frank Act gave the Fed expanded authority over the financial system. New powers included the ability to regulate compensation practices of financial institutions. The Fed was also responsible for enforcing resolution regimes for systemically important financial institutions (SIFIs) in the event they had to be wound down. There was an expanded emphasis on a much broader definition of financial firms, well beyond banks, with emphasis on governance, risk management, capital and liquidity. In effect, the Fed was empowered to address regulatory and systematic risk challenges in the "shadow banking system." The Fed also became the central regulatory institution for international coordination of financial system supervision, which includes the negotiations for reciprocal recognition of comparable institutions, such as exchanges or clearing houses, with foreign governments and regulatory bodies.

OBSERVATIONS ON THE REGULATORY CHALLENGES OF DIFFERENT MISSIONS

How do these three regulatory regimes differ? What affects their ability to work together on regulatory reform or impacts the compliance structures required of regulated financial institutions?

Our perspective is that the different historical contexts and varying focuses of regulation that were incorporated into the creation of each of the major derivative regulatory institutions has shaped their style and approach to market supervision. That is, the CFTC's primary emphasis on market integrity, contrasts with the SEC's central focus on investor protections and the Fed's mission regarding the containment of systematic risks.

Transparency

Take transparency as an example. The SEC puts transparency on a pedestal in attempts to protect investors and level the trading playing field. Mutual funds and asset managers have to report positions quarterly, which are made available publicly by the SEC. By contrast, the CFTC has tended to preserve the confidentiality of positions. The CFTC's commitment-of-traders report gives an aggregated sense of the positioning of large groups of specific types of traders, but there is no ability to back into the positions of any one trading firm. Individual business strategies involving price hedging are kept confidential in the CFTC regulatory structure in contrast to the SEC's requirement to disclose ownership positions in public companies.

The inherent differences between risk management instruments, such as exchange-traded futures and options, compared to capital formation instruments, such as stocks and bonds, underlie the contrasting approaches of the CFTC and SEC and help explain why their philosophical approaches to transparency policies are also different. We also note that the Fed focuses on financial confidentiality, although not nearly to the degree that the transparency issue challenges the different instruments regulated by the CFTC and the SEC.

Market direction

Then, there is the embedded view on market direction. The CFTC, with a focus on risk management tools, is neutral – price protection (hedging) in both directions is actively desired and derivative markets are considered a zero sum game. The SEC has a distinct emphasis on promoting economic growth through capital formation and this is reflected in specific restrictions on short selling. The Fed also seeks to promote economic growth, which can lead to a bias in favor of equity bull markets, although the latter has been occasionally tempered by fears of systematic risk coming from "exuberant" markets.

SROs

There are also significant differences in the approach to financial oversight through the use of SROs. With the CFTC's emphasis on market integrity and SEC's focus on investor protections, both regulators have embraced SROs. By contrast, the Fed's role in the banking system and focus on systematic risk has kept its attention

12 Federal Reserve System, www.FederalReserve.gov/AbouttheFed/Mission.htm

on individual financial institutions. We may be stretching the point here, however we believe that these differences in approaches to SROs may be more related to budgets than to mission and focus.

The Fed has a very different budget structure than either the CFTC or the SEC. While the Fed receives user fees for its financial institution supervision and bank payments system services, in a manner not dissimilar to the fees generated by the CFTC and SEC, the Fed also has a very large net income coming from its asset-liability structure. That is, the Fed has a large portfolio of interest-bearing securities funded by virtue of its powers to issue zero interest currency as well as to set the interest rates it pays on required and excess reserves. As a result, the Fed generates substantial portfolio earnings and is typically able to return a considerable portion of its net interest income to the U.S. Treasury.[13] Thus, while the Fed sends an annual report to Congress every year, unlike the SEC and CFTC, the Fed does not need to get its budget approved, giving it considerably more independence than enjoyed by the SEC and CFTC.

Both the CFTC and the SEC examined their use of SROs after the 2008 financial crisis and passing of the Dodd-Frank Act. The SEC previously viewed SROs as partners but recently has been bringing enforcement actions against them. This raises the question of whether SROs continue to be effective if they are placed in an adversarial position with their primary regulating agency. In addition, as exchanges have gone public, the regulatory authorities have had to assess the unavoidable conflict of interest between the business side of the exchange and its traditional self-regulatory responsibilities. While these conflicts appear manageable, the need to clarify roles is critical.

Trade-offs between containing systematic risk and encouraging market liquidity and efficiency

There are inherent philosophical debates that are becoming more obvious depending on whether the focus is on systematic risk or the efficient functioning of markets. For example, the Volcker Rule, which seeks to limit proprietary trading by certain types of financial institutions, especially banks, is part of an attempt to reduce the risk of failure leading to systematic problems. The unintended side-effect, however, is to reduce the amount of risk capital and trading activity in certain markets, potentially adversely impacting market liquidity and the costs of trading and capital formation for users of the markets.

Also, the Dodd-Frank legislation appears to have made regulatory compliance tasks more complex for financial companies. The SEC and the CFTC both have an interest in the regulation of securities and related derivative products, often with different missions and objectives that are not always easily compatible. For example, index-based contracts trade on futures exchanges, while index-linked exchange traded funds (ETFs) trade on securities exchanges, yet often utilize futures contracts to track their benchmarks. Further, the SEC and CFTC often find themselves with challenging overlapping market concerns with the Fed regarding trading in U.S. Treasury securities and on bank trading practices involving securities and futures contracts.

Capital adequacy and too big to fail (TBTF)

TBTF will be an issue as long as economies of scale exist. Moreover, different approaches to managing the systematic risks of large institutions are likely to create considerable debate, even among the various regulators. For example, to mitigate the systematic risk of the failure of one large institution spreading through the financial network, the Dodd-Frank Act mandated that many OTC swaps now be settled through a central counterparty clearing facility. By mutualizing risk, that is, putting the clearing house in between buyers and sellers, the Act reduced the risk of a domino effect from the bankruptcy of a large institution while making clearing houses more critical to the functioning of the system. This required intermediation may reduce swap participants' contract flexibility while possibly improving liquidity, especially for exiting swap contracts.

In addition, TBTF issues spillover into capital adequacy questions. The Fed has traditionally been a regulator of banks, which are leveraged lending institutions, and capital requirements are a key part of the Fed's supervision and oversight. As the Fed's jurisdiction has expanded to non-bank institutions, with containing systematic risk as the key focus, there has been a tendency to apply bank type rules to institutions that have little in common with banks, such as insurance companies. Moreover, some clearing houses are designated as systematically important institutions for certain purposes, and, thus, the Fed may weigh in on issues impacting clearing house capital requirements, and not necessarily from the same regulatory perspective as the CFTC or SEC. If the various regulatory requirements become too onerous or costly, we may see financial institutions move offshore. Internationally, we also observe that the Bank of England is moving in this direction of using heightened capital requirements for a variety of non-bank institutions in very different types of businesses.

13 Prior to the 2008 financial crisis and the expansion of the Fed's balance sheet through asset purchases (i.e., quantitative easing), the Fed typically returned around U.S.$20 billion per year to the U.S. Treasury from its net earnings. In the 2012-2014 period for example, with a much larger balance sheet, the Fed provided the US Treasury with U.S.$80 to U.S.$90 billion dollars annually from its net earnings.

Impact of macro-prudential regulation on the effectiveness of monetary policy

Indeed, the focus on additional capital charges for the largest banks proposed by the U.S. bank regulators and risk-based capital charges (equity or debt) for Global Systemically Important Banks (GSIFI) proposed by the Basel Committee for Banking Supervision (BCBS) are designed to mitigate the challenges of systematic risk. But a reliance on capital ratios and charges by a central bank can raise new issues with regards to the unintended side-effects related to the interaction of the conduct of monetary policy aimed at managing economic risks and regulatory activities focused on macro-prudential systematic risks. We would broadly define macro-prudential regulation as using supervisory tools to control perceived financial bubbles or asset price movements that are considered by the regulator as undeserved. These types of actions can have the unintended effect of rendering traditional monetary policy considerably less effective.

For example, in the aftermath of the Great Recession of 2008-2009, the Fed, the European Central Bank (ECB), and the Bank of Japan (BoJ) all expressed concerns, to varying degrees, about the potential for deflation. Neither zero short-term interest rates nor massive asset purchases (i.e., quantitative easing or QE) had any observed ability to encourage inflation.[14] One very powerful reason for the inability of extraordinary monetary policy measures to promote an increase in inflation pressures is that the link between the credit creation process and both short-term interest rate policy and the size of the central bank's balance sheet has been severed by more stringent capital controls and macro-prudential regulation.

That is, if a central bank buys the government debt of its country it may put some limited downward pressure on bond yields, as it did in the U.S. during 2012 and early 2013, but it is not clear at all if such actions impact the decision by capital-constrained financial institutions to increase lending. What seems to matter much more for the credit creation process are the expectations of financial institutions about the state of the economy and the perceived risk of extending new loans with a careful eye on capital preservation and capital ratios. On net, in the era of expanded central bank balance sheets, central banks will own a much higher percentage of their country's outstanding government debt while the private sector will own a smaller proportion. One could even see credit agencies viewing this development as a positive factor for their sovereign credit ratings, but central bank asset purchases will not have made any difference in creating inflation. Similarly, zero short-term rates have not ignited the kind of lending boom necessary to fuel inflation pressures, because banks are much more worried about their own profitability and risks. In short, at low rates, the link between central bank policies and credit expansion is very loose if almost non-existent.

We are not arguing against expanded macro-prudential regulation. What we are observing, however, is that one form of regulation designed to mitigate systematic risk may well render other policy tools used for managing the economy less effective. And, there is the plausible scenario that relying more heavily on macro-prudential regulations, such as very large mandated capital ratios, may curtail risk capital allocated to trading activities and potentially make markets used for risk management purposes less liquid and not as efficient. These types of trade-offs are often at the heart of regulatory debates, especially when the focus of the regulators differs. The Fed's focus on economic management and systematic risk, in this sense, places it in a different philosophical position compared to the CFTC's focus on market integrity and efficiency as well as with the SEC's primary emphasis on investor protections.

CONCLUDING THOUGHT ON CULTURE AND ORIGIN OF REGULATORY INSTITUTIONS AND THE NEED TO BE WARY OF UNINTENDED CONSEQUENCES

The types of unintended consequences from the multi-institutional regulatory structure in the U.S. seem bound to become more challenging as regulators seek to achieve different objectives, ranging from improving market integrity, to enhancing investor protections, to containing systematic risk. In essence, we are brought back to two important strands of market structure and regulatory theory – namely, (1) the causes of market failure and (2) the public choice theories of why any political system creates the regulatory system that it does. Each market failure, whether the banking panics of the 1800s, or the old-style delivery squeezes in futures markets, or the stock market crash of 1929 and the Great Depression, tend to give way to new legislation and new regulatory powers specific to the last crisis or market failure. Viewed in this historical light, it is not so surprising that the U.S. has one of the more complex financial regulatory systems leading to regulatory institutions approaching similar market challenges from different philosophical approaches based on their birth stories and missions.

14 Putnam, B. H., 2013, "Essential concepts necessary to consider when evaluating the efficacy of quantitative easing," Review of Financial Economics 22, 1-7.

Investment

Knowledge Management in Asset Management

Private Equity Capital Commitments: An Options-Theoretic Risk Management Approach

Credit Risk Decomposition for Asset Allocation

Time to Rethink the "Sophisticated Investor"

Fund Transfer Pricing for Bank Deposits: The Case of Products with Undefined Maturity

Delegated Portfolio Management, Benchmarking, and the Effects on Financial Markets

Knowledge Management in Asset Management

Eduard v. Gelderen – Chief Investment Officer, APG Groep N.V.
Ashby Monk – Executive Director, Global Projects Center, Stanford University[1]

Abstract
The idea that superior knowledge is required to drive financial outperformance runs counter to some of the most pervasive theoretical frameworks used by investors today. The Efficient Market Hypothesis and the Capital Asset Pricing Model, for example, posit that capital markets are efficient and that no consistent outperformance can be generated without increasing risk. Active asset managers, however, argue differently and claim that skills and knowledge are critical for capturing excess returns. We agree. In fact, in this paper we argue that knowledge assets and the use of superior knowledge are crucial to the success of all asset managers and, in particular, active managers. And yet, despite its clear importance, very little is known about knowledge management in asset management. This article thus seeks to remedy this by offering insight into the role that knowledge plays in the investment process and, more specifically, into the adoption of knowledge management by asset managers. The paper concludes with a roadmap that offers a way for investors to become knowledge and asset managers.

1 This paper was supported by the Institutional Investor Research Club of Stanford University's Global Projects Center. We would like to acknowledge the considerable support and insight of Prof. Gordon L. Clark of Oxford University. We would also like to thank Adam Dixon, Joop Huij, Ray Levitt, Caroline Nowacki, Dane Rook, Rajiv Sharma and Allan Wain for comments on a prior draft. None of the above is responsible for any errors or omissions.

INTRODUCTION

Asset management (AM) refers to the professional administration and investment of financial assets to achieve specified investment goals and objectives. On the surface, asset managers have a simple and attractive business: they take an initial stock of money – what we call financial capital – and put it to work through the application of human capital (i.e., people), market intelligence (i.e., research, technology, and networks), and governance (i.e., policies, processes, and procedures). When these three inputs are combined effectively with an initial stock of capital, asset managers can generate attractive investment returns for clients and, in turn, revenues for their business and employees. Generally speaking, then, a successful investment organization is one that is adept at employing talented individuals in operating environments constrained by policies, processes and procedures in order to identify and then exploit informational advantages in a timely manner. This may seem to be a simple formula for success, but it raises important and complex questions. For example, what are the factors that allow for investment organizations – be they for-profit asset managers, such as hedge funds, or beneficial investment organizations, such as endowments or pensions – to develop and mobilize the inputs listed above? And, in turn, once the inputs are mobilized, can these investors substantiate their value? It is in answering these questions that the business of AM becomes rather complicated. In our opinion, the creation, maintenance and exploitation of "knowledge" are critical to the success of any investment organization.

Since Coase's (1937) paper on "the nature of the firm," many theories have been developed to explain the core essence of firms and the large diversity among them. According to Kraaijenbrink and Spender (2011), at least twenty "theories of the firm" have been put forward, originating from different disciplinary perspectives, such as economic, organizational, and behavioral theories. These can be grouped into four buckets: 1) the firm as a bundle of assets; 2) the firm as a bundle of people; 3) the firm as a production system; and 4) the firm as an interest-alignment system. In order to differentiate further among the prevalent theories of the firm, we can also distinguish the different ways in which firms create value. Of particular interest to our work, Penrose (1959) argued that the ability to bring different intellectual resources together, as part of the production system, is the main driver behind a firm's success. This early research resulted in the knowledge-based view of the firm, but it was not until Nonaka (1991) that the practical implications of this theory were recognized. Specifically, it became accepted that new knowledge, i.e., value, could be created by means of the continuous interaction between explicit and tacit knowledge. In this respect O'Leary (2002) talked about knowledge management as the organizational efforts to: 1) capture knowledge, 2) convert personal knowledge to group-available knowledge, 3) connect people to people, people to knowledge, knowledge to people, and knowledge to knowledge, and 4) measure that knowledge to facilitate management of resources and to help understand its evolution. This is true for firms and investment organizations. For example, investment firms with good governance and an optimal set-up of rules and procedures are able to outperform [Moussavou (2006); Clark and Urwin (2008); Clark and Monk (2013; forthcoming)].[2]

As Nonaka and Takeuchi (1995) define it, knowledge is about forming beliefs and making commitments; it is about putting information and data into action. As this implies, knowledge also goes to the heart of investment decision-making. And, if we assume that active management is a zero-sum game (or at least close to it), superior knowledge would seem to be the only way to achieve excess investment returns. While this may seem an obvious observation, it is worth noting that this view actually runs counter to some of the dominant frameworks used by investors today [see Clark (2014)]. For example, the Efficient Market Hypothesis [Fama (1965)] and the Capital Asset Pricing Model [Treynor (1961); Sharpe (1964); Lintner (1965)] are based on the premise that capital markets are efficient and that no asset manager has superior knowledge over the broader market, believing that all possible information is reflected in current market prices and excess returns are simply a function of the level of risk taken.[3] But, as you might expect, the community of active asset managers disagrees with these mainstream views, arguing that informational advantages do exist and that opportunities for generating excess returns can be identified in the market.[4] This is a view that also seems to be in line with recent empirical research on factor investing. For example, Harvey et al. (2014) identified more than 300 factors that affect equity returns in empirical literature. However, gathering and leveraging those factors in the context of trading requires developing formal policies for knowledge management. More general research on organizational behavior also shows that all organizations, independent of industry, get value from knowledge management and that knowledge carries

2 Investment organizations with high employee ownership and low turnover underpin investment success [Finstad (2005)]. Even organizational size has been directly linked to investment performance [Beckers and Vaughan (2001); Pozen and Hamacher (2011)].

3 Ross' (1976) Arbitrage Pricing Theory made these (pricing) models more profound by allowing the use of multiple risk factors rather than a single market factor. Several macroeconomic factors, as well as style factors, have been suggested in this respect [see Ang (2014)].

4 For example, Goldman Sachs Asset Management stated in one of their Perspectives: "There are many reasons to believe active portfolio management can effectively transform active risk into active returns. These are well documented in investment literature and include time-varying risk premiums, the tendency of investors to underreact and over-react to different types of information, the existence of investors with motives other than pure risk/return optimization, and a variety of frictions and pockets of illiquidity." [Goldman Sachs Asset Management (2005)]

as much value as financial or even human capital [Grant (1996); Spender (1996)]. In short, the way an organization is structured will inevitably affect its ability to create, maintain, and use knowledge – and it is in the context of the organization's design that knowledge ultimately drives performance.

Given the importance of superior knowledge in performance, you would be forgiven for assuming that knowledge management (KM) – or how human capital, market intelligence and governance is combined to get to grips with O'Leary's approach – was a top priority of all active asset managers. Oddly, it is not. Most asset managers could not be described as knowledge managers at all. Many do not even use publicly available knowledge effectively [Huij and van Gelderen (2014), often relying on the tacit knowledge of an individual investor who is not willing to share his or her knowledge [Gertler (2002)]. In fact, very little is known about KM in AM. This article seeks to remedy this by providing insights into the adoption of KM by asset managers and, more specifically, to the role that knowledge can (and in certain cases does) play in shaping the investment process.

In order to develop our arguments, we adopt a multi-method approach grounded in proprietary expert surveys and elite interviews (as per Strauss and Corbin (1998); Denzin (1970)]. Specifically, we delivered two surveys to investment professionals – first in the Netherlands and then in the U.S. In addition to the two surveys, we also interviewed a group of 20 asset managers between September 2012 and December 2015. We use these qualitative and quantitative results in order to develop a better understanding of the role that KM is playing, and can play in the future, in AM. The rest of the paper proceeds as follows: section 2 presents a theoretical KM framework related to investment processes. Section 3 discusses in more detail the methodologies used in this research. Section 4 offers a series of findings from the research, while section 5 provides a roadmap for how KM could be better integrated into AM. We conclude that, despite the knowledge intensive nature of the AM industry, many aspects of KM are still left implicit and not dealt with in a structural or strategic manner. A more visionary KM approach could still provide investors with a true competitive edge over peers.

ACTIVE AM IS KM

Leibowitz (2005) describes active AM as encompassing four steps: (1) ascertaining why a market is priced where it is; (2) understanding the basis for any mispricing of opportunities; (3) developing a view of the true market equilibrium; and (4) concluding that this "discernment" will transpire within a relevant time span. Active AM thus demands an ability to identify, explain, and act on market inefficiencies and anomalies. As you can imagine, this demands considerable and often privileged knowledge of markets. As Grinold states (1989, p. 35): "The strongest assumption behind the law [of active management] is that the manager will gauge the value of information accurately and build portfolios that use that information in an optimal way. This requires insight, self-examination, and a skill level in the investment manager that may be rarely achieved, no matter how admirable the goal."

The "law of active management" states that any added value from active AM (which, fittingly, is known as an "information ratio") is calculated by multiplying the managers' skill (the information coefficient) by the breadth of the investment opportunities [Grinhold (1989)]. While the term "breadth" is clearly defined as the number of distinct, independent investment decisions possible over a certain time period, the term "skill" (or information coefficient) is not clarified other than the technical definition that the information coefficient is the correlation between ex-ante and ex-post performance. Common practice is to determine a manager's skill using indirect and statistical methods applied to the manager's historical performance record, despite its doubtful statistical significance [Harvey and Liu (2016)]. The idea behind this approach is that if a manager's skill is the driver of excess returns, then the investment returns should differ from random (market) returns. However, the required number of data points is often lacking. A direct and forward-looking approach would be to link excess return to the collection of specific sets of data and information and the development and mobilization of unique and superior knowledge. Accordingly, we believe skilled active AM is tantamount to KM.

However, this then raises the question of what types of knowledge and skills are required to be a successful active asset manager. Knowledge in the case of AM means a deep understanding of the functioning of capital markets and its value drivers, which is a combination of two important factors: 1) explicit knowledge and 2) tacit knowledge:

- **Explicit knowledge:** is primarily gained by means of formal training. Professional training has been linked with performance by academic research. For example, De Franco and Zhou (2007) looked into the value of the CFA designation by comparing the performance of sell-side analysts with and without the CFA designation. They found that analysts with the CFA designation showed better performance. These results were confirmed by Fang and Wang (2015) with regards to stock picking skills in the Chinese capital markets. These results show that the CFA training is successful in providing market knowledge. Chaudhuri

(2013) also showed that managers with a high number of PhDs also provide superior performance. The explanation is found in the typical training PhDs receive in the analysis of complex problems. This result comes closer to our definition of superior knowledge; PhDs are trained to ask for the right information.

- **Tacit knowledge:** is earned over time through experience. Again, research shows that this experiential knowledge is also linked with performance. For example, Greenwood and Nagel (2006, 2008) found a clear positive difference in performance in favor of seasoned investors. Although the younger investors had gone through professional training, "inexperienced investors form their beliefs about future price changes by extrapolating past price trends from limited data" (p. 16). As a result, younger investors missed sharp changes in market sentiment and more frequently ended up in lossmaking positions.

This combination of training and experience forms the basis of knowledge and, ultimately, skill. And skill, in Grinold's statement at least, is the capacity to build optimal portfolios to exploit market inefficiencies and anomalies. Put differently, a skillful asset manager maintains and creates superior knowledge and knows how to apply that knowledge effectively.

Superior knowledge may, however, become obsolete over time. After all, market participants quickly become aware of how pioneers exploit market inefficiencies and anomalies and copy their approach. The result is that these investment opportunities are arbitraged away very quickly and no longer offer profitable strategies for active management [Ineichen (2004)]. Lo's Adaptive Market Hypothesis (2004) touched on this as well by postulating that the drivers of markets change over time and new inefficiencies and anomalies inevitably emerge. Consequently, a skillful manager is also typified by the ability to act on changing market conditions by creating new superior knowledge and abandoning obsolete ones. The true impact of skills on investment performance, it turns out, is largely dependent on an organization's ability to foster enduring and valuable knowledge, and to adjust investment strategies accordingly.

RESEARCH METHODS

Over the course of this three-year research project, we have conducted two surveys and interviewed dozens of investment professionals. We believe this an appropriate methodological approach, as this paper does not seek to establish causality or even correlation. Rather, this paper seeks to "map out" the current KM landscape in AM and makes some rudimentary assessments and predictions about its future prospects. Expert surveys and elite interviews with decision-makers provided us with a detailed understanding of the current – and indeed potential – role of KM in AM.

In terms of surveys, both the American and Dutch surveys were constructed as "expert" opinion surveys. Expert surveys like these are important tools in social science research where quantitative, primary data is missing, as was the case for KM in AM [Castles and Mair (1984)]. The first survey was delivered by the Dutch Investment Professionals Association (VBA), which helped to coordinate an online survey in 2012 that had 74 expert respondents. The survey was written in Dutch and consisted of twenty multiple-choice questions, of which five related to the profile of the respondent and fifteen to KM. The majority of respondents (54%) held a senior executive position as board member or managing director at asset managers and asset owners. This survey's aim was to gain a general understanding of investment professionals' views on: (I) the basics of KM's value to an asset manager; (II) the type of knowledge that is related to investment performance; and (III) the ways in which investment organizations can operationalize KM.

Based on the results and experience with the pilot survey in the Netherlands, we conducted another survey on the same topic that targeted U.S. investment professionals. This survey consisted of 19 multiple-choice questions, of which four were related to each respondent's profile. The three focus areas remained the same: (A) the added value of KM, (B) the type of knowledge related to investment performance, and (C) points of particular interest in KM. The advantage of this survey over the VBA survey was that it was possible to drill down into the responses according to specific respondent-groups. Moreover, some questions were adjusted to gain additional insights. Pension & Investments distributed this survey electronically to their subscribers. The survey remained open for three weeks from August 19 till September 9, 2013. A total of 243 responses were received.

Next to the two surveys, a group of 20 asset managers were interviewed during the period September 2012 to December 2015. This fieldwork was used to develop a set of detailed KM case studies [as per Helper (2000); Feldstein (2000); Aberbach and Rockman (2002)]. The organizations in this fieldwork included: JP Morgan IM, State Street Global Advisors, Blackrock, AXA IM, Robeco, GMO, Bridgewater, Templeton, Pimco, Lombard Odier, Blenheim, MAN Group, Blackstone AM, KKR, Neuberger Berman, Goldman Sachs Asset Management, Kepos Capital, PDT Partners, Stanford Management Company and AQR. This group of 20 was selected to ensure a representation of a variety of differing asset management business models. The interviews took place face-to-face with senior executives and were often followed up with an email exchange for further clarification and additional questions. Although the names

of the asset managers are noted above, these organizations will receive anonymity for the remainder of the paper. In securing privileged access to these organizations, we agreed to respect the social science guidelines concerning confidentiality and anonymity of respondents [in line with the approach of Clark and Urwin (2008)].

In summary, over the past three years we have sought to investigate KM and AM in a variety of ways. The key research findings from this work are synthesized below. Details of the survey results are provided in the appendix.

KEY RESEARCH FINDINGS

At a high level, most of our respondents saw knowledge as overwhelmingly positive and a beneficial asset to an investment organization. However, most of our respondents also lacked a deep understanding of KM and identified many barriers hindering its implementation. In what follows, we provide the key insights from the research project to date:

- **Appreciation and (un)familiarity:** in our surveys and interviews, we defined KM as the explicit and systematic management of knowledge – and its associated processes of creation, organization, diffusion, use, and exploitation – in pursuit of business objectives. We sought to register a distinct difference between data, information and knowledge within our respondent groups and focus their thinking around the action of using knowledge to make investment decisions. Despite that, the findings of our research painted a picture of an AM industry largely indifferent to KM. The survey respondents, for example, suggested that a majority of the industry was only vaguely familiar with the concept of KM. Indeed, few organizations in our research had a clear definition for what KM was, let alone tracked the benefits of KM activities. Many investors were also confused about what KM was and how it could be applied within their organizations to create value. For example, a significant number of respondents pointed to "data and information" as the primary focus of KM, which, again, is misguided. Respondents also failed to recognize the gap between the types of data and information they received and the type of information they required to implement successful investment strategies, let alone to create new knowledge. Only the hedge funds in our research emphasized that having access to unparalleled data and knowing how to apply information was at the core of their business. Interestingly, KM was so poorly understood among our respondents that even those asset managers with clear KM strategies in place did not actually recognize them as such; it was often framed as just "good organizational practice." Key takeaway: Our respondents – from surveys and interviews – proclaimed to appreciate KM and even noted its important role in superior investment results. This appreciation, however, rarely translated into pro-active KM policies, let alone KM resources being allocated deliberately.
- **Significance and relevance:** among those investors that actually did value knowledge in our research, the value of KM was perceived very differently depending on the organization. For example, several interviewed asset managers expressed the importance of knowledge in their organizations, even noting that knowledge was part of their competitive edge and that this edge would grow more important over time. However, these same organizations differed considerably in the value they assigned to explicit and tacit knowledge. The quant-oriented asset managers did not believe in the value of tacit knowledge at all, as their strategies were often fully coded and made accessible to the whole organization. Other asset managers expressed that their star-performers have specific traits; for example, they are quicker to act, are "street-savvy" and know how to draw connections between rare events and asset pricing. Additionally, consensus was that academic research played an important role in the industry, and several asset managers in our research had even established intensive working relationships with academics. In spite of this, there was considerable ambiguity with regards to the value-add of academic research, especially when it is already published. It was for this reason that several of the interviewed asset managers fiercely protected their proprietary research. Yet, others claimed that publishing research was part of their business model to support the industry's thinking, but that the operationalization of academic research often failed. In addition, a large majority of survey respondents felt that knowledge was context specific; that it would be very hard to generalize knowledge from setting to setting or even organization to organization. Several interviewed asset managers also pointed out that successful portfolio managers often failed when they moved companies. An explanation for this observation could well be that the skills of the portfolio manager are no longer a match with the available data and information in the new environment. Take away: even within investment organizations that have a strong appreciation for KM, the value of KM is often perceived differently among them. There was no consensus as to the kinds of knowledge that were particularly valuable, albeit tacit knowledge was more directly linked to excess returns. Nonaka's approach that knowledge is created by means of a continuous cycle between explicit and tacit knowledge is absent. Nor was there a consensus on the drivers of KM's value – for people or organizations – which suggests that even among these leaders there was room for a more structured understanding of KM.

- **Measurement and calibration:** although it may be difficult to measure the value of knowledge in monetary terms, we found that measuring the knowledge ecosystem was critical to the success of KM in AM. Indeed, to ensure proper resourcing and structuring of KM operations first required that the organizations track and communicate the benefits of KM by means of key performance indicators (KPIs). This, in turn, helped the organizations develop internal legitimacy for a KM culture and dedicated KM resources. Depending on the degree of complexity, transparency, profitability, and costs involved, technology was highlighted by many funds in our research as critical to evaluating and delivering KM value to the investment professionals (see below). Similarly, KM technology platforms often provided a venue to challenge existing knowledge, which was something our respondents flagged as critical. Indeed, it was noted that there is no place for complacency in KM, and the possession of superior knowledge should be challenged regularly. Yet, many asset managers are in an early stage and costs precede unfamiliar (see first key insight) benefits. Take away: developing a coherent and well-designed KM organization can be costly. Justifying this cost – to leadership and indeed the board – demands that KPIs be developed that allow for the assessment of KM policies. In addition, these KPIs also help with the assessment of the on-going value of existing knowledge.
- **Technology and infrastructure:** effective, transparent and quantifiable KM programs and policies will inevitably require new technologies. For example, large asset managers in our research specifically noted that technology was crucial in realizing operational efficiency gains as well as helping to improve on communication by bridging physical distances. Technology was also shown to facilitate the creation of collective knowledge by means of intranets, libraries and staff directories, among other things. More specialized managers used technology to code their in-depth knowledge, and hedge funds coded and stored almost everything that was codeable and storeable. Still, it also became clear that many asset managers were struggling to get their basic diagnostics in place. Data management (collecting, cleansing, and integrating data) is in place and provides standard descriptive information. This is often restricted to traditional data, such as statistics issued by government bodies, company data, and market data. But data intelligence (filtering, combining, and extracting relationships from data) is often still a challenge, especially when new data sources (big data) come into play.[5] This means that KM is little more than a long-term ambition. Despite the surge in FinTech companies, it became clear that technology companies had also not caught up with the financial industry's fast development and focus on KM. As a result, many AM firms were frustrated by having to rely on a panoply of scattered and legacy technology platforms that could hardly support traditional investment strategies (let alone anything more innovative). In fact, investment teams often relied on their own models and data sources, which lacked in quality, documentation and transferability. Take away: embedding KM into AM organizations will inevitably require technological sophistication to allow for transparency, institutional memory, rapid query, and communication. That being said, while it is common to associate KM with information technology [Ball (2006)], IT is insufficient. Technology must deal with more than data and information; it must also help to store and distribute knowledge, and support knowledge creation. As such, teams of IT specialists may need to work very closely with the investment professionals to make sure that the right data and information is in the systems.
- **Governance and leadership:** a percentage of our respondents seemed to be of the opinion that KM was not a board responsibility. Similarly, few of the respondents saw KM as the right of the Chief Investment Officer (CIO). And yet, consistent with the idea that knowledge provides a competitive edge and should guide investment decision-making, especially in active management, research would suggest that the boards or CIOs should in fact seek oversight and responsibility of KM. Moreover, our respondents noted that creating new tools and processes to collect and pool knowledge was critical to KM. They also noted that for KM to succeed, barriers to knowledge transfers should be dismantled. A lack of incentives (financial and otherwise) was deemed to be the key impediment to overcoming KM logjams. It was thus noted that actively supporting the existence of knowledge assets is also something that should be embedded in compensation schemes. All of these critical elements to the success of KM are the responsibility of boards and the C-suite. Take away: the Board and C-suite should be leading the way in defining the strategic benefits of KM and not treat KM as a by-product of its operating model. Moreover, KM is not a costless exercise; it requires people, process and technology to get right. As such, it will require sufficient resourcing.
- **Culture:** We found throughout our research that organizations must create a culture that supports KM. Indeed, culture works as a catalyst related to corporate goals. We found that this meant, in practice, mixing professionalism, creativity, collaboration, and hard work. Whatever form or shape of asset manager, knowledge capital is perceived as the differentiating factor for an investor's success. Making better investment decisions is also an important common goal; one should feel free enough

5 This finding was confirmed in Citi's "Big data & investment management" (2015) stating: "...for most investment managers these changes in approach are still highly aspirational and there are still several obstacles limiting big data adoption."

to express opinions and ideas and to give and receive criticism. In managing culture, several asset managers in our research used their founders and senior partners to protect the firm's uniqueness and investment philosophy by coaching younger staff. This way the (tacit) knowledge, which is considered the company's competitive edge, is passed on. Transparency and consonance were also identified as important elements of the corporate culture too, as these factors often triggered the right questions and lead to loyalty and low turnover. It also appeared to be easier for the smaller firms and partnerships to create the right culture; the larger firms needed to introduce more formal structures. Moreover, many respondents viewed that individual proprietary knowledge was a source of power within an organizational context and would not want to cede that power. Take away: Human capital and culture are of utmost importance to developing knowledge, which means AM organizations must focus on hiring people with different backgrounds and traits, and prioritize collective knowledge as a core value.

In summary, our research has showed that the large majority of asset managers have not adopted KM practices, and most viewed it as a subset of IT rather than a strategic lever to guide decision-making. Worse still, neither the boards nor the C-suite have prioritized KM efforts, still relying instead on their star-performers. The lack of understanding of technological developments by boards and c-suites only reinforces the underinvestment in KM. This helps to explain why knowledge transfers are often made more difficult due to organizational constraints. In order to improve KM practices, asset managers recognize a need to reorganize their operations. They pointed towards new technologies and new incentives that could help investment organizations mobilize knowledge. They also recognized the importance of people, culture, and organizational design. These findings are far from earth shattering as they touch on the three drivers behind an asset manager's business model mentioned in the introduction: human capital, market intelligence, and governance. But the crucial point of these findings is that knowledge has not been appreciated as the factor that binds these three drivers together. In the section that follows, we use our research findings to provide an initial "KM roadmap" for those investment organizations that would like to become better stewards of knowledge.

THE "KM ROADMAP" FOR AM

Knowledge is about converting information into action. Superior knowledge refers to the understanding of how to successfully apply the appropriate information through skill and process. For knowledge to provide an investment organization value, it has to be accessible. As Javernick-Will and Levitt (2010) remind us, most organizations do not know what they do know let alone what they do not know, which means they require structured ways of learning and sharing. And, as we found in our research, this is particularly true in AM. In this section, then, we build on the findings from our research above and offer an initial KM roadmap that could help AM organizations capture the value of knowledge. Readers should note that this roadmap takes the strategic goals, market positioning, and governance of the AM firm as a given and focuses entirely on the investment process.[6]

Beliefs: The CIO's first "knowledge" task is to come up with a set of investment beliefs, which provide guidance to the type of investment strategies and styles pursued. These investment beliefs are firmly held opinions, but often lack proof. Still, theoretical groundings must be in place, which must be well documented and made available to the whole organization. The investment beliefs should be made part of the investment culture of the firm, i.e., the firm's pride of ownership rather than a proclamation from the top. Every employee should feel accountable for these investment beliefs.

Investment strategy: investment beliefs are often stated in general terms, but need to be translated into clear investment strategies that add value. This is typically where the CIO relies on the investment experts per asset class as a source of in-depth knowledge of market structure, market dynamics, and instruments. Investment strategies need to be clearly substantiated by means of in-depth empirical research and regularly tested on their merits. In that respect, both supporting and falsifying evidence should be assessed. The next question is whether the market offers enough investment opportunities to add value. To answer this question, Grinhold's "breadth" component in the law of active management provides a useful point of action. As was explained, the breadth of the market implies the potential for active investment opportunities. It is important to make this assessment as explicit as possible in order to test the true merits of the proposed investment strategies, but also to assess the alpha capacity. For example, a distinct alpha source pursued by many asset managers leads to a crowded market, which limits the alpha potential. Based on KM consideration, a decision must be made about whether it still makes sense to pursue an investment strategy related to that alpha source.

6 KM can be applied to all aspects of the AM business and is not constrained to the investment process. Marketing, product development, account management, and operations all benefit from a strategic KM approach.

Skill alignment: given the investment beliefs and the investment strategies, the CIO should decide what "skill" component (possession and use of superior knowledge) in Grinhold's equation is required. Formal training, competences, traits, and experience should match the investment strategies and styles the asset manager wants to pursue. For example, a fundamental analyst is not very likely to exploit complex derivatives and arbitrage opportunities or see use for high frequency trading. Likewise, a quant portfolio manager is less likely to be involved in a focused strategy with a lot of engagement with the companies in the portfolio. But likely important is the required diversity to be able to change quickly to changing market conditions and the level of experience to implement investment strategies successfully. To truly get a grip on the available skills, an asset manager should start measuring the skills of its own investment people. Skill, or information coefficient, was defined as the correlation between ex-ante and ex-post performance. To put this differently: how many times is the investment manager right? A methodical analysis needs to be put in place to measure the information coefficient. This requires that much more detail about trades and holdings in the investment portfolio be registered. Not only does this lead to an overall number indicating the level of skill,[7] but it also provides information on the specific strengths and weaknesses of the investment manager.

Data and technology: different investment strategies and styles go hand-in-hand with specific datasets and information requests. Every mismatch and/or inferior quality of data and information could jeopardize the validity of the chosen investment strategy. Digitalization means that an ever increasing number of datasets become available. Still, more data does not mean more knowledge. Investment skills should include the ability to think about new relationships between data and asset prices. New technology can support finding these new relationships. In addition, decision support tools are not limited to individual trades and portfolio construction, but extend to risk analytics[8] and transaction costs analysis as well. The goal of these decision support tools is to optimize the return potential as much as possible and to avoid any form of performance leakage.

In summary, the challenge related to KM is to find the right match between the investment beliefs and investment strategies on the one side and the required skills, data, and technology on the other. A very first step is to determine KPIs, as illustrated in Table 1.[9] However, there is no rulebook regarding the optimal set up. Asset managers must measure different KPIs over time and analyze their impact on the overall investment performance. In order to facilitate statistical analysis, a KPI indicator and/or sub-indicators can be developed that the CIO could share with the senior investment professionals, who could then relate these findings to changes in the investment environment and their performance. This feedback loop in itself will lead to a better understanding (new knowledge) of the investment process and provide guidance for further improvements and adjustments of that process.[10]

Investment beliefs	• Number of years the investment beliefs have been in place • Number of adjustments to the investment beliefs within a certain period • Number of supporting/falsifying research papers taken into account • Number of internal meetings on investment beliefs • Number of training sessions/workshops held on investment beliefs • Number of meetings with academics/external think tanks to discuss beliefs
Investment strategy	• Number of fully documented asset classes • Number of updated market documents • Quantified value add per investment strategy • Number of new investment strategies proposed versus strategies canceled • Number of different instruments required per investment strategy • Turnover per investment strategy
Skill alignment	• Number of staff per investment strategy • Inventory of team characteristics per investment strategy • Information coefficient per investment manager • Number of identified knowledge assets within the firm • Amount spent on formal training per investment manager • Ratio of front office to back office
Data and technology	• Number of issues reported by data integrity board • Number of internal and external data sources • "Actual-target" comparison of data and information • Number of system used and number of system upgrades within a certain period • Computational power of the different systems • Number of positive sign-offs by investment staff on technological changes

Table 1 – KPIs

7 In general, investment managers are considered skillful when they get more than half of the investment decisions right.

8 For example, market risk, counterparty risk, liquidity risk, and operational risk.

9 The KPIs given were just a limited and straightforward set for illustration purposes only.

10 Clearly, next to the internal analysis, the CIO should be very interested in the set-up of its main competitors in order to find specific strengths and weaknesses.

IMPLICATIONS AND CONCLUSIONS

Financial markets have been the beneficiaries of a three-decade decline in interest rates. This has meant that generous passive market returns have contributed significantly to overall portfolio returns. Adding value above the market in this period was nice, but it was not critical for funds to achieve their objectives. Looking to the future, we are facing a more modest outlook for long-term financial market returns, heightening the importance of adding value above benchmarks. Indeed, value added returns will inevitably become a significant contributor to overall portfolio returns in the future. And delivering these value-added returns will require rethinking the way we assess, access and manage investment opportunities. It will require far more sophisticated KM.

It is important to note that the best investments tend to be found in areas where markets are inefficient and where information does not freely travel. It is perhaps an oversimplification to say it, but if an opportunity fits in a box or a silo, it is likely overbid and overvalued. The best investors thus use their unique characteristics in a deliberate attempt to move into markets with minimal competition. For example, being a long-term investor offers additional options to what short-term investors can do. Moreover, being a local trusted partner to companies and project developers in a given jurisdiction can create unique and proprietary opportunities. Finally, a large investor may be constrained in its ability to access top managers, pushing it into alternative access points for similar risk exposures.

It is important we understand and include the unique characteristics of our investment organization in any strategy we formulate to guide our investing. Generally, the unique characteristics of an investor can be broken down into three categories: people, market intelligence, and governance. Persistent outperformance requires an investment organization to apply high caliber people and efficient processes in creative ways to develop proprietary sources of information and, ultimately, knowledge. And it is this knowledge that allows investors to generate outperformance, which will go hand-in-hand with statistical proof on historical track records. .

Put another way, maximizing the returns that can be achieved per unit of risk and per fee dollar spent (implicit and explicit) requires an organization that is thoughtful about its own advantages and proactively seeks to use those advantages in the context of broader market forces. In our view, the AM industry has underappreciated the power of KM in this regard, but this will soon change.

REFERENCES

- Aberbach, J. F., and B. A. Rockman, 2002, "Conducting and coding elite interviews," Political Science and Politics 35:4, 673–676
- Ang, A., 2014, Asset management; a systematic approach to factor investing, Oxford University Press
- Ball, D. E. A., 2006, International business: the challenge of global competition, McGrawHill/Irwin
- Beckers, S., and G. Vaughan, 2001, "Small is beautiful," Journal of Portfolio Management 27:4, 9-17
- Castles, F. G., and P. Mair, 1984, "Left right political scales – some expert judgments," European Journal of Political Research 12:1, 73-88
- Chaudhuri, R., Z. Ivkovich, J. M. Pollet, and C. Trzcinka, 2013, "What a difference a Ph.D. makes: more than three little letters," SSRN: 2344938
- Citi Business Advisory Services, 2014, "Big Data & Investment Management," report. Available at: https://www.managedfunds.org/industry-resources/industry-research/big-data-investment-management-potential-quantify-traditionally-qualitative-factors-citi/
- Clark, G. L., 2014, "Information, knowledge, and investing in offshore financial markets," Journal of Sustainable Finance & Investment 4:4, 299-320
- Clark, G. L. and R. Urwin, 2008, "Best-practice pension fund governance," Journal of Asset Management 9:1, 2-21
- Clark, G. L., and A. H. B. Monk, 2013, "Financial institutions, information, & investing-at-a-distance," Environment & Planning A. 45:6, 1318 – 1336
- Clark, G. L., and A. H. B. Monk, forthcoming, "The production of investment returns in spatially extensive financial markets," Professional Geographer
- Coase, R. H., 1937, "The nature of the firm," Economica 4:16, 386-405
- Connor, K. R., and C. K. Prahalad, 1996, "A resource-based theory of the firm: knowledge versus opportunism," Organization Science 7:5, 477-501
- DeFranco, G., and Y. Zhou, 2007, "The performance of analysts with a CFA designation: the role of human-capital and signaling theories," SSRN: 1019292
- Denzin, N. K., 1970, The research act in sociology: a theoretical introduction to sociological methods, Butterworths
- Fama, E. F., 1965, "The behavior of stock-market prices," The Journal of Business 38:1, 34-105
- Fang, Y., and H. Wang, 2015, "Fund manager characteristics and performance," Investment Analyst Journal 44:1, 102-116
- Feldstein, M., 2000, "The NBER-Sloan project on productivity change," AEA introduction, available online: http://www.nber.org/sloan/AEAintro.html
- Finstad, D., 2005, "Institutional or entrepreneurial management? An analysis of organizational factors and their effect on manager performance," Canadian Investment Review, Spring, 17-24
- Gertler, J. E., 2002, "Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there)," The Journal of Economic Geography 3:1, 75-99
- Grant, R. M., 1996, "Toward a knowledge-based theory of the firm," Strategic Management Journal 17, 109-122
- Greenwood, R., and S. Nagel, 2006, "Inexperienced investors and bubbles," AFA meetings paper, New Orleans
- Greenwood, R., and S. Nagel, 2008, "Inexperienced investors and bubbles," NBER Working Paper No. 14111
- Grinold, R. C., 1989, "The fundamental law of active management," The Journal of Portfolio Management 15:3, 30-37
- Grinhold, R. C., and R. N. Kahn, 2000, Active portfolio management, McGraw Hill
- Harvey, C. R., Y. Liu and H. Zhu, 2014, "….and the cross-section of expected return," Fuqua School of Business, National Bureau of Economic Research working paper no. 20592
- Harvey, C. R., and Y. Liu, 2016, "Rethinking performance evaluation," Working paper, SSRN: 2691658
- Helper, S., 2000, "Economists and field research: "You can observe a lot just by watching," American Economic Review 90:2, 228-232
- Huij, J., and E. van Gelderen, 2014, "Academic knowledge dissemination in the mutual fund industry: can mutual funds successfully adopt factor investing strategies," Journal of Portfolio Management 40:4, 157-167

- Ineichen, A. M., 2004, "Absolute returns; the future in wealth management?," Journal of Wealth Management 7:1, 64-74
- Javernick-Will, A. N., and R. E. Levitt, 2010, "Mobilizing institutional knowledge for international projects," Journal of Construction Engineering and Management 136:4, 430-441
- Kraaijenbrink, J., and J. C. Spender, 2011, "Theories of the firm and their value creation assumptions," paper presented at the SMS 31st Annual International Conference, Miami, U.S.
- Leibowitz, M. L., 2005, "Alpha hunters and beta grazers," Financial Analysts Journal 61:5, 32-39
- Lintner, J., 1965, "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets," Review of Economics and Statistics 47:1, 13–37
- Lo, A., 2004, "The adaptive market hypothesis: market efficiency from an evolutionary perspective," Journal of Portfolio Management 30:5, 15-29
- Mische, M. A., 2001, Strategic renewal, organizational change for competitive advantage, Prentice Hall
- Moussavou, J., 2006, "Organizational architecture and decision making," Journal of Portfolio Management 33:1, 103-111
- Nonaka, I., 1991, "The knowledge-creating company," Harvard Business Review 69:6, 96-104
- Nonaka, I., and H. Takeuchi, 1995, The knowledge creating company, Oxford University Press
- O'Leary, D. E., 2002a, "Knowledge management in accounting and professional services," in: Arnold V., and S. G. Sutton (eds.), Research accounting as an information systems discipline, American Accounting Association, 273-283
- O'Leary, D. E., 2002b, "Knowledge management across the enterprise resource planning systems life cycle," International Journal of Accounting Information Systems 3:2, 99-110
- Penrose, E. G., 1959, The theory of the firm, Wiley
- Pozen, R., and T. Hamacher, 2011, "Most likely to succeed: leadership in the fund industry," Financial Analysts Journal 67:6, 21-28
- Prusak, L., 2001, "Where did knowledge management come from?," IBM Systems Journal 40:4, 1002-1007
- Ross, S., 1976, "The arbitrage theory of capital asset pricing," Journal of Economic Theory 13:3, 341-360
- Sharpe, W. F., 1964, "Capital asset prices: a theory of market equilibrium under conditions of risk." Journal of Finance 19:3, 425-442
- Spender, J. C., 1996, "Making knowledge the basis of a dynamic theory of the firm," Strategic Management Journal 17, 45-62
- Strauss, A. L., and J. M. Corbin, 1998, Basics of qualitative research: techniques and procedures for developing grounded theory, Thousand Oaks
- Treynor, J. L., 1961, "Towards a theory of market value of risky assets," unpublished manuscript
- Zack, M. H., 1999, "Developing a knowledge strategy," California Management Review 41:3, 125-145

APPENDIX 1: VBA SURVEY

Table A1.1: KM's value to an asset manager

1. What is the primary focus of KM?	
Data and Information	21%
The firm's knowledge and expertise	35%
Knowledge assets, being the informal structure of knowledge	28%
Intangible assets, such as brand name, partnerships, and goodwill	14%
Don't know	2%
2. What activity can benefit the most from KM?	
Investments	33%
Research and strategy	35%
Operations and IT	12%
Marketing and sales	19%
Don't know	1%
3. What activity relies on implicit knowledge the most?	
Investments	40%
Research and strategy	28%
Operations and IT	8%
Marketing and sales	24%
4. What activity would benefit the most from turning implicit into explicit knowledge?	
Investments	35%
Research and strategy	30%
Operations and IT	11%
Marketing and Sales	24%
5. Who should be responsible for KM?	
Chief Executive Officer	26%
Chief Investment Officer	37%
Chief Financial Officer	3%
Chief Information and Technology Officer	11%
Chief Marketing and Sales Officer	3%
It concerns a line-responsibility	20%

Table A1.2: Type of knowledge related to investment performance

1. The generation of market performance (beta) is a function of	
Explicit knowledge	30%
Implicit knowledge	3%
Explicit and implicit knowledge	55%
Don't know	12%
2. The generation of excess performance (alpha) is a function of	
Explicit knowledge	5%
Implicit knowledge	7%
Explicit and implicit knowledge	80%
Don't know	8%
3. Would knowledge management harm the performance of a star-performer	
Yes	21%
No	49%
Don't know	30%
4. Does a direct relation exist between knowledge assets and academic research?	
Yes	50%
No	27%
Don't know	23%

Table A1.3: Ways in which investment organizations can operationalize KM

1. What is the most effective manner to share knowledge?	
Informal and regular talks	39%
Formal meeting schedule	16%
Training-on-the-job	22%
Internal courses	17%
External courses	5%
Don't know	1%
2. What is the main barrier to overcome in KM?	
People don't share knowledge that gives them a competitive edge	30%
There is no individual financial reward for sharing knowledge	32%
Knowledge is too specific; sharing has no impact	12%
Most knowledge is tacit and cannot be coded	18%
There is no barrier	7%
Don't know	1%
3. How can knowledge assets be protected?	
Specific clauses in labor contracts	13%
Continuing education and innovation	42%
Treat staff on a "need to know" basis	3%
Patents	2%
Knowledge assets cannot be protected	39%
Don't know	1%
4. Is the value of knowledge assets context dependent?	
Yes	77%
No	14%
Don't know	8%
5. Can the value of knowledge assets be measured in terms of money?	
Yes	34%
No	34%
Don't know	32%
6. Does your organization use KPIs re KM?	
Yes	9%
No	84%
Don't know	7%

APPENDIX 2: P&I SURVEY

Table A2.1: Added value KM

1. Given the definition of KM, please select the statement that best reflects when KM would be of importance for your organization	
KM is never important for my organization	2%
KM is only important for my organization during times of "normal" market activity	3%
KM is only important for my organization during times of "abnormal" market activity	3%
KM is sometimes important for my organization for reasons that do not depend on market activity	18%
KM is always important for my organization	74%
2. If an AM firm does not have a KM system in place, do you think it could achieve a more stable business model by using one?	
Yes	55%
No	6%
Do not know	32%
3. If an AM firm does have a KM system, do you think that system contributes to a more stable business model?	
Yes	65%
No	5%
Do not know	25%
4. Do you believe that building or improving KM systems justifies higher fee levels?	
Yes	15%
No	68%
Do not know	17%
5. If KM were a board responsibility, who should be responsible?	
Chief Executive Officer	27%
Chief Financial Officer	4%
Chief Operating Officer	16%
Chief Client Officer	1%
Chief Technology Officer	2%
Chief Investment Officer	30%
KM is not a board responsibility	17%
Other	4%

Table A2.2: Type of knowledge related to investment performance

1. For generating "beta" return (market return), which of the following is most important?	
Explicit knowledge (what can be codified)	25%
Implicit knowledge (what is difficult to codify, such as experience)	6%
Explicit and implicit knowledge	59%
Neither are important	4%
Don't know	7%
2. For generating "alpha" return (excess return), which of the following is most important?	
Explicit knowledge (what can be codified)	14%
Implicit knowledge (what is difficult to codify, such as experience)	25%
Explicit and implicit knowledge	56%
Neither are important	1%
Don't know	4%
3. Do you believe that the collective knowledge of investment teams is more critical to generating excess returns than the individual knowledge of a star performer?	
Yes	73%
No	15%
Don't know	12%
4. Do you believe that integrating a star performer's knowledge into the organization's pool of shared knowledge would help or harm the star performer's investment performance?	
Harm	8%
Help	62%
Neither harm or help	17%
Don't know	13%
5. Do you believe that results found in academic research will lead to better investment strategies?	
Yes	64%
No	18%
Don't know	18%

Table A2.3: Points of particular interest in KM

1. What would be the most effective knowledge transfer process?	
Daily, informal one-on-one meetings	38%
Formal business meetings	9%
Training-on-the-job	17%
Internal professional training	20%
External professional training	5%
Other	11%
2. What is the biggest hurdle to setting up KM within an AM firm?	
Portfolio managers protect "their" knowledge as it gives them a competitive edge	32%
Compensation structures are not linked to sharing knowledge	33%
Most knowledge is so specialized that it doesn't make sense to share	3%
It is simply not possible to make most investment knowledge explicit and/or to codify	17%
There is no need: all necessary knowledge is readily available	4%
Other	12%
3. Do you believe that firms can and should set KPIs that are specific to KM?	
Yes	59%
No	14%
Don't know	27%
4. Which of the following incentives should be successful in encouraging transfer of tacit knowledge (i.e., knowledge that is difficult to codify)?	
Bonuses linked to transferring tacit knowledge to the organization (e.g., reports)	24%
Bonuses linked to transferring tacit knowledge between individuals (e.g., mentoring)	23%
Non-monetary incentives linked to transferring tacit knowledge to the organizations	15%
Non-monetary incentives linked to transferring tacit knowledge between individuals	17%
Incentives do not work for tacit knowledge transfer	13%
Do not know	8%
5. Can incentives be used to make managers self-critical in examining the knowledge they possess?	
Yes, monetary incentives can be used	17%
Yes, non-monetary incentives can be used	9%
Both can be used	43%
No, incentives do not help managers to become self-critical of the knowledge they possess	16%
Do not know	15%

6. Can KM systems be constructed so that they successfully delete or update knowledge that has become obsolete?	
Yes	33%
No, deletion of obsolete knowledge could not occur in a timely manner	3%
No, it would be too difficult to accurately identify obsolete knowledge for deletion	22%
Both "no" answers apply	14%
Do not know	28%

Private Equity Capital Commitments: An Options-Theoretic Risk Management Approach

Andrew Freeman – Fellow, Centre for Risk Studies, Judge Business School, Cambridge University
D. Sykes Wilford – Hipp Chair Professor of Business and Finance, the Citadel[1]

Abstract
The capital call for a private equity (PE) firm has been described in various modeling approaches almost entirely from the perspective of the investor, from the liquidity implications, in the context of Modigliani-Miller, a Merton type environment, or using forms of Markowitz allocation modeling. In most articles, the nature of the call option (written by the investor and owned by the fund), as it relates to providing liquidity, is assumed. This article narrows the discussion by focusing on the risk of PE firms during a financial or economic crisis. Two sets of options are analyzed. First, we examine the ability to call funds when opportunities arise during periods of market stress. Second, a PE firm's highly flexible ability to "put" holdings to the markets, by waiting for more opportune times to do so, is discussed. Our aim is to better understand the market risk associated with any single PE fund and in aggregate the risk of the PE firm, with respect to these options held by the fund or firm. In an attempt to quantify the risk of the PE fund or firm it is essential to understand from various perspectives the option-like qualities of the contracts that the fund has with its investors. The conclusions should be obvious to the risk manager of a firm or fund, but are often blurred in an attempt to make these investments fit into simple VaR systems or more complex theoretical models. The implicit option-like characteristics create a set of offsets to potential mark-to-market losses as market volatility changes, especially with respect to markets in crisis. The old saw that a financial crisis is a "friend" of the PE firm is vindicated to some extent by the analysis in this article from the perspective of risk measurement and management.

1 The authors wish to thank Charles W. Smithson, Bluford Putnam, and Martin Pergler for their comments. Any errors are those of the authors.

INTRODUCTION

When a private equity (PE) fund is raised, investors give commitments to make an agreed amount of investment capital available at the behest of the fund manager. The contract with the investor allows the PE firm to draw down the commitments via a series of capital calls. In a typical fund with a 10-year life cycle, the expectation is that commitments will have been called fully or almost fully by no later than the end of the fourth or fifth year and that during the second half of the fund's life cycle capital will steadily, if unpredictably, be returned to investors via distributions as a consequence, first, of dividends from investments made, and, second, of eventual exits from past transactions.

This arrangement appears straightforward enough. In fact, applying the lens of options theory suggests that it is anything but. It can shed interesting light on an important strategic question of risk measurement for PE funds, PE firms, for their investors, as well as for regulators and institutions concerned with macroeconomic and financial system stability. Although there are many questions to be addressed vis-à-vis the different regulators, risk managers, investors, and the PE funds themselves, most of these questions link back to the relationship of the PE fund structure to public markets. For example, (1) how vulnerable are PE firms to volatility in public equity markets? (2) In the scenario of an extreme market decline, would PE firms be to any extent immune from systemic collapse? And (3) would that immunity be size-dependent? To answer basic risk-related questions for PE funds or PE firms, we must focus our analysis on the option-like capital call provisions associated with any PE fund structure.

The right to make capital calls is a time-variant call option that is written to the PE fund manager by investors. Deriving its value is complex because the option contains a significant element of liquidity premium; the ability of the PE fund (and therefore the PE firm) to call capital when it is becoming more valuable, in part because liquidity is becoming scarcer, is a powerful right during times of stress. While we can posit with confidence that the value of a PE firm would be negatively impacted by a sharp fall in public market valuations, the call option forms a natural offset both in terms of liquidity and by giving the PE firm (through the funds it manages) the opportunity to buy assets at distressed prices. In "normal" times, the call option has clear value. However, we argue that its value is actually greatest during a crisis and its aftermath.

Further, in contrast to an active investor in public markets with readily tradable securities and consequent exposure to liquidity risk in times of crisis, PE fund managers also have the option to put individual investments in the fund to the market largely, but not wholly, at times of their choosing. In effect, they can to some extent wait out a crisis. Again, this provides an interesting option to the PE fund, and therefore the investor in a fund, relative to a typical equity investment. The combination of options that the PE fund and fund manager hold can potentially be exploited, thereby neutralizing certain risks faced by other types of equity-based funds.

PE FIRMS DURING CRISES: OPTIONS-BASED CRISIS IMMUNITY BUILT IN?

Using a simple schema we can observe that PE funds bring a degree of liquidity to otherwise illiquid corporate assets, and over time they tend to deliver higher return-on-equity (RoE) than diversified portfolios of public equities.[2] For this reason, PE fund managers receive rewards in the form of management fees and carried interest. Their opportunity to deploy liquidity expands across more than one dimension in times of financial or economic crisis when public valuations are depressed. Liquidity by itself becomes more valuable. Assets become cheaper. The range of assets available to a PE fund expands because many more companies will fall into the enterprise value parameters typically written in PE fund governing documents.

We can make the following assertions about the liquidity call:

- PE funds appear to be paid differently for the liquidity they put to work
  - Before a crisis
  - During a crisis
  - Following a crisis
- The value of the capital call option changes during these three periods
  - Before a crisis it can be valued based upon observed prices in the market and low or stable volatilities
  - During a crisis volatility has risen sharply and equity prices are typically lower
  - The demand for liquidity is greater
  - The return on liquidity is greater
- Once a crisis is resolved, PE funds not yet put to work return to a lower volatility-defined value

2 We will argue that even if the RoE is lower, it may still be preferable due to the risk mitigating nature of the call option provided to the fund and therefore may be a risk diversifier for the typical investment entity.

The capital call option is not like a typical option, which gives the owner the right but not the obligation to exercise and will expire worthless if not exercised. It gives the PE firm the right to call for capital, just like a normal option, but it also includes an element of obligation – the PE firm must invest an amount of the committed capital by a set date or it will breach the governing documents of the relevant fund, triggering investor rights for redemptions, fund closure, etc. However, from the perspective of the PE firm, it has an unusually long time horizon, typically measured in years or months rather than weeks or days.[3] While one can measure the runoff of the options held by the PE firm as a redemption date nears, for illustrative purposes we will consider the options as long dated and not deal with this issue here.[4]

The risk profile of a PE fund is, therefore, correlated with the extent of the value of its capital call options at a given moment in time. As such, the maturity structure of a PE firm's portfolio of funds will ultimately determine the extent of its robustness over the economic cycle, in particular its ability to deploy precious capital during periods of distress. PE firms – managers of the fund – that are unable to make such deployments are at a strategic disadvantage versus their competitors. This insight might help to explain industry dynamics such as fundraising cycles. However, it raises additional non-trivial problems. Some of the large PE firms are themselves listed on stock markets, meaning their own equity is likely correlated with public valuations in ways that may not be fully understood. Moreover, they often have hundreds of funds, some huge some small, across a range of different maturities – some will be newly raised, others winding down.[5]

Further, PE funds in wind-down (i.e., they are returning investment proceeds to investors) also have interesting option-like characteristics, which suggest that PE funds actually own puts, as well as the calls noted above. Because they must sell assets, including via flotations on public markets, they are at first glance vulnerable during times of stress and will suffer from lower valuations and reduced liquidity. In most cases, however, they have considerable choice as to exactly when to do so. In effect, the PE fund owns a set of options to put assets back, either directly to its investors, to other trade buyers or to public markets. However, depending on the precise conditions of a portfolio of funds, some of the put options will retain significant time value because they might not need to be fully exercised for months or even years. In other words, the PE fund might be immune, even if only partially, to short-term crises or even a crisis of considerable duration. In addition, a fund that is winding down will tend to be reducing its leverage, meaning that it is progressively less vulnerable to market conditions. The liquidity call options, combined with the flexibility to put existing assets to the market at the choice of the fund, potentially smooth out risk over time.

Another feature of the capital call option is that by definition its "crisis premium" decays in proportion to the extent that the relevant fund makes investments; more investments means less cash to call. The less capital that remains to be called, the less value the PE (fund) firm can extract from distressed markets. For a typical PE firm that is diversified across time, sectors, geographies, and asset classes the timing option to put an investment to the market certainly mitigates the risk of a particular fund that is winding down, with respect to the declining value of the call option. By maintaining funds at different stages of maturity or drawdown, a PE firm in effect hedges itself against extremes of volatility.[6] This insight offers a new dimension for thinking about the influence of "dry powder" – this is the collective amount of capital available at any one time for deployment by the PE industry. Whereas dry powder is typically, and not incorrectly, viewed as a measure of the competitive pressure on PE firms to do deals at more aggressive prices in good times, it might just as significantly be seen as a gauge of the industry's likely resilience in the event of a crisis.

And, some degree of "crisis" should be viewed as the norm. Karagiannidis and Wilford (2015) suggest that crises in equity markets occur every five to 10 years. However, demonstrably each crisis is different and, therefore, has different implications for valuations, not only with regard to the underlying assets in a PE fund, but also to the call option, owing to typically wide swings in volatility. For example, prior to the recent financial crisis, the volatility of the S&P jumped from about 16% (average for the previous six years) to almost 60%, before returning to the lower post-crisis levels.[7] A glance at the historical movements in the VIX index shows a similar picture. In the mid-2000s it was in the low teens before spiking to above 60 during the crisis. And even after the crisis, spikes occurred that reached the high 30s.[8]

3 For our purposes we will not attempt to model the implied run-offs in the value of the call due to the contractual obligation embedded in the PE Fund governing documents that committed funds must be invested by a set date. Any attempt to measure the changing values of the call option must a priori take this obligation into account. Indeed, in valuing the call option, one should consider the extent of decay as a fund reaches the point where it needs to be nearly or fully invested.

4 Our analysis is necessarily limited. We do not examine the embedded options referred to above. Some PE firms begin fund-raising on a next-generation fund once the existing fund passes its investment thresholds, since the ability to refresh its contingent capital via closing of a new fund appears to be critical to the PE firm's ongoing financial resilience.

5 We do not address the potential for a mispricing of the publically traded funds in this analysis, but it may be the case that the market trades these firms at a much higher risk premium, especially during a crisis, than is warranted. Further research is needed on this interesting question.

6 A much more formal approach to the issues related to valuations relative to publicly traded positions is found in Sorenson et al. (2013).

7 See Wilford (2014) for more discussion of the relevant volatility measures.

8 See http://www.cboe.com/delayedquote/advchart.aspx?ticker=VIX.

Spikes in volatility are the "lifeblood" of options. It is these spikes that make the call/put provisions of a PE firm so valuable; in essence providing a cushion to the risk associated with funds and, by implication, PE firms. Interestingly, size matters with respect to how the risk is mitigated by options ownership for the PE firm. Indeed, small PE firm (and perhaps venture) assets are more likely to become attractive to large PE firms during a crisis because they may carry less options-based immunity and will, therefore, decline in value to a greater extent than large PE firm assets. This helps to explain why the valuation challenge is so severe.

Whereas a generalized model is not attempted in this analysis, careful illustration of the options implicit in the structure of the PE fund can demonstrate the potential this perspective offers for risk mitigation.[9] Using a simple construct, the PE firm's risk may be mitigated by the aggregation of the embedded options in the funds that the firm manages.[10]

VALUING THE OWNERSHIP OF THE PE CALL OPTION – FROM WRITER TO THE FUND, AND THE PE FIRM

It should already be clear that valuing the PE call at the fund level is required before statements about the value of these options to the PE firm can be made. Any application to risk measurement must take into account the different maturity characteristics of individual funds in such a way that these can be aggregated into a meaningful picture of a PE firm's portfolio where the firm operates multiple funds.

Moreover, the term "PE firm" is something of a misnomer, because the leading firms have increasingly diversified across asset classes since the recent financial crisis. As banks reacted to new regulatory capital and liquidity requirements, they reduced their lending to parts of the capital markets, including middle-market commercial lending and mezzanine financing. This created an opportunity for PE firms to offer funding in non-pure-equity areas of companies' capital structures, including the flexibility to create combined equity/debt solutions. The addition of debt capital instruments created a new layer of complexity to PE risk profiles and has yet to be clearly analyzed. But, to the extent that a PE firm has non-PE assets, then it will react differently to external market conditions, particularly during periods of stress or crises. Existing debt structures, especially mezzanine financing, imply unique pricing problems, almost all of which are based on options theory. As per above, the risk of government debt rose dramatically during the crisis as measured by volatility.[11] Based upon Merton (1974), corporate debt can be viewed as a put option on the value of the firm plus a close to risk-free government bond. Combine this fact with the right to invest in such debt at moments of extreme risk (when the put is most negatively affected by a sharp rise in volatility) and the call provision again offers a cushion to the risk of a PE fund; potentially offsetting the negative implications for increasing spreads on high yield instruments that may be held in, say, a mezzanine fund.[12]

To simplify matters we note that owning the right to call – whether it is to invest in equity directly, indirectly through mezzanine debt, leveraged through warrants that typically are attached to mezzanine transactions, or any combination thereof – is valuable and may offset some of the negative implications for a fund or the fund manager with respect to the inherent risks resulting from a crisis movement in market prices (and their declines). Particular circumstances will dictate the methodological approaches necessary to measure the value of these options correctly; the key is to recognize that they are valuable and that their value changes radically during a crisis in ways that may be inversely related to movements in public markets.

Will the investor deliver on the call and PE firm (fund) implications?

An underlying assumption of an options-based approach to measuring the volatility of PE firm valuations is that the capital call options change in value over time depending on market conditions and related asset prices, but that they can always be exercised because investors are contractually obliged to honor all capital calls. Investors subscribe to funds governed by by-laws and agreements that are set out in advance. In effect, they would be in serious breach of contract were they to refuse to meet a capital call. The same would apply if they were unable to meet a call because of distress and a subsequent lack of necessary liquidity. Under the terms of a typical fund, the PE firm would then have the right to eject the investor from the fund and seek a secondary sale of the relevant commitment. It could also revert to legal proceedings for breach of contractual obligations, particularly if the investor's unwillingness to pay could be shown to have damaged the interests of other investors in the fund.

9 It can be argued that a "building block" approach to measuring the risk mitigation implicit in the call and put options can yield very practical results in pricing and measuring risk. This is based on the building block focus of Smithson (1998) in valuing derivatives and financial structures.

10 For a typical PE fund investor knowledge about the degree of risk mitigation at the firm level (perhaps size is correlated, ceteris paribus) may be important information in choosing a fund in which to invest.

11 We are ignoring direction of prices and, therefore, correlations to the S&P.

12 See Jensen and Meckling (1976) and Wibaut and Wilford (2009) for theoretical analyses, as well as implications of volatility spikes on corporate debt.

In practice, however, would a PE firm resort to legal action? It would have to take into account overlapping factors. Suing an investor in distress might lead to serious reputational harm for the PE firm, even if the legal case for doing so were ironclad. The investor concerned might have investments across multiple funds, meaning it would have an important relationship with the PE firm that would then be in jeopardy. A lawsuit could also send a negative signal to other investors, both existing and prospective, decreasing their appetite for future commitments. These real costs might outweigh the cost of the missing capital that was called and not delivered. Assuming the capital call was to fund an investment, the PE firm might be better off making good the opportunity from its own capital in order to drive home its advantage.

Depending on the circumstances, then, a PE firm might decline to exercise its right to enforce a call option. Again, this further complicates the underlying problem. At some point, the risk can become so high that the writer of the call may default on their obligation. To the extent that the PE firm's utilization of this option to measure (offset) the increased risk as suggested by the market is compromised, this benefit is lost. From an options-theoretic perspective, such an event suggests that a risk manager should consider vega risk for the PE firm, although the lack of information about the circumstances under which default on the contract may occur means that it is difficult to measure such a risk. In practice, the recent financial crisis witnessed many institutional investors in PE having to find strategies to meet their commitments that on the surface caused significant difficulties in valuing the assets of the institutions and perhaps led to some wealth destruction.[13]

A SIMPLE ILLUSTRATION OF THE VALUE OF THE OPTION TO THE PE FIRM

Previously, we have stressed the actual complexity of the liquidity call embedded in the PE fund contracts. In effect, to apply this methodology to an existing fund, in an attempt to enhance any risk measurement system, each of these complex options must be broken into its component parts and recombined into whatever system of risk measurement is appropriate for the PE fund or firm. Alternatively, it would be possible to create an overall model that considers all of the risks inherent in the contract.

Our analysis will focus on some of the simpler aspects of the call provisions (as well as the put options implicit in a fund with respect to timing of placement of invested capital). In doing so, we begin the process of measuring the implications of the ownership of the call (put) options by the PE fund and in addition, the firm. Armed with simple snapshots of the change in value of the option to the PE fund we can then suggest some logical implications for risk measurement and perhaps the management of risk.

The liquidity call – a simple approach

The impact of a financial crisis on the implied value of the call can easily be illustrated with a simple pricing model. Using measures of volatility from Wilford (2014) as guides, we can calculate the value of the call from the perspective of the fund by simply allowing volatility to vary with the pre, post and crisis periods. This can provide us with snapshot comparisons of the value of a call.

	Pre crisis	Crisis	Post crisis	Recent period
Volatility	16%	57%	17%	13%
Value of call as a percentage of the notional principal	18.54	49.91	19.44	16.25

Table 1 – Call option value

As a first exercise, let us assume a five-year call period (typical of many funds) and set the strike price at 100% of the exercise price. Using an interest rate of 2% a rough calculation of the value of the call can be estimated. For our purposes, we simply desire to understand the implication of the rise in volatility on the value of the call.

This illustrates how a crisis will drive up the value of the call. In this exercise, each period was calculated assuming the call starts during the relevant period. This is not reality, of course, but a near tripling of the value of the call illustrates the point that the holder of that call has had an increase in value just when it is needed most

13 See, for example, Phalippou and Westerfield, 2014, "Cash-poor LPs face capital-call pressure," citing Private Equity Insider, footnote 3. November 5, 2008: "Brown University, Calpers and Carnegie Corp. are suddenly finding it hard to meet capital calls from private equity fund managers (...) A growing set of limited partners find themselves short on cash amid the financial crisis and thus are scrambling for ways to make good on undrawn obligations to private equity vehicles. Among those in the same boat: Duke University Management, Stanford Management, University of Chicago and University of Virginia... Brown, whose $2.3 billion endowment has a 15% allocation for private equity products, is apparently thinking about redeeming capital from hedge funds to raise the money it needs to meet upcoming capital calls from private equity firms. That's similar to a strategy that University of Virginia is employing... Carnegie, a $3.1 billion charitable foundation, is also in a squeeze. Its managers have been calling on commitments faster than expected, while distributions from older funds have slowed down, creating a cash shortfall. As for Duke, the university's endowment has been named as one of the players most likely to default on private equity fund commitments. That partly explains a massive secondary-market offering that the school floated last month, as it sought to raise much-needed cash and get off the hook for undrawn obligations by unloading most of its $2 billion of holdings in the sector... Some of the bigger investors are considering tapping credit facilities to meet near-term capital calls."

from the perspective of measuring and managing risk.

Alternatively, once a crisis passes and the rest of the investment portfolio of a PE fund or firm regains some of its mark-to-market value, the value of the call falls.[14] The opposite movement in risk measures will never be symmetrical in nature, but the point should be clear. Just when a fund may find itself under pressure from falling value of existing holdings, the opposite is occurring with respect to the call provisions and perhaps sufficiently to have a significant impact on overall values of the fund (or firm).

To make the exercise more meaningful, let us now assume that a firm entered into a contract with an investor one year before the onslaught of the financial crisis. After one year the call will have had time decay. Thus, instead of the call being valued at nearly 50% of the notional principal, as per above it, will have a value of 5% less. Of course, as the time period changes, so will the value and the decay factor will eventually have a waterfall effect on the gains from an increase in volatility. Still, the fundamental point holds that during a crisis the value of the call may offset to a great extent the implied mark-to-market decline in the value of the PE fund's holdings resulting from the onset of a crisis.[15]

Implied price	100	80	60	50
Value of call	50%	52%	59%	64%

Table 2 – Pricing the strike

As noted above, a crisis might offer better opportunities for a PE firm than a bull market. If this is the case, then the exercise price of the call and the implied spot price will vary with a shock (implied price means the percentage of the underlying price). Indeed, one could argue that as the general market decline occurs so will the value of the company to be targeted. One way to model this then is to allow the exercise price to decline in the option calculation. During the financial crisis, equity markets plummeted by some 50%, more or less, depending on the market. In a simple risk analytics mark-to-market model one could argue that the existing investments of some PE firms would also be impacted by 50%. If so, then how is the crisis the lifeblood of PE investing? The answer is obvious; with the effective cash available due to the call provision the PE fund can buy assets at these reduced prices. Put another way, from a risk management perspective, the call offsets, to some extent, the decline in mark-to-market prices.

Let us now assume a 50% fall in prices and examine what happens to the option as it is effectively coming more into the money. In our case above of a one-year time decay, the value of the option moves from approximately 45% to 64%, or an increase of 50% in value. Such a large fall in general market prices is unusual, so perhaps by doing the same exercise but holding pre and crisis volatilities the same while allowing the implicit fall in the exercise price we can see more clearly the risk-mitigating potential of the call for the PE fund (firm).

The implied price decline of a targeted investment for a fund that would require a call may or may not fall so significantly with the rise in volatility.[16] It is, however, a factor that needs to be considered in understanding the overall impact of the call provision on risk measurement and obviously there are implications for risk management of the PE fund or firm.[17] In a simplified form, at a market price of 50 the fund would be able to buy twice as much as if the price was 100, but from an options value perspective the fact that the option calculation implies that it is much more in the money suggests that its impact is less relative to the underlying impact of lower prices.

Clearly a fund that has just finished fund raising and is about to begin deploying capital has a huge advantage over one that is nearing maturity with respect to the value of the call for risk measurement and management purposes. One may think of the combination of the call versus the percentage of the call contracts called as a vega risk problem. For the sake of simplicity, however, we think that the measurement of the risk of the actual investments resulting from a crisis should be considered separately from that of the call itself. The overall risk of the fund can thus be aggregated based on the percentage called (and therefore invested). Although in combination it may be intellectually interesting to look at the percentage of the call outstanding versus what has been called as a vega calculation problem, the simpler aggregation of implied values should yield similar and intuitively understood results, especially when combined with mark-to-market or other valuation implications from general market movements in underlying prices.[18] Again, the problem is made more complex by the fact that PE valuations are rarely derived solely from observable market prices, but are constructed

14 Implicitly we have assumed that a serious crisis would be correlated with a decline in publically traded equity prices and a recovery would imply a movement toward the long-term trend of rising equity prices.

15 In any attempt to model the timing of a call occurring, one would need a simulation approach based on some random process. Here we are suggesting the basic concept, not any particular time when a call would be made.

16 Logically, these estimates of the impact on the value of the call due to a drop in the price of the underlying are subjective.

17 Fairly sophisticated modeling by Sorenson et al. (2013) integrates some of these issues. However, from an applied perspective we believe that it is necessary to consider the subjective nature of the risk measures for a crisis. Looking for perfect or near-perfect quantitative measures will often miss the point in practice.

18 Whatever system is chosen for measuring the variable risk of an asset, it must be considered along with the value of the options held by the fund because it has an inevitable implication for the correct input values for the option pricing scheme utilized.

by a process of peer comparisons and earnings discounting techniques.

Valuing the implied put option

So far we have focused on the value of the call to the PE fund (firm); however, as noted, PE firms have another set of options: they can time the putting of the private investment to the market. Tables 1 and 2 provide an insight into the value of this timing as well. It is precisely when the market is most risky in a crisis, with an overall decline in opportunity to sell the existing assets, that owning the put is most valuable to the fund (firm). A simple illustration of this is presented in Table 3.

Again, the value occurs in the crisis period when the market has been hurt so badly and the fund manager may be forced to write down the value of the holdings not yet put to the market. As the market recovers, the value of the put declines, as one would expect. For the risk manager of the PE fund or firm both of these implied options – call and put – are helping to control risk just when normal mark-to-market would suggest a large write down in the fund's value. The natural cushion imbedded in the PE structure itself mitigates the overall market risk associated with the investments.

It is interesting that from a risk management perspective both the puts and the calls implicit in the structure of PE funds mitigate the overall risk of a crisis significantly while not hurting the long-term performance of the fund. Existence of these options could easily offset declines in the market itself, and depending on the amount of the fund invested, but not brought to market, the crisis could actually raise the value of the PE fund while the market is in disarray. As such, the notion that crises are the lifeblood of the PE firm may not be a strange comment at all, but is in fact supported by the large movements in value of the options implied in the structure of the funds. Of course, the value of the call, as well as the put, will depend upon the percentage called and invested at any point in time, but this is mechanical and can easily be considered by the PE fund risk manager as well as aggregated at the PE firm level.

The correct value of the put, however, is much more difficult to determine. In Table 3 we made simple assumptions about the strike price for the put. In simulating over any meaningful period of time, the problem becomes much more difficult. What is the strike price to be considered? Is it the market price, some forward price, etc.? Obviously, if the market simply goes down and the investment itself cannot be taken to market, then eventually the value of the put will go to zero, no matter the volatility. Alternatively, the strike price could be held at its accounting value – what was paid for it or 100 – but this seems obviously incorrect in the case of a drawn out market downdraft. Equally, in the case of a significant rise in the market, choosing a strike price of 100 would imply that the put becomes worthless – we would never simply put it to the market at 100, so in principle the strike should rise along with the market price. Again, this assumption has obvious flaws in the event that we attempt a simulation of the put option over time. Our analysis is deliberately simpler in order to illustrate the value of the put under different conditions, so "knowing" the correct strike price is less of an issue.

	Pre crisis	Crisis	Post crisis	Recent period
Volatility	16%	57%	17%	13%
Value of put as a percentage of the notional principal	9.25	40.52	10.05	6.86

Table 3 – Put option value

GIVEN THE RISK, ARE INVESTORS PAID SUFFICIENTLY FOR THE OPTIONS THEY WRITE THE PE FIRM?

By definition, investors consent to writing a call option to the PE firm. According to options theory in a conventional traded market, any increase in value of the option to the PE firm must be offset by the same loss of value to the investor. During a crisis when volatility spikes, the option writer should feel that there is an implied drop in the value of the portfolio from which it was written, i.e., they have incurred a loss in favor of the PE firm.

Anecdotal evidence suggests that this is not the case. As one foundation manager has queried, "doesn't PE's lack of trading help dampen our risk?" There are several reasons why this could be the case. One relates to the obvious fact that PE firms largely buy non-traded equities, so the options noted above are intuitively considered. They may understand that when there has been a large general loss in public markets there are significant opportunities for PE firms to lock-in long-term bargains, so they tolerate the call option as part of the PE offering of outperformance, including its special role in times of crisis. Finally, the opportunity cost of the option they have written may be offset in some cases by the implied put's rise in value (per Table 3), at least to some extent, for the funds already called and invested. For the moment we ignore this point, although it is critical to the statement above, and we focus mainly on the call written.

If any of these arguments are correct, then the writer of the call option (investor) does not have a decrease in value when volatility rises to the extent that would be the case for a normal call. Purely

quantitative assessment of the gain/loss trade-off would struggle to reflect this complexity. We believe the implications are case specific. To find the key to why an otherwise asymmetrical deal in favor of the PE firm is not generally contested by investors, we consider the general conditions of the portfolio of the investor. The written call provision will have to be funded but how will it be funded may vary. In a simple valuation model, one would assume certainty in funding from the signing of a contract. In reality, actual investment calls come unevenly through the investment period of the fund. As we have shown above, the value of the call to the PE firm can be calculated given certain assumptions and information, but whether there is negative value to the investor, and, if so, its extent, is not nearly as clear.

Let us assume three possible conditions for a PE investor committed to a fund. At this stage we make no assumption as to the lifecycle position of the fund, but we can observe that it will still be in its investment period so that the call option has not expired and therefore retains value:

- The investor holds cash or T-bills "in case" of a call
- The investor holds investments with a beta of 1 that will have to be liquidated when a call is made
- The investor uses a readily available line of credit to fund the call – behind the PE firm's contingent capital lies another supplier of contingent funding in the form of a revolving credit facility

Let us now examine each condition in turn.

Condition 1

Here, the investor obtains the expected return on cash or cash alternatives, so has, in effect, a zero risk-return position (although as we know from recent history this might not be the case in reality, as the concept of a "risk-free asset" has been shown to be of theoretic, but not practical use). The call option does have an opportunity cost, but to the extent that the Treasury position reflects the intended asset allocation there is no opportunity cost, or one that is limited. In other words, large investors with small allocations to cash might routinely hold sufficient liquidity to be able to fund calls even in times of distress. The risk of not being able to meet a cash call is much greater for smaller investors. Similarly, pension fund investors with predictable cash flows are relatively well placed to predict their cash position and to have contingencies for capital calls. They hold a quasi-permanent call option over the contributions from fund members by way of monthly payroll, so even in crises their ability to raise cash is stable. In the real world, we would need to parse the investor base to determine which PE firm might have a particular vulnerability to its investors' collective cash position.

In our stylized simple example, if volatility rises, creating a higher value to the PE fund for the call option, then the investor will not see a consequent reduction in the value of its portfolio.

In Condition 1, the change in value is not negative for the investor. As the call rises in value for the PE firm, the investor does not necessarily lose value. An increase in volatility, as during a period of crisis, does not cost the investor, but it does create a positive value change for the PE firm that should ultimately benefit the investor (assuming the PE firm can take advantage of the crisis). Investment funds that meet Condition 1 logically have gains to trade by seeking investments in PE funds.

Condition 2

The investor maintains an investment in a beta of 1 investment. Here, we choose a beta of 1 for simplicity, but the logic carries for other equity-type choices, hedge funds and other alternative assets, provided the betas are known.

If volatility rises in the market due to a crisis, equity prices will tend to fall sharply. In this case, the value of the call to the PE firm rises, but the value of the assets the investor has to sell in order to fund the call falls accordingly. This is a fine example of a risk or liquidity cascade, where a supposedly safe asset must be sold in distressed circumstances, leading to further falls in asset values and a further reduction in liquidity and so on. Conventional risk management traditionally underestimates the impact of such cascades.

In extremis, this condition could mean that the increase in the value of the call option to the PE firm is reduced or negated by the impact on the ability of the investor to liquidate other assets in order to fund the call. The critical moment is when the investor proves unable to fund the call because of the compromising of its 1-Beta assets. Some investors will be able to fund the call using emergency capital or other measures (unanticipated sale of other assets/emergency borrowing against collateral – see below Condition 3). Let us assume for the moment, however, that it is funded, so as not to introduce a new variable in considering the option value.

The investor's loss (realized in forced liquidation during the cascading market) is thus the PE fund's gain. To the extent, however, that the crisis enables the PE fund to invest at better prices, thereby providing a better return to the investor due to the provision of liquidity during a liquidity-challenged period, then there may be some positive offset to the investor. In the end, the impact could be felt more through the carried interest agreement than through the declared returns on the fund. Even in this case, assuming funding occurs there may be gains to trade for the investor; the extent of the gains due to a risk offset is not easily discerned.

Condition 3

If the investor funds the call via borrowing with established lines of credit, then the offsetting value may be different depending on:

- The cost of creating the line of credit at the time of writing the call to the PE firm
- The interest paid on the credit line once the funds are called
- The possible alternative uses for the credit

Although in theory the capital-structure argument should not change the value of the negative call to the investor, it could do so depending upon the three issues above.

If the investor obtains a guarantee over the necessary credit lines during a period of low volatility, then an increase in the volatility of the market may mean that the call option it has purchased rises in value simultaneously, although not necessarily symmetrically, with the call it has written to the PE firm. Hence, a significant change in volatility may not imply a loss to the investor equivalent to the size of the gain to the PE firm.

If the investor has to borrow during a crisis because it had not secured its line of credit at the time of commitment, then the situation is entirely different. Now the investor must incur the cost of additional volatility in its purchase of a now more expensive credit line and this will at least partially, and possibly fully, offset any advantage it might derive from the gain in value of the PE firm's call option. This assumes that the PE firm is indeed able to find more attractive deals in which to deploy the contingent capital during distressed markets. If it does not, then the net result for the investor in this condition is even worse.

If the investor can deploy its available credit more efficiently, then there is an opportunity cost resulting from having written the call to the PE firm, assuming the capital call takes precedence over other opportunities for purchasing distressed assets and that the call will be met. Does the increase in the value of the call for the PE firm preclude the investor from using its credit to purchase Beta under duress? Is the investor in fact able to make the decision to divert otherwise committed funds to the PE firm towards other distressed assets? We can see that this will be highly idiosyncratic. A few large investors might have the flexibility to go beyond the capital call and make distressed asset purchases of their own – this includes taking advantage of the opportunity to "double down" on public equity markets. Many, however, will lack the optionality to do so.

Perhaps the true value of the capital call option is that it passes from the investor to the PE fund (firm) the opportunity to purchase distressed assets during crises and thereby tends to enhance fund returns, such that management fees and carried interest are "justified." This is because the eventual overall outperformance across a 10-year cycle, riding through any crisis or crises, is notable.

Summary of investor conditions

In Condition 1, it appears that the gain to the PE firm does not have a proportional offset loss to the investor and may not imply any loss in value to the investor if the allocation to cash or Treasuries is not affected by the investment decision.

Under Condition 2, the investor is, in most cases, losing, because the PE firm is gaining value with a crisis or a sharp increase in volatility (this was clearly the case during the recent 2007-09 crisis, for example). This may be mitigated to some extent by the richness of the opportunities created for the PE firm. From a risk perspective, however, the investor has written an insurance policy to the PE firm and bears the risk.

Under Condition 3, multiple outcomes are possible with respect to the size of the loss for the investor given the gain for the PE firm due to a rise in volatility. This is because the "insurance" against having written a call option to the PE firm contemporaneously takes the form of buying a call from another provider of credit. If the option to obtain credit is purchased in a low volatility environment then it is clear that such an arrangement will have value, which may offset the gain in value to the PE fund due to a rise in market volatility. If, however, secure funding to cover the call is obtained during the crisis, gains from trade may be lost due to increased funding costs directly incurred by the investor.

For each of these cases, however, over time the call's rise in value in a crisis will be offset to the extent that part of the funds have been called and replaced by the put that they now own, supporting the notion that the overall implied risk in the PE fund itself should not be managed in the same manner as general market risk. Although the embedded options in the PE fund process are complex, our simple analysis suggests that they may offset much of the downside risk normally associated with investing in equities. If this is the case then, depending on the conditions above, PE fund investing may be less risky to some extent than one would conventionally think. Even if the long-term returns are simply those of the market, then the PE fund would have "correctly measured" a much better information or Sharpe ratio than market investing, arguing for significant inclusion of such investments in diversified portfolios. And to the extent that the option characteristics are orthogonal (or low to negative correlation with general equity prices) such investments may decrease significantly portfolio risk at the margin.

SUMMARY REMARKS

We have assessed the complex optionality of the contract between a PE firm and its investors during the period of fund commitment and subsequent investments. The value of the capital call must be viewed as an asset for the PE (fund) firm. How the PE (fund) firm itself will change in value will depend upon the degree to which the call option has decayed with time or, more interestingly, with the possibility of a non-funding of part of the commitment, thereby creating the potential for vega risk associated with the specific PE fund. Across multiple funds within a single PE firm this is a non-trivial problem in the management of risk for the PE firm itself.

The positive value of the call option in a traded market would, in theory, be of equal portion negative for the writer of the option under normal circumstances if the markets to trade them were fully developed. We have shown that in the PE world it seems clear that risks and values are asymmetrical. For the investor writing the option, the trade-off appears to depend on how the call is funded. It seems clear that if the call is funded by the investor merely as a way of increasing its exposure to equity markets (for example, maintaining a beta exposure while simultaneously writing the call), then an increase in risk (volatility) will hurt the writer via losses, potentially severe losses, on the beta portion of its investment during a crisis. The increased volatility should be reflected in the overall increase in the risk of the investor's portfolio, suggesting an opportunity for more efficiency in the market to develop. If this is indeed the case then a greater development of the PE fund market would help complete the proper pricing of these embedded options vis-a-vis the portfolio allocation to PE funds by investors relative to typical asset allocations suggested by naïve models.[19]

	Zero % called	Half called	Fully called
PE fund firm call	Highly valuable	Valuable	No value
PE fund put	Zero value	Valuable	Highly valuable
Investor call condition 1	Neutral	Neutral	No value
Investor call condition 2	Negative in value	Negative in value	No value
Investor call condition 3	Neutral to negative in value	Neutral to negative in value	No value
Investor put condition 1	Potentially valuable	Valuable	Highly valuable
Investor put condition 2	Potentially valuable	Valuable	Highly valuable
Investor put condition 3	Potentially valuable	Valuable	Highly valuable

Table 4 – Assume a crisis

Under the other funding choices for investors, risk levels are not so easily understood. If the call is funded by cash equivalents or highly liquid government bonds, then an increase in market volatility may benefit the investor, at least with respect to this marginal risk portion of the portfolio. Whereas the value of the call option to the PE firm rises and therefore its risk falls, the risk savings to the PE firm do not come at the expense of the investor. Credit funding may or may not have a similar effect, depending on the planned capital structure of the investor's portfolio and leverage. It is all about the extent to which the relevant parties have thought ahead and created risk-tolerant positions.

From the perspective of the PE firm holding the capital call option, ceteris paribus, increased volatility (or a crisis) implies an increase in the value of this key asset of the fund and thus the firm. Gamma risk is positive. Only if there is a cap on the ability to call funds from investors should one consider the vega risk associated with the option. This is a disaster risk a PE fund may consider and treat any vega risk accordingly.

From the perspective of this narrow measure of risk for the PE firm itself, owning the call may simply be an offset to rising risk in times of crisis or high-volatility markets. And once the put provisions of the contracts are better assessed it becomes less and less clear when the investor is hurt by a crisis. No doubt the Condition 2 investor has more at risk; perhaps this was the real world point faced by the investors noted above in footnote 7. In other cases, however, just as the existence of the options smooth out the implied risk for the PE fund (firm) during a crisis, it may also be the case for the investor under certain conditions. Table 4 lays out the crisis situation and the implications for the crisis at various points in the life of a fund with respect to the imbedded options in the fund.

CONCLUSIONS

Recognition by investors and managers of the implicit options in private-equity investing is essential to determining the risk associated with those investments, particularly as it relates to general market risk. A simple approach to breaking out the options implicit in the contracts entered into by PE fund (firm) managers and investors

19 This observation was highlighted thanks to Charles Smithson who pointed out that swaptions "completed" the corporate callable bond market where calls were typically underpriced before the development of the options market.

yields interesting insights into each group's risk profiles. It would be a significant mistake of risk analysis to ignore the implicit options embedded in the PE fund. For the risk manager or any interested party (such as an investor into publicly listed PE firms) recognition of the role these options have in mitigating the market risk of the fund is crucial.

Although we have taken a simplistic approach to analyzing the options themselves, this does not negate the implications that a VaR, or any similar market-based, approach to risk valuation of a PE fund (firm) is missing many of the most important factors. In spite of the widely held beliefs that many "smart" investors were damaged by their reliance on PE fund investing during the crisis, this need not at all be the case for most investors. We can see why Condition 2-type investors would be damaged by being over leveraged. For other investors, it is not clear, or in some cases quite the opposite; the crisis makes these investment vehicles risk-reducing on the margin when considered in a portfolio context.

Moreover, given the conclusion that the existence of these options naturally mitigates risk, more efficient investment portfolios may be created by including PE fund allocations. The "correct" proportion of an allocation is dependent on many factors (some noted above). However, if PE fund risk is much lower than is priced by the market in part because there is poor understanding of the options involved, then systematic under-allocation to PE funds (and under-pricing of PE firms' equity) is likely.

Although we have used snapshots of period volatility and assumed time periods to measure the value of both the calls and later the puts owned by the PE fund and ultimately the PE firm, these snapshots highlight the need for an options-theoretic approach to measuring risk for the PE fund (firm). A next step would be to use a typical time horizon for a fund, a typical period to call the funds, a typical period to place the funds and a typical period for returning investments and earned returns to the client, as well as reinvestment along the way. This requires considerably more sophisticated analysis, but is potentially feasible. One could create a rolling valuation of the implied options, taking time snapshots over the life of the fund. Using a simplistic backward looking volatility pattern would be a first step to seeing how the values of the options unfold over time.

Further, crisis modeling can be utilized. We chose the recent financial crisis, but others, such as the LTCM crisis, or Asian Contagion, or the 9/11 shock, could also be employed. In all cases, the ownership of options by the PE firm may significantly mitigate the shock and the downward implications of such tail events. Any risk management system for the PE fund or firm should include the consideration of fully pricing these options in the event of a shock. This has implications for PE firm risk managers who must consider stress testing as part of their regulatory requirements.

One must also note that the typical PE firm has a portfolio of funds. As such, the long-dated options it holds are also a portfolio. A portfolio of options will likely be less valuable than simply adding up the value of the individual options themselves, an effect that is another avenue for future research.

Finally, we want to stress the difference in risk measurement and risk management. Implementation of a strategy to capture the value of the options that are owned by the PE firm goes beyond knowing that they have an offsetting risk capability during a crisis. The use of OTC or exchange-traded baskets of options may provide a way to capture this value during the heat of a crisis. Depth of markets, types of baskets that may be created, and the ability to manage effectively the vega and theta risk of the options sold to capture the value embedded in a volatility spike are issues that need addressing if risk measurement is to be turned into truly effective risk management.

We expect that further study will uncover academically sound analytics to support the simple notion that a crisis is indeed the "best friend" of the PE fund.

REFERENCES

- Chicago Board Options Exchange, "VIX (CBOE market volatility)," http://www.cboe.com/delayedquote/advchart.aspx?ticker=VIX.
- Black, F., and M. Scholes, 1973, "The pricing of options and corporate liabilities," Journal of Political Economy 81:3, 637-654
- Jensen, M., and W. H. Meckling, 1976, "Theory of the firm: managerial behavior, agency costs and ownership structure," Journal of Financial Economics 3:4, 305-360
- Karagiannidis, I., and D. Sykes Wilford, 2015, "Modeling fund and portfolio risk: a bi-modal approach to analyzing risk in turbulent markets," Review of Financial Economics 25:C, 19-26
- Markowitz, H. M., 1952, "Portfolio Selection," The Journal of Finance 7:1, 77-91
- Merton, R., 1973, "Theory of rational options pricing," Bell Journal of Economics and Management Science 4, 141-183
- Merton, R., 1974, "On the pricing of corporate debt: the risk structure of interest rates," Journal of Finance 29, 449-470
- Phalippou, L., and M. M. Westerfield, 2014, "Capital commitment and illiquidity risks in private equity," working paper
- Siegel, L. B., 2008, "Alternatives and liquidity: will spending and capital calls eat your modern portfolio?" Journal of Portfolio Management 35:1, 103-114
- Smithson, C., W., 1998, Managing financial risk: a guide to derivative products, financial engineering and value maximization, McGraw-Hill Professional
- Sorensen, M., N. Wang, and J. Yang, 2013, "Valuing private equity," Review of Financial Studies 27:7, 1977-2021
- Wibaut, S., and D. S. Wilford, 2009, "Holding equity and debt of the same firms can prove suboptimal," Journal of Applied Finance 19:1&2, 19-27
- Wilford, D. S., 2014, "Risk management insights from Markowitz optimization for constructing portfolios with commodity futures," Journal of Financial Perspectives 2:2, 47-55

Credit Risk Decomposition for Asset Allocation

Álvaro Mª Chamizo Cana – Executive Director, Credit Risk Portfolio Management, BBVA[1]
Alfonso Novales Cinca – Professor, Department of Quantitative Economics, Universidad Complutense

Abstract
We provide a methodology for credit risk analysis that can be embedded into a risk appetite framework. We analyze the information content in CDS spreads to estimate the systematic and idiosyncratic components of credit risk for CDS issuers in the industrial sector of Europe. Such decomposition should be an important tool for the evaluation of the diversification possibilities of credit portfolios or for the design of appropriate hedging strategies. It could be used by financial institutions to maintain their risk limits when taking their asset allocation decisions as well as by supervisors investigating potential systematic risk problems. The analysis could be extended to other sectors.

1 The authors acknowledge comments received from J.A. De Juan Herrero. Financial support from Spanish Ministry of Education through grant ECO2015-67305-P, the PROMETEO 2013 Program of Comunidad Valenciana and from Banco de España through a Programa de Ayudas a la Investigacion grant is gratefully acknowledged. This article reflects the opinions of the authors, and not those of BBVA.

INTRODUCTION

The most widely used measures of credit risk use information on CDS spreads, which are forward-looking and reflect the market perception of the credit risk of the issuer. A firm with a large idiosyncratic component of credit risk could default with a minor impact on its sector or the economy, while the opposite will happen for a firm that has important sectorial or systematic components of risk. If the systematic component is important, the behavior of its CDS will tend to follow that of the market, leaving few possibilities for hedging a credit position on that firm. Hence, from the point of view of implementing the risk policy at a given financial institution, as well as evaluating the possibilities for hedging a credit portfolio, estimating the relevance of the systematic, sectoria,l and idiosyncratic components of risk for a given creditor is critical.

We propose a simple methodology for the estimation of these different components of credit risk. We use the information provided by a wide set of financial indicators to decompose the credit risk of each firm into systematic, sectorial, and idiosyncratic components. Such decomposition should be central to evaluating which firms have more potential to produce systematic risk problems. This information would clearly be essential for the policymakers responsible for supervision and regulation. It will also be extremely useful for companies, investors, hedgers, and speculators who are involved in the credit markets and in the pricing of credit, since it provides some insights on the possibilities of hedging the credit risk of a given position. Our analysis is based on the degree of commonality among CDS spreads across sectors, as well as on the correlation among CDS spreads of firms operating in a given sector. A principal component analysis of the mentioned set of financial indicators is used to characterize the systematic components of credit risk, while a principal component analysis of CDS spreads across firms in a given sector is used to characterize the sectorial component of credit risk in those firms. The idiosyncratic component is what is left after estimation of the systematic and sectorial components of credit risk.

An alternative methodology using the first component of sectorial indices of CDS spreads to identify the systematic component of credit risk yields a very similar decomposition. Further research should examine the relationship between our estimated risk components and certain characteristics of firms, such as the size of assets and liabilities, profit and loss, equity and bond prices, and market share. That would allow us to extend the risk evaluation results obtained in this paper for CDS issuers to any other firm.

LITERATURE REVIEW

Given the importance of the topic for researchers and for market regulators after the financial crisis, the recent literature on measuring systematic risk has been quite extensive. We briefly review in this section those we consider most relevant for our work.

Ericsson et al. (2009) analyzed the relationship between theoretical determinants of default risk, such as firm leverage, volatility, risk-free interest rate, and actual market premium, using the CDS on senior debt for the period 1999-2002. Using time series regressions, they found that these variables explain approximately 60% of the variations of CDS premia, while the R-squared for changes in default swap premia is approximately 23%. Tang and Yan (2013) used transactions data from 2002 to 2009, covering 861 North American corporates, to find that CDS spreads are mostly driven by fundamental variables such as firm volatility and leverage, market conditions, and investor risk aversion. Hence, even if actual default risk remains constant, CDS spreads may increase when investors become more pessimistic and more risk averse. A 1% increase in the VIX index, interpreted as a measure of market sentiment or investor risk aversion, is shown to be associated with about 1% increase in CDS spreads.

Some studies have used synthetic risk indicators, illustrating the existence of a strong degree of commonality in credit risk. Rodríguez-Moreno and Pena (2013) analyzed two groups of systematic risk measures when searching for the best systematic indicator over the January 2004-November 2009 sample period. The first group contained indicators related to the overall tension in the market, while the second group was made up of indicators related to the contributions of individual institutions to systematic risk. In a sample of 20 European banks and 13 U.S. banks they found that the first principal component of CDS spreads performed better than measures of market stress. For a sample of 150 European firms from January 2003 to July 2007, Berndt and Obreja (2010) showed that the first principal component of CDS returns explained less than 30% of the variation in weekly CDS returns, but that fraction surged to 50% during the crisis, from August 2007 to December 2008. The shift in the correlation structure of European equity returns was more modest when compared to CDS returns.

Bhansali et al. (2008) used a three-jump model to carry out a decomposition of CDS spreads among systematic risk, sector risk, and idiosyncratic risk as we attempt to do in this paper, although their methodological approach is quite different.

DECOMPOSITION OF RISK IN SPECIFIC SECTORS: SYSTEMATIC, SECTORIAL, AND IDIOSYNCRATIC RISKS

For asset allocation purposes, it is central to have some knowledge of the nature of risk involved in a given credit position. We aim to measure the degree to which firms in a given sector are subject to systematic, as well as to sectorial, risk and determine the relevance of idiosyncratic risk.

We consider systematic risk events as those that have an influence across the global credit markets. Consequently, our approach to decompose risk is based on the use of a set of financial factors, which we split into two groups. The initial set of seven credit market indicators include: Markit iTraxx Europe Index, Markit iTraxx Europe HiVol Index, Markit CDX North American Investment Grade Index, Markit CDX North American Investment Grade Index High Yield, 3-month ATM iTraxx Europe Index Option, 3-month ATM CDX North American Investment Grade Index Option, iTraxx Japan IG. A second set of 21 indicators include: the 3-month EURIBOR interest rate, the 3-month EONIA Index, the Euro liquidity premium, measured by the absolute difference between 3-month EURIBOR and 3-month EONIA, both in euros, the 1-, 5- and 10-year Euro swap rates, the 3-month/ 5-year ATM Euro swaption, the VSTOXX index, the 5-year German government yield, the 3-month USD LIBOR interest rate, the 3-month USD overnight index, the USD liquidity premium, measured by the absolute difference between 3-month LIBOR and the 3-month USD overnight index, the 1-, 5- and 10-year USD swap rate, the 3-month/5-year ATM USD swaption rate, the VIX index, the 5-year US Treasury Rate, the EUR/USD FX spot rate, the EUR/USD 3-month ATM option, the 5-year JPY swap rate. Data were obtained from Bloomberg.[2]

We determine common risk factors among CDS spreads from the different sectors using the principal component methodology to the covariance matrix of weekly returns on CDSs. Two principal components of the subset of credit market variables and three principal components of the subset of other financial indicators explain more than 98% of the fluctuations in each group of indicators.

Sectorial risk events are those that impact all of the firms in a given sector, with no major implications elsewhere. For a given sector, a principal component will contain some features common across firms in the sector, possibly combined with some elements of systematic risk.

European industrial sector

Our sample contains CDS spreads for 30 issuers in the European industrial sector, with daily quoted prices for the 2006-2012 period. There is important commonality among the time evolution of these spreads, but there are also significant risk components that are specific to each issuer in the sector. The first principal component of the time series for the 30 CDS spreads is an approximate average of CDS prices across the sector, with all the firms entering with a similar load in its definition. It has a linear correlation coefficient with the iTraxx Index of 0.72, and it explains 64.4% of the joint fluctuation in the set of spreads. We would need to consider at least six principal components to explain more than 80% of the volatility in the vector of CDS prices. Firms like Rentokil Initial 1927 Plc, Heidelberg Cement AG, Invensys plc, Alstom, and Siemens AG have a significant presence in defining the successive principal components. Hence, the first intrasector principal component can be safely used as an indicator of sectorial risk, since most of what it is unable to explain is due to idiosyncratic risk that is captured by further principal components, which we do not use.

Column 2 in Table 1 shows the adjusted R-squared from a regression of CDS spreads on the first two principal components of credit indicators plus the first three principal components of non-credit indicators. These R-squared statistics, between 0.22 and 0.57, can be interpreted as an estimate of the size of the systematic risk component for each firm. They are very close to the R-squared statistics obtained by explaining CDS spreads with just the two first principal components obtained from credit market indicators.[3] To be conservative, we have chosen to maintain the two sets of principal components in these projections to obtain the R-squared values shown in column 2.

To estimate a sectorial component of risk, we use the first principal component of CDS spreads for the 30 issuers as a sectorial credit risk factor. Column 3 shows the R-squared statistics from regressions of CDS changes on this risk factor, with values of between 0.29 and 0.79. They show a higher explanatory power than the credit and financial risk indicators taken together. Furthermore, when we put together all these factors in the regressions in column 4, the explanatory power is barely higher than the one obtained by the sectorial risk factor alone. Obviously, this factor reflects some sectorial implications, besides capturing some influences from the global credit markets. To segregate the implications of each component, we take in column 5 the difference between the numerical R-squared values of columns 4 and 2 as an estimate of the relevance of sectorial risk. Finally, what remains unexplained by the regression on credit and financial risk factors and the sectorial

2 The 3-month ATM iTraxx Europe Index Option and the 3-month ATM CDX North American Investment Grade Index Option are provided by JP Morgan.

3 The R-squared from regressions of CDS spreads on the credit indicators fall between 0.21 and 0.57, while the R-squared from regressions on the rest of the financial indicators are lower, between 0.05 and 0.27. For reasons of space, these regressions are not shown in the table.

(1) Issuer	(2) Systematic risk	(3) Sectorial PC	(4) Joint regression	(5) Sectorial risk	(6) Idiosyncratic risk
AB Volvo	**57.20%**	73.70%	74.20%	17.00%	25.80%
Cie de St Gobain	**56.90%**	78.40%	78.80%	21.90%	21.20%
Holcim Ltd	**56.70%**	79.30%	79.80%	23.10%	20.20%
Rolls-Royce Plc	**54.90%**	71.00%	73.80%	18.90%	26.20%
Lafarge	**54.70%**	79.10%	79.80%	25.10%	20.20%
Scania Ab	**54.60%**	70.80%	71.40%	16.80%	28.60%
Thales	**52.20%**	77.90%	80.00%	27.80%	20.00%
Finmeccanica S.p.A	**51.70%**	66.50%	68.30%	16.60%	31.70%
Vinci	**51.50%**	73.90%	74.50%	22.90%	25.50%
Volvo Treas AB	**51.00%**	69.20%	70.20%	19.20%	29.80%
Adecco S A	**48.40%**	68.60%	69.10%	20.70%	30.90%
BAE Systems PLC	**48.00%**	71.80%	72.10%	24.10%	27.90%
Deutsche Lufthansa AG	**47.20%**	66.00%	65.80%	18.70%	34.20%
Deutsche Post AG	**44.60%**	58.90%	59.80%	15.20%	40.20%
Eurpopean Aero Defence & Space Co Eads N V	**44.50%**	70.30%	70.80%	26.30%	29.20%
Rexam plc	**44.20%**	67.10%	67.00%	22.80%	33.00%
Metso Corp	**43.40%**	62.10%	63.00%	19.60%	37.00%
HeidelbergCement AG	**42.90%**	58.40%	59.90%	17.00%	40.10%
Societe Air France	**42.10%**	63.80%	63.80%	21.70%	36.30%
Assa Abloy Ab	**41.20%**	62.90%	63.00%	21.80%	37.00%
Alstom	**40.70%**	62.30%	62.10%	21.50%	37.90%
Securitas AB	40.60%	57.40%	59.20%	18.60%	**40.80%**
Siemens AG	39.80%	57.50%	60.10%	20.30%	**39.90%**
Atlas Copco AB	39.30%	59.20%	58.90%	19.60%	**41.10%**
Brit Airways plc	36.40%	53.40%	55.10%	18.70%	**44.90%**
Schneider Electric SA	36.40%	55.80%	55.70%	19.40%	**44.30%**
Smiths Group Plc	30.20%	51.00%	51.40%	21.20%	**48.60%**
Ab Skf	27.90%	45.20%	45.80%	17.90%	**54.20%**
Rentokil Initial 1927 Plc	23.00%	29.30%	30.90%	7.80%	**69.10%**
Invensys plc	21.80%	37.80%	39.30%	17.50%	**60.80%**

Note: Column1 shows the company name from Markit database. Column 2 displays the adjusted R-squared from a regression on two first principal components of credit indicators and the three first principal components of non-credit financial indicators. Column 3 shows the adjusted R-squared from a regression on the first principal component of the European industrial CDS spreads in the sample. Column 4 shows the adjusted R-squared from a regression on the set of explanatory variables in the two previous regressions. Column 5 displays the difference between the numerical R-squared values in columns 4 and 2 as an estimate of the relevance of sectorial risk. Finally, column 6 displays the size of idiosyncratic risk, computed as 1 minus the adjusted R-squared in column 4. Bold figures indicate the most important factor in the risk decomposition for each CDS issuer. All regressions are estimated in weekly changes of the mentioned variables.

Table 1 – European industrial issuer CDS spread decomposition

factor can be naturally interpreted as the size of the idiosyncratic component of risk. This way, we have a decomposition of CDS risk in systematic risk (column 2), sector-specific risk (column 5) and firm-specific risk (column 6), adding up to +100%. Firms in Table 1 are ranked by the size of their systematic components of risk. Firms with a high idiosyncratic component of risk should be preferred by financial institutions, since they offer better prospects for build a well-diversified credit portfolio. On the other hand, it would also be unwise to take a credit position in a few firms with large idiosyncratic risk components.

Sectorial risk oscillates between 8% and 28%, while the idiosyncratic component of risk ranges from 20% and 69%. In most issuers (27 out of 30), the idiosyncratic component is below 50% of total risk. Bold figures in the table denote the most important component for each single issuer. For 21 of the 30 issuers, global risk factors are

the most important components of CDS risk, while firm-specific factors are the most important component for the other nine issuers. In our estimates, sector-specific components were never the most important source of fluctuations. Using median values, the systematic component of risk for the European industrial sector is 44% of total risk, sectorial risk is 20%, and the idiosyncratic component amounts to 35%.[4]

An alternative decomposition of risk

To develop an alternative method of decomposing risk, we initially select a set of 5-year CDSs trading as senior unsecured debt, SNRFOR, with 1825 daily observations on approximately 2500 issuers, from the eleven industries and the thirteen geographical areas.[5]

We then construct CDS indices for each sector by taking the median CDS spread traded each day in that sector across all regions. To reduce the possibility of excessive noise due to low trading, we aggregate over time, taking weekly averages of sectorial indices. Finally, we compute logarithmic changes of weekly CDS spreads, obtaining a total of 365 weekly observations for each sector index over the 2006-2012 period. We used this data to characterize common risk factors among CDS spreads from the different sectors using the principal component methodology. The first principal component, by itself, explains 68% of the fluctuations in the set of eleven sectorial indices, indicating that there is strong commonality among the sectors. This is a higher percentage than the one estimated by Berndt and Obreja (2010) for European firms during the 2003 to 2008 period, but it is very close to the average explanatory power estimated by Chen and Härdle (2012) for the pre- (58.7%) and post-crisis periods (72.3%).

Since the first principal component explains more than two thirds of the fluctuations in the whole set of CDS issues from all sectors and geographical areas, it can naturally be interpreted as representing a global risk factor, capturing the systematic elements of risk. Hence, an alternative decomposition to the one we used earlier would estimate the relevance of systematic risk by the adjusted R-squared of CDS spreads for each firm on the global risk factor. The first intra-sector principal component adds some sector-specific information to the global risk factor, and we take the difference between their joint explanatory powers and that of the global risk factor alone as an estimate of the relevance of sectorial risk. The residual in that joint regression is an estimate of the idiosyncratic component of risk; its relevance being estimated as 1 minus the R-squared in such regression.

Surprisingly, estimates of risk components by both procedures are quite similar. The rank correlation coefficient between the estimated relevance of systematic risk by both approaches is 0.78, with a linear correlation coefficient of 0.86. The similarity between the estimated relevance of idiosyncratic components is still higher, with a rank correlation coefficient and a linear correlation coefficient of 0.99.

An analysis of the estimated idiosyncratic components of credit risk[6]

The estimated idiosyncratic component of CDS risk turns out to be quite large in many firms, which might be due to the fact that our estimated idiosyncratic component could still contain some systematic risk elements. To test for the effectiveness of our methodology we examine whether our estimates of the idiosyncratic component of credit risk have the appropriate features.

A first test consists of examining the possibility of diversification. If idiosyncratic components are relatively important, then a well-diversified portfolio should be much easier to hedge than positions on individual assets. In the European industrial sector, hedging positions on CDS from an individual firm using a contrary position on iTraxx leads to a significant decline in variance,[7] with a median reduction of 14.1%. On the other hand, for the equally weighted portfolio we would achieve a reduction in variance of 30.0%. The fact that the hedge is much more successful for the equally weighted portfolio than hedging a position in any single firm in the sector suggests that idiosyncratic components of credit risk are indeed important.

A second test considers whether the hedging possibilities increase with the size of the idiosyncratic component of risk. This is clearly the case: the reduction in variance from hedging the portfolios made up of the 5 or 10 firms with the highest idiosyncratic components of risk is of 62% and 65%, respectively, while the reduction in variance from hedging a portfolio of the 5 or 10 firms with the lowest idiosyncratic components of risk is 43% and 54%, respectively. Hence, hedging efficiency is clearly higher for portfolios made up of firms with high idiosyncratic risk. Among portfolios with low idiosyncratic risk, a sufficient hedging efficiency would require considering portfolios made up of a larger number of firms.

4 Being median values they may not add up exactly to 100%.

5 We use Markit industry levels, which considers eleven industries: basic materials, consumer goods, consumer services, energy, financials, health care, industrials, technology, telecommunication services, utilities, and government. Government is another category considered by Markit but not included in the Industry Classification Benchmark. Finally, Markit considers thirteen different regions: Africa, Asia, Caribbean, Eastern Europe, Europe, India, Latin America, Middle East, North America, Oceania, Offshore, Pacific, and Supranational.

6 Since both approaches lead to similar decompositions of credit risk, we just interpret the results obtained with the use of 28 financial indicators to estimate the systematic component of credit risk.

7 We consider a least-squares hedge, with the hedge ratio being the negative of the estimated slope in a regression of the CDS spread for a given issuer on the iTraxx index.

(1) Issuer	(2) Systematic risk	(3) Sectorial PC	(4) Joint regression	(5) Sectorial risk	(6) Idiosyncratic risk
AB Volvo	**59.60%**	73.70%	73.70%	14.10%	26.30%
Cie de St Gobain	**66.00%**	78.40%	78.30%	12.30%	21.70%
Holcim Ltd	**65.20%**	79.30%	79.20%	14.10%	20.80%
Rolls-Royce Plc	**52.60%**	71.00%	72.10%	19.50%	27.90%
Lafarge	**67.70%**	79.10%	79.10%	11.40%	20.90%
Scania Ab	**60.70%**	70.80%	70.80%	10.10%	29.30%
THALES	**62.90%**	77.90%	78.00%	15.00%	22.10%
Finmeccanica S.p.A	**52.60%**	66.50%	66.70%	14.10%	33.30%
Vinci	**59.50%**	73.90%	74.00%	14.50%	26.00%
Volvo Treas AB	**58.70%**	69.20%	69.10%	10.40%	30.90%
Adecco S A	**58.90%**	68.60%	68.60%	9.70%	31.40%
BAE Systems PLC	**57.00%**	71.80%	71.90%	14.90%	28.10%
Deutsche Lufthansa AG	**53.30%**	66.00%	66.00%	12.70%	34.00%
Deutsche Post AG	**48.70%**	58.90%	58.80%	10.10%	41.20%
European Aero Defence & Space Co Eads N V	**57.10%**	70.30%	70.30%	13.20%	29.70%
Rexam plc	**58.70%**	67.10%	67.20%	8.50%	32.80%
Metso Corp	**56.50%**	62.10%	62.60%	6.20%	37.40%
HeidelbergCement AG	**46.70%**	58.40%	58.40%	11.70%	41.60%
Societe Air France	**53.80%**	63.80%	63.70%	9.80%	36.30%
Assa Abloy Ab	**59.20%**	62.90%	64.00%	4.80%	36.00%
Alstom	**51.60%**	62.30%	62.20%	10.60%	37.80%
Securitas AB	**49.00%**	57.40%	57.30%	8.40%	42.70%
Siemens AG	**53.50%**	57.50%	58.30%	4.80%	41.70%
Atlas Copco AB	**58.20%**	59.20%	61.30%	3.10%	38.70%
British Airways plc	**47.60%**	53.40%	53.60%	5.90%	46.40%
Schneider Electric SA	**50.00%**	55.80%	56.10%	6.10%	43.90%
Smiths Group Plc	41.30%	51.00%	51.00%	9.70%	**49.00%**
Ab Skf	45.30%	45.20%	47.20%	1.80%	**52.80%**
Rentokil Initial 1927 Plc	24.10%	29.30%	29.10%	5.10%	**70.90%**
Invensys plc	29.10%	37.80%	37.90%	8.80%	**62.10%**

Note: Column1 shows the company name from Markit database. Column 2 displays the adjusted R-squared from a regression on the global risk factor, which is estimated as the first principal component of sectorial CDS indices. Column 3 shows the adjusted R-squared from a regression on the first principal component of the European industrial CDS spreads in the sample. Column 4 shows the adjusted R-squared from a regression on the explanatory variables in the two previous regressions. Column 5 displays the difference between the numerical R-squared values in columns 4 and 2 as an estimate of the relevance of sectorial risk. Finally, column 6 displays the size of idiosyncratic risk, computed as 1 minus the adjusted R-squared in column 4. Bold figures indicate the most important factors in the risk decomposition for each CDS issuer. All regressions are estimated using weekly changes of the mentioned variables.

Table 2 – European industrial issuer CDS spread decomposition using GRF as the systematic explanatory variable

The last test is based on the fact that the estimated idiosyncratic components turn out to be essentially uncorrelated across firms, which is a necessary condition for the interpretation we give to this component. There are 30 issuers in the European industrial sector, implying 435 pairwise correlations between idiosyncratic components, with a low median correlation of -0.05. Ninety percent of them are in absolute value below 0.23. These are all low levels that justify an interpretation of our estimated idiosyncratic components as being firm-specific in nature.

Taken together, the possibilities for hedging the risk of a well-diversified sectorial portfolio, the higher efficiency in hedging portfolios made up of firms with the highest idiosyncratic components of risk, and the low pairwise correlations across firms in the European

industrial sector, suggest that our estimates of such components are appropriate.

However, a question remains: what is causing the large idiosyncratic component of risk? A possible conjecture would be that the large idiosyncratic components of risk could be just a reflection of the low liquidity in some issues. To examine the validity of this assumption, we could relate the size of the estimated idiosyncratic risk component with either the number of contributors giving price to the 5-year CDS, the quality rating of the data provided by Markit, or the volatility of CDS spreads. In the latter case, the argument would be that illiquid CDSs would often repeat price in the Markit database, the time series of CDS spreads then having a relatively low variance. Hence, we would expect a negative correlation between the size of the idiosyncratic component of risk and the volatility of CDS spreads. That correlation between the size of the idiosyncratic risk component and the annual volatility of CDS weekly changes among European industrial issuers is equal to -0.30. Hence, the large size of the idiosyncratic risk component for some issuers could in part be due to the low liquidity of their CDS spreads.

CONCLUSIONS

A central component of a risk appetite framework at financial institutions would be a mechanism to decompose asset risk into systematic, sectorial and idiosyncratic components. We use a large set of 28 credit and non-credit financial indicators to estimate the systematic component of credit risk. A regression model to explain CDS spreads on five principal components summarizing the commonality in these indicators provides an estimate of the market perception of systematic risk for each firm. Next, we use a principal component of CDS spreads across firms in the sector to estimate the relevance of the sectorial component of credit risk. The idiosyncratic component of risk is the remaining CDS spreads for a given firm after extracting the systematic and sectorial components of risk. An alternative decomposition using the first principal component for sectorial CDS indices to estimate systematic risk yields a similar decomposition of credit risk.

This evaluation of the relevance of risk components has obvious implications for the asset allocation strategy by a given financial institution that wants to diversify its credit portfolio in that sector. When designing their credit policy, financial institutions should avoid firms with a large systematic risk component in favor of those with larger idiosyncratic risk components, always trying to form sufficiently diversified portfolios, thereby maintaining their risk limits when taking their asset allocation decisions.

We have provided some evidence that the estimated idiosyncratic components are due in part to lack of liquidity. We have also shown evidence suggesting that portfolios made up of firms with higher idiosyncratic components are easier to hedge, contrary to what happens with portfolios made up of firms with lower idiosyncratic risk components. By and large, the estimated idiosyncratic risk components turn out to be uncorrelated across firms in the sector.

By evaluating the firms with the most potential to produce systematic risk problems, our analysis should also be considered to be crucial for supervisors and regulators. Even though we restrict our analysis to CDS issuers, further research should attempt to relate our estimated risk components to firms' characteristics such as size of assets and liabilities, profit and loss, the leverage ratio, the EBITDA, bond prices, market share, or the market value of equity. This is an open question that would allow for extending the evaluation of credit risk components for CDS issuers to any other firm, even if it is not a CDS issuer. A further issue would consider the dynamics of defaults, analyzing how the stand alone default of a given issuer might affect other companies in its sector. Characterizing the interconnection between CDS issuers [as in Kanno (2016)] would provide us with information to identify the firms that play a central role in their network, thereby allowing for a more efficient coverage of credit risks at financial institutions.

REFERENCES

- Berndt, A., and I. Obreja, 2010, "Decomposing European CDS returns," Review of Finance 14:2, 189-233
- Bhansali, V., R, Gingrich, and F. A. Longstaff, 2008, "Systemic credit risk: what is the market telling us?" Financial Analysts Journal 64:4,16-24
- Chen, Y.-H., and W. K., Härdle, 2012."Common factors in credit defaults swaps markets," Technical Report, SFB 649 Discussion Paper
- Ericsson, J., K. Jacobs, and R. Oviedo, 2009, "The determinants of credit default swap premia," Journal of Financial and Quantitative Analysis 44:1, 109–132
- Kanno, M., 2016, "Interconnectedness and systemic risk in the US CDS market," Available at SSRN 2711112
- Rodríguez-Moreno, M., and J. L. Peña, 2013, "Systemic risk measures: the simpler the better?" Journal of Banking & Finance 37:6, 1817–1831
- Tang, D. Y., and H. Yan, 2013, "What moves CDS spreads?" Available at SSRN 1786354

Time to Rethink the "Sophisticated Investor"

Peter Morris – Independent Researcher[1]

Abstract

Policymakers need to change the way they think about so-called "sophisticated investors." The way they think about these organizations now disenfranchises the millions of ordinary people these big investors represent and makes it literally impossible to hold such big investors to account. This creates a dangerous flaw at the heart of the way financial markets are organized. This is not just abstract musing: it is demonstrably leading to poor outcomes for the ordinary people who depend on big investors. The good news is that policymakers can make a difference by applying a simple principle to "sophisticated investors": accountability. It need not cost a lot or involve a lot of bureaucracy. They must demand that big investors, and the fund managers they hire, disclose more to the public. What they disclose must allow (truly) independent outsiders to analyze how well the big investors have performed, including how cost-effective they are. Anyone who believes in markets knows that harnessing people's self-interest helps to make markets work. If policymakers choose to enfranchise the rest of society, they will be doing just that. Vested interests – including much of the financial services sector, many big investors, and even some policymakers – will call this idea outlandish. Some will portray it as an attack on financial markets. Even observers with no vested interest may worry that it will damage markets or the economy or both. Nothing could be further from the truth. An 80-year-old parallel shows the way. This change in approach would help to ensure that financial markets serve society as a whole, rather than just the people who work in them.

1 The author worked as a credit analyst for twenty-five years, most recently at Goldman Sachs and Morgan Stanley. Since 2009 he has been a consultant working in various areas including social investment. His independent research on private equity has appeared in academic and other publications, as well as at www.ssrn.com. He can be contacted at morrisp1@aol.com.

THE OFFICIAL STORY

The consensus view of "sophisticated investors" goes like this.[2] Unlike retail investors, big institutions like banks, insurance companies, and pension funds can look after themselves.[3] For that reason, policymakers give these organizations a free hand when it comes to investing (subject to any other rules they have to follow).

These organizations control most of society's wealth. That means the rest of society needs them to make good investments. The reason policymakers give these organizations so much freedom is that they assume they will in fact make good investments. Over the last thirty years, "sophisticated investors" have put an increasing proportion of their portfolios into so-called "alternative investments": hedge funds, private equity, commodities, and the like.[4] According to the official story, that just shows by definition that these must be good investments; hence there is no need for policymakers (or anyone else) to check if that is true.

Suppose outsiders want to check if these really have been good investments. Policymakers make it hard for them to do so. That is because policymakers give "sophisticated investors" a second big break: they allow these organizations to operate in private. Some big investors disclose headline results to the public. But they do not have to supply data that would allow outsiders to review their performance in detail. This applies even to their own end-investors or beneficiaries. Take an individual member of a big pension fund, or a taxpayer who contributes to it, as an example.[5] They are unable to obtain data that would allow them, or an expert they employ, to assess how well either the pension fund's managers, or any firms they in turn hire, are performing on their behalf.[6]

"Sophisticated investors" themselves have generally been happy to go along with the official story that allows them so much freedom and privacy. They have argued that they know what they are doing, they are doing a good job for their stakeholders and that policymakers have no need to get involved.

The consensus view of "sophisticated investors" is reassuring. If it is accurate, then regulators are right to allow big investors so much freedom and privacy. But evidence of poor outcomes and excessive fees raises doubts about how accurate the story really is.

PROBLEMS WITH THE OFFICIAL STORY

The most dramatic proof that the "sophisticated investor" doctrine does not stack up is, of course, the banking sector. Banks are the archetype of the "sophisticated investor." Yet they made enough bad investments to bring down the global economy if taxpayers had not bailed them out.

But the story does not end there, and the high profile nature of the problem with banks must not be allowed to obscure a bigger issue. However unlikely this may sound, banks are only one symptom of a problem that extends much wider.

The rest of this article will look at a generic large defined benefit pension scheme, to which ordinary people, including pension scheme members and taxpayers, entrust their cash. The pension fund hires specialist third-party fund managers to invest some of its funds in "alternative investments." Many of the examples used will relate to private equity, but the underlying issues relate to all "alternative investments."

Even their proponents admit that on average, "alternative investments" make a lot of money for the fund managers who run them,

2 Despite being so widely used, the term "sophisticated investor" has no formal status. Rule 501 of Regulation D in the U.S. federal securities laws defines an "accredited investor." The U.K.'s Financial Services Authority maintains a "Qualified Investor register" and Chapter 3 of its Conduct of Business Sourcebook defines "professional investor" without ever using the word "sophisticated." However, both media and regulators regularly use this informal term – see for example Financial Services Authority (2006), 3.132: "As private equity investors are generally sophisticated..." This is misleading: using the word "sophisticated" implies, without any evidence, that qualified (or accredited) investors will by definition make good investments. This article uses quotation marks in order to indicate the term's informal status. Examples in this article are drawn from the U.S. and the U.K., but most developed economies' regulations make a similar broad distinction between classes of investor.

3 In more formal economic terms, policymakers assume that that these organizations do not suffer from information asymmetries, and that they act like principals rather than agents.

4 This article leaves the term "alternative investments" in quotation marks to signal that it lacks any clear or consistent definition. It uses the term "private equity" generally to refer to buyouts (formerly known as leveraged buyouts), not venture capital. The Universities Superannuation Scheme (USS), one of the U.K.'s largest pension funds, provides a typical example of the trend. At March 2001, the USS investment portfolio contained no "alternative investments." Fourteen years later at March 2015, about one-quarter of the portfolio (25.3%, or £12.5 billion) was invested in assets that appear to meet USS's definition of "alternative assets." See USS (2015).

5 The term "pension fund" in this article refers to a defined benefit plan.

6 The U.K.'s Occupational Pension Schemes (Disclosure of Information) Regulations 1996 make clear that individuals are only entitled to receive information about benefits. U.S. public pension funds each year file a Comprehensive Annual Financial Report (CAFR). The CAFR contains headline information about returns and some information about costs. As will be discussed later, however, the information does not enable a detailed judgment on how good a job either the pension fund or its chosen fund managers are doing. See Dang et al. (2015).

but rather less for the people who invest in them. A 2011 report by the World Economic Forum observed that "... the [private equity] industry has been organized so that most of the rents (profits) from these skills go to the fund managers themselves, rather than to the [investors]."[7] The situation is not so different for hedge funds. A comprehensive academic study in 2011 found that over time, investors in hedge funds receive on average net returns that are little better than the returns on cash [Dichev and Yu (2011); Aiken et al. (2013)].

Most are familiar with the fact that retail investors run the risk of paying too much for supposed fund management skills. In their case, the reason is easy to find: they know less than professional fund managers.[8] But the official story suggests that "sophisticated investors" are different. These organizations can afford to employ skilled professionals, hence it seems strange that they have received only mediocre net returns from "alternative investments."

One of the main reasons why big investors are earning poor net returns in "alternative investments" is that they are paying excessive fees to the fund managers they hire. Over the last couple of years, this subject has received some long overdue attention. In May 2014, a Securities and Exchange Commission (SEC) official gave a high-profile speech that cited a lack of transparency and high fees in private equity [Bowden (2014)]. Two weeks later, he told the New York Times that "In some instances, investors' pockets are being picked. These investors may be sophisticated and they may be capable of protecting themselves, but much of what we're uncovering is undetectable by even the most sophisticated investor" [Morgenson (2014)].

U.S. state pension funds are among the biggest investors in the world. The largest of them, CalPERS, controls investments of over U.S.$300 billion, of which it allocates about one-tenth to private equity. Their size would suggest that state pension funds should also be considered among the most "sophisticated investors." But in April 2015, a well-known pension fund consulting firm issued a report that stated "[l]ess than one-half of the very substantial [private equity] costs incurred by U.S. pension funds are currently being disclosed" [Dang et al. (2015)]. The rest of the year saw this topic receiving a good deal of media attention.

Why are these "sophisticated investors" under-reporting these expenses? The answer appears to be that they do not fully understand what expenses they are paying in the first place. In July 2015, senior financial officials from twelve large U.S. states and cities wrote a joint letter to the SEC, asking the regulator "to require [private equity fund managers] to make better disclosure of private equity expenses to [investors]" [SEC (2015)]. In October 2015, California's State Treasurer wrote to the investment committees of California's two largest pension funds, stating that: "Pension funds and other limited partners pay excessive fees to private equity firms and do not have sufficient visibility into the nature and amount of those fees (...) The current lack of transparency undermines our fiduciary duty to protect our members and the public at large. Without it, how can we ever hope to have a meaningful dialogue with private equity firms, regulators, and other investors about the appropriate level of fees that should be paid?" [Chiang (2015); FT (2015a, b, c); NYT (2015); WSJ (2015)].[9]

A picture has begun to emerge. "Sophisticated investors" earn mediocre returns (in aggregate) on "alternative investments." One key reason is that they pay excessive fees, which they themselves do not fully understand. It is hard to square this picture with the official story about "sophisticated investors."

The gap that is appearing between the official story and what actually happens in practice is quite wide. It would be easier to understand such differences if there were a plausible explanation for why things are not working the way they are supposed to.

Fortunately – or rather, unfortunately – there is a plausible explanation: incentives matter. The people who work for "sophisticated investors" may be acting perfectly sensibly. But if their incentives are wrong, the outcomes will be, too.

INCENTIVES MATTER: TOO BIG TO FAIL (TBTF) AND CONFLICTS OF INTEREST

Several factors may distort incentives for the people who manage large investments. As far as banks are concerned, the most obvious problem is moral hazard; the issue colloquially known as TBTF. A belief that an organization is TBTF can result in situations where skillful individuals end up making poor investment decisions.

Although moral hazard is usually associated with large banks, it also applies to some other "sophisticated investors" as well. Consider a U.S. state pension fund, which unlike a bank faces little, to no risk of failing overnight. Over a longer period, though, it may fall

7 World Economic Forum (2011), page 60. The report's academic research was supervised by Professor Josh Lerner of Harvard University, one of the world's leading experts on private equity.

8 In technical terms, there is asymmetry of information between the investor and the fund manager.

9 Also see coverage under "Private equity" at www.nakedcapitalism.com

short of its obligations. Such a "failure" would lead to difficult decisions about whether to reduce pension benefits or to increase contributions from taxpayers. In practice, as long as states remain unwilling to see pensioners starve on the streets, politicians will likely call on taxpayers to make up at least some of the shortfall. Public pension funds share with TBTF banks an implicit ability to call on taxpayers in an emergency. The moral hazard created is the same. In the case of a pension fund, its potential effects will simply unfold more slowly.[10]

A second distortion of incentives can arise from what economists call "agency problems" – in plainer English, conflicts of interest [Jensen and Meckling (1976)]. The staff who work for a "sophisticated investor" do not have the same interests as the people whose money it is looking after. After all, it is not their money.

Just like the directors of quoted companies, pension fund trustees are there to look after the interests of a large number of widely-dispersed stakeholders: that is, other people's money. Everyone, including policymakers, understands that looking after other people's money creates a potential conflict of interest for the directors of quoted companies. It should be easy to see that the same applies to pension fund trustees. They, too, are looking after other people's money. Aligning different people's interests is always hard, irrespective of the circumstances. There is no reason why big investors, such as pension funds, should be uniquely immune to this issue.

WHY HAVE PROBLEMS NOT BEEN ADDRESSED?

The principles of both moral hazard (TBTF) and agency problems (other people's money) are generally well understood. In some areas, policymakers recognize these problems and take active steps to address them. Think, for example, of all the effort that policymakers have put into trying to solve TBTF problems within the banking system since the recent crisis. It is hard to discern just how successful they have been in this regard, though the nature of the problem leaves policymakers no choice but to maintain in public that they have been successful.

As far as agency problems are concerned, policymakers have been trying for years to address the conflicts of interest that exist with publicly quoted companies. The U.K. alone has seen a string of reports, the Cadbury Report (1992) and Greenbury Report (1995), among others.

Where big investors are concerned, though, policymakers seem to overlook the agency problems that are inherent within public companies. They studiously ignore the idea that these same problems apply to big investors as well, leading them to make poor investments. This is either naive or disingenuous. Why have so many people placed such excessive faith in the "sophisticated investor"?

One possible answer would involve the notion of "fiduciary duty." Supporters of the status quo might point out that pension fund trustees have a fiduciary duty to look after the interests of their members. They might suggest that this is sufficient by itself to ensure that trustees will be effective in looking after their members' interests. But it is not.

The easiest way to see why it is not sufficient is to compare the position of a pension fund trustee with that of the chief executive of a quoted company. Both the trustee and the CEO are looking after other people's money. Both have a fiduciary duty to widely-dispersed stakeholders.[11] No one, however, would suggest that fiduciary duty is enough to ensure that a CEO will always act in the best interests of the company's shareholders. If fiduciary duty was sufficient, there would be no need to require companies to provide so much public disclosure. Nor would policymakers have put so much time and energy over the years into studies such as the Cadbury Code, the Greenbury Code, and so on.

Quoted companies have to disclose information that will allow stakeholders to assess the performance of fiduciaries, such as the CEO. Meanwhile, the fiduciaries who run pension funds are exempt from such requirements. This begs an obvious question: why does the consensus treat one set of fiduciaries (CEOs) so differently from another (pension fund trustees)?

The conventional answer might be that, unlike CEOs, pension fund trustees do not have the opportunity to enrich themselves at the expense of the stakeholders, whose interests they represent. Serving as a pension fund trustee is generally seen more as a public service than as an opportunity to make money. While CEOs can become rich if they perform well, pension fund trustees are paid more modestly and on a more or less fixed basis. Perhaps this explains why the consensus assumes that for trustees, as opposed to CEOs, fiduciary duty will be enough.

10 The moral hazard present within large public pension funds puts them in a different position from some other "sophisticated investors," such as private university endowments. The trustees of a college endowment know that if they make poor investment decisions, their organization will have no claim on the public purse.

11 In the case of CEOs, this is a simplification. Strictly speaking, CEOs' fiduciary duty is to their employers (the company) rather than to the company's shareholders.

If so, the consensus is missing a crucial point: there is more than one way for interests to be poorly aligned. Pension fund trustees may not be able get rich at the expense of stakeholders, but that does not mean their interests are well aligned. Trustees' incentives are not skewed to the upside, like CEOs'. They are, however, skewed to the downside. Pension fund trustees face at least two forms of downside risk. Breaking legal obligations carries a cost for trustees. Less tangible, but arguably more significant, is reputation risk. Trustees whose funds underperform their peers will face criticism.

Being exposed only to downside risk is a recipe for risk-averse behavior. Unlike CEOs, pension fund trustees have no incentive to do anything different from their peers. Rather, they have every incentive to follow the herd. That is the best way for them to minimize the risk of underperforming their peers. Being too conservative can do just as much damage as taking too much risk.[12]

One way to think about fiduciary duty is as a legal device for trying to make agents (such as CEOs and pension fund trustees) act like principals (that is, as though they were looking after their own cash). Fiduciary duty gives pension fund trustees one incentive, but it is not the only one they face. The consensus is wrong to assume that because pension fund trustees cannot get rich, fiduciary duty alone will ensure they make optimal decisions on behalf the people they represent.

It may seem strange that the consensus has failed to think clearly about how agency problems affect big investors. Incentives may be at work once again, however, because most of those who support the "sophisticated investor" story have a vested interest in it. The people who work for big investors are bound to like the label "sophisticated" thanks to the freedom and lack of scrutiny that comes with it. They will earn more for looking after a complex "alternative investment" than for supervising simpler (and cheaper) investments.

A second group has an even bigger financial incentive to support the status quo: intermediaries such as banks, asset managers, consultants and lawyers. All of them do well if big investors use more complexity. For example, big investors pay consultants to advise them how and where to invest. It should come as no surprise to find that consultants have mostly supported the steady growth in complex (and expensive) "alternative investments."

A third and last group has (one hopes) influenced policymakers more than either of the first two. Its vested interest is also of a different kind. Conventional economic theory makes it an article of faith that big investors will (on average, if unconstrained) make good investments. Conventional theory chooses to ignore the moral hazard and agency problems discussed above. This means that mainstream economics has an intellectual vested interest in the "sophisticated investor" story. It has an incentive to turn a blind eye to what could be wrong with the official story.

PROPOSAL: MANDATORY DISCLOSURE

The flaws in the "sophisticated investor" story are easy to see. A simplistic view of fiduciary duty, combined with vested interests, may explain why these flaws have received less attention than they deserve until now. Finding neat and tidy solutions is hard.

But the people who depend on big investors cannot afford to wait for the perfect answer. This is not an abstract issue. If "sophisticated investors" do a poor job, ordinary people suffer. When ordinary people's own cash is involved, the impact is direct. The SEC's Andrew Bowden noted in 2014 that what happens in private equity "affects the retirement savings of teachers, firemen, police officers, and other workers across the U.S." [Bowden (2014)].[13] Even if their own money is not involved, ordinary people are vulnerable to indirect effects. At best, too much capital flowing into "alternative investments" reduces economic growth for the whole of society. At worst, taxpayers may find themselves bailing out banks in the short term and pension funds in the long term.

Both ends of the ideological spectrum will offer simplistic solutions. One end will suggest a raft of new regulations that prescribes in detail who can do what. For example, someone might suggest that big investors have to "prove" to regulators that they are good investors before being allowed to take on "alternative investments." The bureaucracy involved would be expensive and would not work.[14]

12 Readers familiar with financial options may find the following analogy useful in explaining how different the incentives are for CEOs and pension fund trustees despite the fact that they are both fiduciaries. A CEO's incentive structure can be compared to a long call option, which gives them an incentive to take risks. A pension fund trustee's incentive structure is the reverse: a short put option. An investor who is short a put option will try to minimize risk. (Thanks for this analogy are due to Jack Edmondson.)

13 In 2007, private equity firm Permira told a U.K. parliamentary committee "We have 30 million pensioners in our pension funds [sic] and millions of them are in the U.K. For instance, we have at least one million local government employees, past and present, who invest in our funds…" (House of Commons 2007, Ev 34.)

14 A variant on this could see "sophisticated investors" split into groups depending on their level of sophistication – see, for example, Tett (2010). It does seem strange, and perhaps dangerous, that the current rules make little distinction between (say) Goldman Sachs and a small local council. But this approach would also involve costly and tricky bureaucracy.

The other end of the spectrum will feature siren voices of the kind that have dominated finance over the last two generations. These will proclaim that the answer is simple: just get rid of agency costs by "aligning the interests" of principals and agents, and the problem will magically disappear.

"Alignment of interests" has acquired totemic status in the world of "alternative investments."[15] But that is a sign of how superficial the thinking behind it is. In reality, "alignment of interests" is a mirage. There is only one way to truly align the interests of a principal and an agent: that is literally to merge them. Anything less can only produce an imperfect alignment of interests. The alignment may be more or less imperfect, but that is all. And, paradoxically, imperfect alignment of interests is more dangerous than none at all. That is because it gives a principal false confidence that they can afford to stop worrying about their agent's conflict.

The agency problems that affect big investors are here to stay. They cannot be either avoided or eliminated. The most we can hope to do is to mitigate them. Anyone who believes in markets will accept that the most effective way to do this is to harness the self-interest of the people that big investors' money actually belongs to. This is where policymakers have a vital role to play.

Policymakers have to make it possible for ordinary people to hold accountable the agents, such as pension funds, to which they have entrusted their cash. This includes millions of pension scheme members and taxpayers. To hold someone accountable, you need to be able to assess their performance. But "sophisticated investors" are allowed to operate in private. That makes it impossible to see fully how they are doing.

Policymakers need to remove this exemption. They must make it mandatory for big investors, and the asset managers they hire, to release more information to the public.[16] The data disclosed must allow a detailed and truly independent analysis of how big investors have performed, including how much they have paid in expenses.[17]

The details of how to put this approach into practice lie beyond the scope of this article. Cracks have already started to appear in the historic consensus that big investors can look after themselves and should be left to operate in private. Events in the U.S. in 2014-15, discussed earlier, are one example. Something similar happened in the U.K. in 2014, though it received less publicity. The Chartered Institute of Public Finance & Accounting (CIPFA) is a body that oversees public sector accounting in the U.K. In June 2014 it issued a report that addressed the way local government pension schemes report the cost of investment management [CIPFA (2014)]. CIPFA essentially pointed out that pension schemes were under-reporting the costs for "alternative investments," in the same way that CEM Benchmarking's 2015 report did for private equity in the U.S.[18]

Assuming policymakers take on this challenge, they will have to be robust about what they do. Even when they have tried to make disclosure mandatory in the past, "sophisticated investors" and intermediaries have tried energetically to get around the rules.[19]

Some vested interests will go on resisting the idea that they should be more open. In November 2014 the Chief Executive of the U.S. private equity industry's lobby group wrote "The argument that [private equity] limited partnership agreements (LPAs) should be accessible to the public is akin to demanding that Coca-Cola publish its famous (and secret) soda recipe."[20] This is disingenuous. Morris and Phalippou (2012) show why this analogy does not apply to private equity. Information that is genuinely time-sensitive creates a real challenge. Zingales (2009) suggests an elegant solution: where appropriate, simply allow a time-delay on its disclosure.

Vested interests may also suggest that it is pointless to release more information because the average person would not be able to interpret it. Once again this is disingenuous. The fact that individuals cannot interpret such information is irrelevant. If the data were publicly available, independent experts without a vested interest (e.g., academics) would analyze them for free. This would

15 Blackstone, a major "alternative investment" manager, writes that "We strive to maintain a work environment that reinforces our culture of collaboration, motivation and **alignment of interests** with investors [emphasis added]" [Blackstone (2010)]. A randomly chosen "sophisticated investor" writes: "One of the greatest strengths of the hedge fund industry is the **alignment of interest** that is created with 'pay for performance' carry fee structure [emphasis added]" [Utah (2009)]. In 2007, private equity manager Permira told a U.K. parliamentary committee that "Our pension fund investors are some of the largest and most sophisticated in the world. They spend a huge amount of time doing due diligence on our funds and an inordinate amount of time looking at the **alignment of interest** between us and them. [emphasis added]." [House of Commons (2007, Ev 50)]

16 Zingales (2009) and Morris and Phalippou (2012) present more detailed arguments for mandatory and standardised disclosure.

17 As discussed earlier, current reporting by U.S. public pension funds is inadequate for this purpose.

18 One pension scheme that adopted CIPFA's new measures showed an almost eight-fold increase in the "investment management expenses" it reported for 2014/15, from £10.7 million to £81.2 million [West Midlands (2015)]. The scheme observed that "this is a change in reporting only and does not represent an actual increase in costs…" But it shows how dramatic the under-reporting of costs for "alternative investments" has been until now. As the Treasurer of California wrote in October 2015, "Without [a clear view of costs], how can we ever hope to have a meaningful dialogue with private equity firms, regulators, and other investors about the appropriate level of fees that should be paid?"

19 Abrahamson et al. (2012) is an alarming cautionary tale about how "sophisticated investors" try to evade even mandatory disclosure rules.

20 Steve Judge, CEO, Private Equity Growth Capital Council, PEHub, 3 November 2014. Available at: https://www.pehub.com/2014/11/confidentiality-of-limited-partnership-agreements-is-paramount/

lead to a more informed debate about an issue that has serious public consequences. No one expects the average person to be able to interpret data about tests for breast cancer. But nor does anyone expect a small group of insiders with a vested interest to be allowed to keep those data private and use them to extract rent from the rest of society.[21]

None of the routine objections from vested interests stands up to scrutiny. But privacy has become deeply ingrained in the world of "sophisticated investors." They have even persuaded many neutral observers that privacy is essential. Some historical context will help show how wrong this is.

A HISTORICAL PERSPECTIVE

In 1913, the American lawyer Louis Brandeis wrote a series of articles about the power of the U.S. finance sector. In one of them he coined the phrase "Sunlight is said to be the best of disinfectants..."[22] 20 years later, opacity was still the norm. Even the financial crash of 1929 and the onset of the Great Depression had brought little change. Quoted U.S. companies were still able to get away with disclosing poor quality information. The incoming President proposed creating a new agency to address this problem. It was called the Securities and Exchange Commission (SEC).[23]

U.S. financial and business interests lobbied against it fiercely. "There is no important economic aspect of the economic life of this country," intones the President of the New York Stock Exchange in a surviving February 1934 newsreel, "whether it be agriculture, industry, banking or commerce, which will not be adversely affected by this Bill. This Bill, if passed by Congress, will not only destroy our security markets, but will as a necessary consequence interrupt the flow of credit and capital into business."[24]

80 years later, no one would suggest that the SEC is perfect – indeed, it receives criticism from both ends of the ideological spectrum.[25] But both academic and anecdotal evidence confirms that its disclosure rules have made U.S. capital markets work better than they would have done otherwise [Fox et al. (2003)].

In effect, the SEC shone the "sunlight" that Brandeis wrote about in 1913. Quoted firms now have to file standard financial reports. These must be timely, relevant and easy to obtain and compare. Doing so has not seriously damaged American firms' ability to compete. Disclosure clearly involves some cost. But any private cost is dwarfed by the public benefits that flow from creating deep and trusted markets.

80 years ago, finance sector lobbyists warned that improved disclosure by quoted companies would bring the U.S. economy down. It did not. Instead, it helped the market for quoted securities work better and regain public trust. Improved public scrutiny of "sophisticated investors" would have the same effect today.

CONCLUSION

Plenty of "sophisticated investors" make good investments. The vast majority of people who work for big investors are acting in good faith. Some "alternative investments" are good value for investors. None of these has anything to fear from improved disclosure. But policymakers have to think about aggregate outcomes: not the better performers, nor the inherited dogma, but overall reality. And outcomes appear to be sub-optimal. Where "alternative investments" are concerned, "sophisticated investors" in aggregate seem to be letting down the ordinary people who depend on them. One key reason is that big investors are over-paying the fund managers they hire.

Some people who believe in markets may find this idea hard to accept. In truth, they should be neither surprised nor downhearted. Agency problems (conflicts of interest) affect most other human institutions. It would be very strange if they did not also affect big investors. Agency problems are here to stay. They provide a very straightforward reason why big investors collectively do not do as good a job as the consensus view has simply assumed they do.

It is poor outcomes that make improved disclosure necessary. This is not a pointless, pro forma fishing expedition: it is the most market-friendly way to try and mitigate agency problems that are causing real harm. Anyone who believes in markets will understand that harnessing people's self-interest is a powerful tool. Policymakers can do that here by ending the monopoly big investors have historically had on key information. They must help outsiders (meaning, the rest of society) look after their own interests.

21 For evidence of finance sector rents, see Philippon and Reshef (2012).

22 Brandeis (1932). The title of Andrew Bowden's SEC speech in April 2014 was "Spreading sunshine in private equity."

23 Zingales (2009) draws a similar parallel with the creation of the SEC.

24 Available at http://www.sechistorical.org/museum/film-radio-television/. For the origins of the SEC, see McCraw (1984).

25 See http://www.reuters.com/article/2011/09/15/sec-schapiro-idUSS1E78D1QL20110915 and http://www.theatlantic.com/business/archive/2011/12/too-big-to-stop-why-big-banks-keep-getting-away-with-breaking-the-law/249952/

Better disclosure is not a panacea: it is necessary, but not sufficient. Nor must it be used as an excuse to dilute fiduciary protection for small investors or beneficiaries. Rather, it is a way to help make sure fiduciaries are doing their job well. Some finance sector insiders may lose out from better disclosure. But financial markets are supposed to serve the interests of society as a whole, not a small group of insiders. Regulators can and must help to make that happen by opening up "sophisticated investors" to proper scrutiny.

REFERENCES

- Abrahamson, M., T. Jenkinson, and M. Sousa, 2012, "Does transparency overcome conflicts of interest? Evidence from investment managers and their brokers," working paper available at www.ssrn.com
- Aiken, A., C. Clifford, and J. Ellis, 2013, "Out of the dark: hedge fund reporting biases and commercial databases," 2013, Review of Financial Studies 26:1, 208-243
- Blackstone, 2010, 10-K filing 2010, available at www.blackstone.com or at www.sec.gov
- Bowden, A. J., 2014, "Spreading sunshine in private equity," speech delivered at the Private Equity International Private Fund Compliance Forum, 6 May, available at https://www.sec.gov/news/speech/2014--spch05062014ab.html
- Brandeis, L., 1932, "What publicity can do" in Brandeis, L., 1932, Other people's money, chapter 5, p. 92, first published in Harper's Weekly, December 20, 1913
- Brandeis, L., and N. Hapgood, 1933, Other people's money: and how the bankers use it, (republished from 1914), National Home Library Foundation, Washington
- Cadbury Report, 1992, "Report on the financial aspects of corporate governance," available at http://www.icaew.com/~/media/corporate/files/library/subjects/corporate%20governance/financial%20aspects%20of%20corporate%20governance.ashx
- CIPFA, 2014, "Accounting for local government pension scheme management costs," Chartered Institute of Public Finance and Accountancy, June
- Chiang, J., 2015, Letter to Investment Committee Chairs, CalPERS and CalSTRS, 12 October, available at http://www.nakedcapitalism.com/wp-content/uploads/2015/10/John-Chiang-Letter-to-CalPERS-and-CalSTRS-Oct-2015.pdf
- Dang, A., D. Dupont, and M. Heale, 2015, "The time has come for standardized total cost disclosure for private equity," CEM Benchmarking, April, available at http://www.cembenchmarking.com/Files/Documents/CEM_article_-_The_time_has_come_for_standardized_total_cost_disclosure_for_private_equity.pdf
- Dichev, I., and G. Yu, 2011, "Higher risk, lower returns: what hedge fund investors really earn," Journal of Financial Economics 100:2, 248–263
- FT, 2015a, "Calpers' private equity problems pile up," Financial Times, 12 July
- FT, 2015b, "Calstrs: U.S. private equity woes deepen," Financial Times, 19 July
- FT, 2015c, "Calstrs under pressure to reveal private equity fees," 6 December
- FSA, 2007, "Private equity: a discussion of risk and regulatory engagement," Financial Services Authority discussion paper 06/6, November; feedback statement 07/3, June, available at http://www.fsa.gov.uk/pubs/discussion/dp06_06.pdf and http://www.fsa.gov.uk/pages/library/policy/dp/2007/fs07_03.shtml
- Fox, M., A. Durnev, R. Morck, and B. Yeung, 2003, "Law, share price accuracy and economic performance: the new evidence," Michigan Law Review 102, 331-386
- Greenbury Report, 1995, "Directors' remuneration – report of a study group chaired by Sir Richard Greenbury," July, available at http://www.ecgi.org/codes/documents/greenbury.pdf
- House of Commons Treasury Committee, 2007, "Private equity – tenth report of Session 2006-07, volume II," available at www.tsoshop.co.uk
- Jensen, M., and W. H. Meckling, 1976, "Theory of the firm: managerial behavior, agency costs and ownership structure," Journal of Financial Economics 3:4, 305-360
- McCraw, T., 1984, Prophets of regulation, Harvard University Press
- Morgenson, G., 2014, "The deal's done. But not the fees," New York Times, 24 May, available at http://www.nytimes.com/2014/05/25/business/the-deals-done-but-not-the-fees.html?_r=0
- Morris, P., and L. Phalippou, 2012, "A new approach to regulating private equity," Journal of Corporate Law Studies 12:1, 59–86
- NYT, 2015, "When private equity firms give retirees the short end," New York Times, 13 June
- Philippon, T., and A. Reshef, 2012, "Wages and human capital in the U.S. financial industry: 1909-2006," Quarterly Journal of Economics 127:4, 1551-1609
- SEC, 2015, "Re: standardized private equity fee disclosure," Letter to SEC Chair, 23 July, available at https://www.sec.gov/rules/petitions/2015/petn4-688.pdf
- Tett, G., 2010, "'Sophistication' debate heats up," Financial Times, 6 May
- The Occupational Pension Schemes (disclosure of information) regulations 1996, available at http://www.legislation.gov.uk/uksi/1996/1655/contents/made
- Universities Superannuation Scheme, Annual Report and Accounts 2015, available at https://www.uss.co.uk/how-uss-is-run/running-uss/annual-reports-and-accounts
- Utah Retirement Systems, 2009, "Summary of preferred hedge fund terms," available at http://online.wsj.com/public/resources/documents/ursnote03102009.pdf
- WSJ, 2015, "Calpers' private-equity fees: $3.4 billion," Wall Street Journal, 24 November
- West Midlands Pension Fund Annual Report 2015, available at http://www.wmpfonline.com/CHttpHandler.ashx?id=7083&p=0
- World Economic Forum, 2011, "The future of long-term investing," February, available at http://www3.weforum.org/docs/WEF_FutureLongTermInvesting_Report_2011.pdf
- Zingales, L., 2009, "The future of securities regulation," The Journal of Accounting Research 47:2, 391-425

Fund Transfer Pricing for Bank Deposits: The Case of Products with Undefined Maturity

Jean Dermine – Professor of Banking and Finance, INSEAD[1]

HIGHLY TECHNICAL

Abstract

The paper presents a pedagogical yet rigorous analysis of fund transfer pricing for deposits with undefined maturity. The objective is to identify the conditions needed to convert the case of deposits with undefined maturity into one with a single effective maturity. This in turn allows us to identify the many circumstances under which the practice of conversion into a single effective maturity is not warranted. Attention is called to the context in which the choice of a maturity is made: pricing, evaluation of performance and hedging of interest rate risk on deposits with undefined maturity.

1 I would like to acknowledge the editorial comments of Hazel Hamelin.

INTRODUCTION

Fund transfer pricing (FTP) is used by bankers to evaluate the profitability of deposits and loans and for pricing. It is used by academics and antitrust authorities to evaluate the degree of competition in banking markets. The challenge, as far as on-balance sheet banking is concerned, is as follows. When one evaluates the profitability of deposits, one knows the cost – the interest paid on deposits and the operating expenses associated with deposits collection, such as employee time and IT. However, determining the return on deposits is more problematic because they can be used to finance various types of assets: consumer loans, corporate loans, interbank assets, bonds, and fixed assets. Revenue – known as the fund transfer price – must be identified to remunerate deposits.

For loans, the problem is symmetrical: the return on loans is known (that is, the interest income net of expected bad debt expense), but not the cost of funding loans. The reason for this is that banks use several sources of funds to finance assets: demand deposits, savings deposits, time deposits, corporate deposits, interbank deposits, subordinated debt, and equity. Again, there will be a need for a specific fund transfer price to evaluate the cost of funding loans. Appropriate identification of the FTP, in particular its maturity, is fundamental for the pricing of commercial products, performance evaluation, bank strategy design and hedging of interest-rate risk.

In three publications [Dermine (2007, 2013, and 2015)], I present foundation and advanced approaches to fund transfer pricing. The foundation approach, used sthroughout the banking world, covers two cases: products with fixed and undefined maturities. I argued that as a result of the global financial crisis, attention should be given to five potential issues: rationing on the interbank market, the funding of a Basel III liquidity coverage ratio, the necessity to adjust FTP to the credit-riskiness of specific assets, the need to include a liquidity premium in the case of long-term funding and, finally, the choice of a consistent methodology to incorporate the credit spread on the bank's own debt due to the perceived risk of bank default. I concluded that an advanced approach to fund transfer pricing must be adopted by banks.

Having observed the heated debate that the choice of a specific maturity for the FTP applicable to deposits with undefined maturity – such as demand and savings deposits – can generate, I propose a pedagogical yet rigorous discussion of the issues involved. More specifically, I have observed on several occasions an attempt to identify a single effective maturity which makes it possible to convert the complex case of products with undefined maturity into one with a fixed effective maturity. Deposits are divided in two (or several) buckets: (1) volatile "transient" deposits with a short-maturity and (2) loyal "core" deposits with a long-maturity. The effective maturity is then a weighted average of short- and long-maturity buckets. Indeed, the notion of "behavioral" maturity is referred to by the European Banking Authority (2015) in a report on the measurement of interest-rate risk on the banking book. My purpose here is to identify the conditions that are necessary to convert the case of deposits with undefined maturity into one with a single effective maturity in order to shed light upon the many circumstances under which this simplification is not warranted and a more complex multi-period setting would then apply.

The choice of an economic maturity as opposed to a contractual maturity arises in three contexts: the pricing of the product, the evaluation of performance of a business unit and the selection of a hedge against interest rate risk.[2] As the relevant maturity might not be the same for the three applications, the choice of the FTP maturity has to be set in a specific context.

To illustrate the nature of the issue, consider the following scenario. A Lebanese bank raises deposits in U.S. dollars. Those in charge of pricing deposits argue that, since these deposits are fairly stable, they could be invested in a 5-year fixed-rate U.S.$-denominated eurobond issued by the Lebanese government. As the return on such bonds is 7%, they propose to pay 5% on a very competitive market for U.S.$-denominated savings deposits. At the time, the 3-month interest rate on a U.S.$-denominated Lebanese government bond is 4%. Those in charge of asset and liability management (ALM) wonder whether the single effective maturity of five years chosen to identify the benchmark market rate is warranted, while those in charge of managing risks wonder about the hazard that investing these savings deposits in Lebanese five-year fixed rate bond represents. This illustrates the problematic nature of choosing an effective maturity for deposits with undefined maturity and the desire to simplify and convert this case into one with a single effective maturity.

The review of the foundation approach to fund transfer pricing in Section 2 is followed by a discussion of FTP for deposits with undefined maturity in Section 3. Numerical examples are used to illustrate the nature of the problem and solutions.

2 An additional issue not discussed in the paper is the measurement of liquidity risk on deposits with undefined maturity.

THE FOUNDATION APPROACH: PRODUCTS WITH DEFINED MATURITY

The foundation approach to fund transfer pricing for products with fixed maturity is represented in Figure 1.

The horizontal line represents the market rate, i.e., the interest rate observed on the interbank market (LIBOR).[3] The line is horizontal as the interest rate is set on large international markets and is independent of the volume of transactions initiated by the bank. The two other lines represent the marginal income on loans and the marginal cost of deposits. As a bank wishes to increase its loan portfolio, the expected income from an additional dollar of loan – the marginal or incremental income – will go down because the bank needs to reduce the interest rate to attract the additional dollar of loan, or because the bank is willing to agree to a loan of lower quality. Similarly, the cost of collecting an additional dollar of deposits – the marginal or incremental cost of deposits – will go up because the bank either needs to raise the deposit rate to attract the additional dollar of deposits or to open additional branches in remote areas. In Figure 1, the optimal volume of deposits, D^{OPT}, is reached when the marginal cost of deposits is equal to the opportunity market rate. One would not want to go beyond D^{OPT} because the incremental cost of deposits would be higher than the return earned on the money markets.[4] Similarly, the optimal volume of loans, L^{OPT}, is reached when the marginal revenue from loans is equal to the marginal investment return, the market rate. One would not want to increase the loan portfolio beyond L^{OPT} because the incremental income on the new loan would be lower than the return available on the money markets. The maturity of the market rate used for fund transfer pricing should correspond to the maturity of the fixed-term product. For shorter maturities (up to one-year) the interbank market rates are frequently used, while for longer fixed-rate maturities the swap rates are used.[5] Matching maturities not only has intuitive appeal for the search of a relevant opportunity cost, it also insulates the commercial units against the impact of interest rate (or currency) fluctuations. Interest rate (or currency) mismatches are transferred to the ALM department in charge of managing these sources of risk [Dermine (2015)], which is implicitly assumed to have the tools necessary to manage the maturity mismatch created by the loan and deposit commercial units. There is a separation between the profit earned from margins on loans and deposits (benchmarked against a matched-maturity market rate) and the profit realized by the ALM department in mismatching the book. This is justified by the respective types of expertise required to price loans and deposits, and in forecasting interest rates.[6]

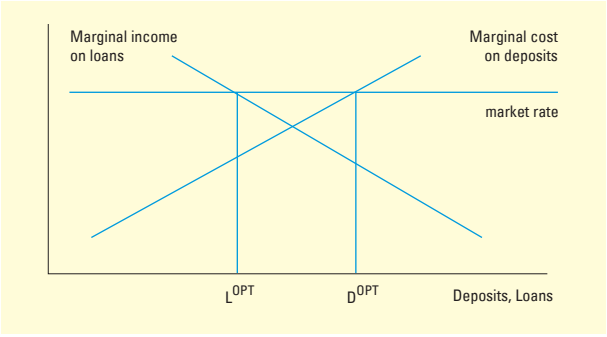


Figure 1 – The separation theorem

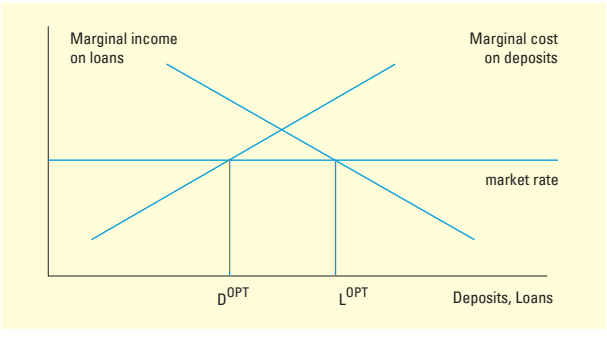


Figure 2 – The separation theorem

Note that there is a separation between the lending and funding decisions. Separation theorem states that loans and deposits must be priced with reference to the market rate and that these decisions are independent of one another. The difference between the optimal volumes of deposits and loans (D^{OPT} - L^{OPT}) is the net position in treasury, bonds or interbank assets. In Figure 1 it is positive with deposits exceeding the volume of loans. The bank is a net lender in the money market. But it could be negative with the bank being a net borrower, as illustrated in Figure 2. In this case, the difference between the volume of loans and deposits (L^{OPT} - D^{OPT}) must be funded in the money markets.

3 In countries with illiquid interbank markets, the relevant market rate is the interest rate on government bonds.

4 We ignore reserve requirements with the central bank, which reduce the revenue earned on deposits.

5 The swap rate gives the long-term cost of the roll-over of short-term interbank funding that is hedged with a swap. This is likely to differ from the actual cost of funding the long-term asset with a long-term debt that would include a liquidity or credit spread. The use of a swap rate is appropriate when the bank performs the traditional function of maturity transformation, funding long-term assets with short-term debt. The case of maturity matching – long-term assets funded with long-term debt – is analyzed in Dermine (2013).

6 When the ALM department does not have access to treasury products to manage the maturity mismatch at reasonable cost, the separation between deposits and loans breaks down because the interest rate risk is the result of the joint decision on loans and deposits. In this more complex case, one needs to find the appropriate joint mix of loans and deposits that maximizes the present value of future expected profits under reasonable interest rate risk constraints.

FOUNDATION APPROACH, PRODUCTS WITH UNDEFINED MATURITY

In the foundation approach to fund transfer pricing, the relevant maturity for the marginal return is that of the deposit or loan. A two-year deposit should be priced against the two-year matched maturity market rate. However, there are several well-known cases, such as demand or savings deposits, for which the contractual maturity (very short as withdrawable on demand) is different from the effective economic maturity. Indeed, many deposits are fairly sticky with a longer effective maturity and the Basel Committee (2015) refers to non-maturity deposits (NMDs). Often bankers and regulators attempt to identify a behavioral maturity that would make it possible to convert a case with undefined maturity into one with a fixed effective maturity. Hence the question as to the conditions needed for the conversion.

Let us return to the example mentioned in introduction. In Lebanon, the deposit dollarization ratio reaches 66.1% in 2013 [Bank Audi (2013)]. Due to stiff competition, the bank proposes to pay 5% on U.S.$ savings deposits withdrawable on demand as it can invest the money in a fixed-rate 7% U.S.$ 5-years-to-maturity bond issued by the Lebanese government in international markets. The choice of a long-maturity is justified by the argument that savings deposits are fairly stable. The above discussion illustrates the search for a single relevant maturity that makes it possible to apply the framework of the foundation approach for products with defined maturity.

In some cases, one assumes that a fraction of the deposits (α) is volatile, equivalent to short-term "transient" deposits, while the complement (1- α) behaves like long-term "core" deposits. The effective maturity becomes a weighted-average maturity of short- and long-term deposits. Again, the conditions needed to apply a weighted-average maturity for the choice of the benchmark rate merits further investigation.

With reference to pricing a product and performance evaluation, a related question arises as to whether the analysis can be conducted over a short period – say one year – and whether this process will lead to optimal decisions. Indeed, to reward bank executives, it is customary to evaluate their performance over a relatively short period, such as one quarter or one year, but it is important to ensure that this does not create sub-optimal short-term biases in decision making.

To illustrate the sources of complexity arising from deposits with undefined maturity, we consider the case of deposits over a 2-year horizon, although this could be extended to a more realistic multi-year setting.[7]

The 1-year-to-maturity bond rate in Year 1, b_1, is 4% and the 1-year-to-maturity bond rate expected at the start of year 2, b_2, is 6%. The fixed coupon c on a 2-year-to-maturity bond is 4.971 %.

The coupon of 4.971% ensures that the fair value of the bond, the present value of future cash flows,[8] is equal to 100: 100 = (4.971/1.04) + (104.971/1.04 x 1.06)

The coupon rate of 4.971% (see calculation in Appendix 1) also ensures that a 1-year investment strategy with roll-over yields the same return as a 2-year investment strategy with reinvestment of the annual coupon. Consider the case of an initial investment of 100:

1-year investment strategy with rollover: 100 x 1.04 x 1.06 = 110.24
2-year investment strategy with reinvestment of interim coupon: (4.971 x 1.06) + 104.971 = 110.24

It is assumed that the ALM department of the bank can hedge the interest rate risk, so that the focus of the commercial units is entirely on the interest margins, taken as given by the current market rates and those implied in the yield curve.[9]

In this 2-year scenario, deposits with undefined maturity are collected. Undefined maturity has two dimensions. It refers first to the fact that some of the deposits collected in year 1 will still be deposited in the bank in year 2. This introduces the notion of temporal dependence – in our case over two years – between the volumes collected in year 1 and year 2. A second potential source of temporal dependence arises from price rigidity when the deposit rate set in year 2 is related to the interest rate set the previous year.

Several cases likely to be observed in the real world will be considered. The objective is to identify the conditions in which deposits with undefined maturity can be converted into a simpler one with a single effective maturity, taking into account three perspectives: pricing, evaluation of performance, and management of interest rate risk.

A parsimonious approach with five cases is chosen to focus and illustrate the sources of the time dependence. They are as follows:

7 A spreadsheet with solutions to examples is available from the author upon request.
8 We implicitly assume a risk neutral world and a liquid bond market. Interest rate and liquidity risk premia are assumed to be 0%.
9 Note that the ALM department might decide not to hedge the position. There is a separation between the commercial units focusing on interest margins based on current interest rates and the role of ALM in managing the maturity mismatch.

Case 1: Independence between deposits collected in year 1 and those collected in year 2, and flexible deposit rates in both years.

In the next four cases, we analyze the implications of temporal dependence originating from deposit volumes and/or deposit rates.
Case 2.1: Log-linear dependence between deposit volumes and flexible deposit rates.
Case 2.2: Complete rigidity of volumes and deposit rates over the two years.
Case 2.3: Linear-additive volume dependence and rigid deposit rate.
Case 2.4: Linear-additive volume dependence and discriminatory pricing.

CASE 1: INDEPENDENCE BETWEEN YEAR 1 AND YEAR 2

Consider a first case in which the supply of deposits in year i (D_i), i = 1, 2, is positively related to the deposit rate d_i offered in year i and negatively related to an investment opportunity competitive bond rate b_i. Let us assume that both the volume of deposits D_2 and the deposit rate d_2 chosen in Year 2 are independent of what happened in Year 1. That is, there is independence between Year 1 and Year 2. The case of complete independence is used as a benchmark to study the sources of temporal dependence.

The supply of deposits in years 1 and 2 are given by the following log-linear relations[10]:

$$D_1 = 100{,}000 \times b_1^{-1.5} \times d_1^2$$

$$D_2 = 100{,}000 \times b_2^{-1.5} \times d_1^2$$

The log-linear function is chosen because, as shown in Appendix 2, the price elasticity is the exponent of the deposit rate variable. Assuming that the deposits collected are invested in 1-year-to-maturity bonds,[11] the maximization of the present value of future profits, evaluated at the end of Year 1, is equal to:

$$\text{MaximizeValue} = (4\% - d_1) \times D_1 + \frac{(6\% - d_2) \times D_2}{1+6\%}$$

In case of independence between years 1 and 2 – the volume of deposits and the deposit rate in year 2 are unrelated to what happened in year 1 – one can maximize the profit of each year separately. In this case, the fund transfer price to be used in pricing and in evaluating the performance in year 1 is the matched-maturity 1-year maturity bond rate b_1 = 4%, and the FTP for the second year is b_2 = 6%.

As is shown in Appendix 2, the optimal deposit rate that maximizes profit in years 1 and 2 are given by the following relations, ε denoting the deposit rate elasticity:

$$d_1^{\text{OPTIMAL}} = b_1 \times \frac{1}{(1+\frac{1}{\varepsilon})} = 4.0\% \times \frac{1}{(1+\frac{1}{2})} = 4.0\% \times 0.6666 = 2.67\%$$

$$\text{Profit in year 1} = (4.0\% - 2.67\%) \times 100{,}000 \times 4.0^{-1.5} \times 2.67^2 = 1{,}185.19$$

$$d_2^{\text{OPTIMAL}} = b_2 \times \frac{1}{(1+\frac{1}{\varepsilon})} = 6.0\% \times \frac{1}{(1+\frac{1}{2})} = 6.0\% \times 0.6666 = 4.0\%$$

$$\text{Profit in year 2} = (6.0\% - 4.0\%) \times 100{,}000 \times 6.0^{-1.5} \times 4.0^2 = 2{,}177.32$$

$$\text{PresentValue of Profits} = 1{,}185.19 + 2{,}177.32/1.06 = 3{,}239.26$$

The current 1-year interest rate – the marginal income on each dollar collected – has to be used for pricing and evaluating the performance of the manager over each year separately. Single-year pricing and performance evaluation is optimal as it will lead to the highest value over the 2-year horizon.

In addition to pricing and performance evaluation, a third question to consider is hedging interest rate risk. What maturity assets should the money collected at the start of year 1 be invested in? A matched-maturity of one year, shorter or longer?

To understand the nature of interest rate risk, we run the following simulation. The current upward rising yield curve being at 4% - 4.971%, deposits in year 1 are priced optimally at 2.67%. But what happens next year if the 1-year-rate falls from 6% to 5%? Let us calculate the profit in year 2 in a lower rate environment:

$$d_2^{\text{OPTIMAL}} = 5.0\% \times \frac{1}{(1+\frac{1}{2})} = 5.0\% \times 0.6666 = 3.33\%$$

$$\text{Profit in year 2} = (5.0\% - 3.33\%) \times 100{,}000 \times 5.0^{-1.5} \times 3.33^2 = 1{,}656.35$$

When market interest rate in year 2 falls from 6% to 5%, profit in that year falls by 24% from 2,177.32 to 1,656.35. This is caused by the relative rigidity of the deposit rate. For a fall in market rates of 1% from 6% to 5%, the deposit rate fell by only 0.67% from 4% to 3.33%, which generated a fall in interest margin. To hedge against

10 In the deposits supply relations, an interest rate of 4% is entered as "4.0."
11 Deposits could be invested in 2-year maturity bond with coupon c. Arbitrage ensures that the two investment strategies yield the same return over two years.
12 An alternative hedging tool would be to invest in a one-year maturity bonds and purchase an interest rate futures contract that creates a gain in case of a fall in interest rates.

the fall in profits in year 2, the bank can increase the maturity of the bond purchased at the start of year 1 beyond one year to create a capital gain when interest rate[12] falls by an amount equal to the fall in profitability in Year 2.

The above example illustrates that the relevant maturity used for hedging interest rate risk does not need to be the same as the maturity used for pricing and performance evaluation. For hedging, one would use a maturity of asset longer than one year, while for pricing and evaluation of performance, one would use the marginal rate, the 1-year-maturity rate.

With regards to evaluating the performances of the business units, a practical question arises as to which business units should receive the benefit of the hedge (capital gains if interest rate falls): the ALM department or the commercial deposit gathering unit? If the objective is to insulate the commercial unit against the negative impact of a lower interest rate environment on profit one could allocate the benefits of the hedge to the commercial unit. However, we take the position that this should not be done for the following reason. Allocation of capital gains to the profit in year 2 would risk distorting the evaluation of profitability in that year. In a low interest rate environment, profit margins are smaller and the correct lower marginal transfer price should be recognized to create managerial incentives to be more efficient and to reduce costs, possibly reducing the number of branches and exiting some locations. This does not imply that managers of commercial units should be penalized with lower bonuses when they operate in a low interest rate environment. To achieve this and reward outstanding performance, bonuses should be based not on the absolute revenue of a business unit but on the difference between realized revenue and a benchmark target. Obviously, in a lower rate environment, the benchmarked target would be reduced.[13]

In the context of independence between the two years, the analysis leads to the following observations. One would use a matched maturity current rate for both pricing and evaluation of performances in years 1 and 2. Maximization of 1-year revenue leads to optimal decisions that maximize total value. Hedging the interest rate risk requires consideration of profits over the two years. Capital gains on the asset is needed in the case of a fall in the interest rates. An additional performance measurement issue relates to the allocation of profits and losses of the hedging strategies among the commercial units. We argue against this in order to ensure that there is a recognition of the current lower interest rate environment. To avoid penalizing commercial units in a low interest rate environment, performances would be compared to a more flexible benchmark target.

CASE 2: DEPENDENCE BETWEEN YEAR 1 AND YEAR 2

In Case 1, the decision taken in year 1 had no impact on the profit in year 2: there was complete separation between the two decisions. However, there are two potential reasons why independence might not be apply: (1) the volume of deposits in year 2 might be related to that collected in Year 1 and (2) the deposit rate applied in year 2 might be related to that set in year 1.

This situation is likely to be observed in retail banking markets with deposits with undefined maturity. Slow adjustments by depositors suggests that some customers will take their time in moving their deposits to another bank or financial product. This creates a time dependence between the volumes of deposits. Deposits in year 2 are partly related to what was collected in year 1. The second reason is that for marketing reasons (menu cost), one wants to avoid changing the interest rate on deposits too often. This creates a second type of dependence: the interest rate paid in year 2 is the rate chosen in year 1. Time dependence will force that bank to analyze the impact of the first year decision on the profit of the second year. Value maximization should be conducted on a multi-period basis.[14]

The stability of deposit volumes and the rigidity of the deposit rates suggest that the effective maturity of short-term deposits is longer than the contractual maturity. Hence we need to analyze how an effective maturity can be identified.

For banks, a standard practice when dealing with retail deposits with undefined maturity is to split them into two categories: volatile "transient" deposits and stable "core" deposits that are equivalent to long-term fixed rate maturity deposits. The fund transfer price is the weighted sum of the short-term and long-term interest rates, with the weights being the volume of volatile and stable deposits. Although one accepts the desire for a simple managerial rule, doubts may persist about the concept of an effective fixed-rate maturity. In the banking world the volume of sticky deposits is not completely fixed and the interest rate chosen in year 2 is not totally rigid. One objective of this paper is to specify the conditions in which the use of an effective long-term maturity can be justified, and to present a coherent value-maximization framework to deal with cases in which these conditions do not apply.

13 And if we leave the capital gains to the ALM department, this should not lead automatically to a bonus for ALM managers. Performance of the ALM department's mismatch strategy should be relative to that of a fully hedged strategy (which in this case would include the capital gains).

14 Although this issue is often discussed with reference to deposits, it applies as well to retail loans, such as consumer and credit card loans with relatively rigid interest rates.

Consider the case where the volume of deposits in Year 2, D_2 (.), is a function not only of the deposit rate paid that period, d_2, but also of the volume of deposits collected in year 1, D_1. Below, a log-linear specification, Case 2.1, is first analyzed, which allows us to compare with Case 1 using a similar specification for the deposit supplies.

Case 2.1: Log-linear dependence of the volume of deposits with flexible deposit rates

The supplies of deposits in years 1 and 2 are equal to:

$$D_1 = 100{,}000 \times b_1^{-1.5} \times d_1^2$$
$$D_2 = 300 \times b_2^{-1.5} \times d_2^2 \times D_1^{0.5}$$

The specification of the volume of deposits in year 2 is log-linear, in which case the elasticity of the volume of deposits in year 2 to deposits collected in year 1 is the exponent 0.5. It is assumed that the deposit rate in year 2 is flexible.

The maximization of the present value of future profits, evaluated at end of year 1, is equal to:

$$\text{MaximizeValue} = (4\% - d_1) \times D_1 + \frac{(6\%-d_2)\times D_2}{1+6\%}$$

Intuitively, one should pay a bit more to attract deposits in year 1 because this will increase the supply of profitable deposits in year 2. In the case of dependence overtime, one needs to work with dynamic optimization [Intriligator (1971)], which has two stages. First, compute the optimal deposit rate in year 2, the last period, and then identify the deposit rate in the first year that will maximize the present value of profits earned in years 1 and 2.

The optimal pricing in the last year, year 2, is identical to that of the case of independence:

$$d_2^{\text{OPTIMAL}} = 6.0\% \times \frac{1}{(1+\frac{1}{2})} = 6.0\% \times 0.6666 = 4.0\%$$

$$\text{Profit in year 2} = (6.0\% - 4.0\%) \times 300 \times 6.0^{-1.5} \times 4.0^2 \times D_1^{0.5}$$

Having computed the optimal deposit rate in year 2, one can then compute the optimal deposit rate in year 1 that maximizes value.

$$\text{MaximizeValue} = (4\% - d_1) \times D_1 + \frac{(6.0\%-4.0\%) \times 300 \times 6.0^{-1.5} \times 4.0^2 \times D_1^{0.5}}{1+6\%}$$

A closed-form solution for optimal pricing in year 1 is given in Appendix 3. Alternatively, one can use the function "optimizer" or "solver" in a spreadsheet to identify the deposit rate in year 1 that maximizes value. The optimal deposit rate in year 1 is 3.235 %, a rate significantly higher than that obtained under myopic optimization of 2.67%. Given the higher market rate of 6% accompanied by a higher margin in year 2, there is an incentive to attract more loyal "core" deposits in year 1. Relative to a myopic optimization, a dynamic optimization generates lower profits in year 1 (1,000.98 versus 1,185.19), but higher profits in year 2 (2,362.25 versus 1,947.46). The impact on the present value of profits over the two years is positive, 3,229.52 versus 3,022.41.

The above case shows that in a situation of intertemporal linkage – in our case deposits in year 2 are related but not identical to those of year 1 and the deposit rate in year 2 is not rigid – maximization over several periods is necessary and there is no simple concept for an effective maturity. One could artificially increase the FTP in year 1 to 4.852% to ensure that single-period optimization leads to the optimal deposit rate of 3.235% (= 4.852% x $(1 + 1/\varepsilon)^{-1}$). This "blown-up" FTP rate is lower than the 2-years-to-maturity coupon rate of 4.973%. It is equal to a weighted average of the 1-year rate b_1 and 2-year fixed coupon interest rate c with weights[15] of 12.25% and 87.75%, respectively. Hence, the market practice of applying weighted average rate to compute the FTP could be used but one must note that the weighting is sensitive to the market rates b_1 and b_2, the price elasticity and the time-dependence factor. This differs from the ad hoc practice of applying shares of volatile deposits and stable deposits; the reason being that long-term deposits are not completely sticky but sensitive to the second year interest rate and that the deposit rate is not constant.

Three partly related issues have been identified in the context of products with undefined maturity: pricing, evaluation of performance and hedging. The analysis of the log-linear case of volume dependence over time with flexible deposit rates leads to the following conclusions:

- **Pricing:** to achieve optimal pricing, one needs to conduct multi-period optimization. If a bank intends to maximize profits over one period, it could artificially increase the FTP to ensure optimal pricing in year 1, but the blown-up transfer price will not be equal to a two-year maturity fixed interest rate or a weighted average of short- and long-term rate; with the weights being the volumes of volatile and stable deposits.
- **Evaluation of performance:** optimal pricing is shown to reduce profit in year 1. Once again, "superior" performance should be evaluated against a benchmark, not in absolute terms. An alternative is to use a blown-up FTP.

15 The weights are obtained from the following relationship: effective FTP = 4.852% = ((1- α) xb1) + (α x c).

- **Hedging:** as was done for case 1, one needs to assess the impact of a change in interest rates on the present value of future profits to determine the hedged maturity for the assets.

In the following case (2.2), we assume that the volume of deposits in year 2 is identical to that of year 1 and that the deposit rate in year 2 is identical to that applied in year 1.

Case 2.2: Complete rigidity of the volume of deposits and of the deposit rate

The log-specification for the deposit volume in year 1 is identical to that of the first case but the deposits volume and the interest rate set in year 2 are identical to those of year 1. This is a case of complete rigidity of both volume and deposit rates.

$D_1 = 100{,}000 \times b_1^{-1.5} \times d_1^2$
$D_2 = D_1$
And $d_2 = d_1$

$$\text{MaximizeValue} = (4\% - d_1) \times D_1 + \frac{(6\%-d_1) \times D_1}{1+6\%}$$

In the case of complete rigidity (Appendix 4), the multi-period maximization problem can be converted into a one-period maximization, with the deposit rate in year 1 being priced against the two-year fixed coupon rate c (calculated in Appendix 1).

$$d_1^{OPT} = c \times \frac{1}{(1+\frac{1}{\varepsilon})}$$

Within the parameters of the example, the optimal deposit rate in the case of constant volume and deposit rate is equal to:

$$d_1^{Optimal} = d_2 = 4.971\% \times \frac{1}{(1+\frac{1}{2})} = 3.31\%$$

With reference to the example in Lebanon, using a 5-year 7% fixed rate to price short-term deposits is only warranted if the volume of deposits is constant and the deposit rate is fixed. This is an extreme situation since in a period of rising interest rates, the case of a positive yield curve, one can anticipate an increase in the deposit rate in the future driven by competition with a corresponding impact on volume. The case of extreme rigidity of both interest rates and volumes is very unlikely in reality, with the consequence that one cannot rely on maximization over one year with an effective maturity interest rate. A more complex multi-period maximization is needed.

One could argue that the above results are due to the log-linear specification in year 2. This specification does not allow for segmentation in year 2 between the "old" loyal deposits collected the previous year and the "new" deposits collected in year 2. Such a segmentation would allow year 1-deposits to be treated as quasi longer-term 2-year-to-maturity deposits. A linear-additive specification is introduced to allow for such segmentation. Again, the purpose of the analysis is to identify conditions that allow the multi-period maximization to be simplified into an effective fixed maturity problem.

Case 2.3: Linear additive volume dependence and fixed deposit rates

The volume of deposits in year 2 is made up of two components: a fraction of the deposits collected in year 1 (the loyal deposits) and new deposits (ND_2).

We consider two settings for pricing. In the first one, case 2.3, the deposit rate chosen in year 1 applies in the second year. In the second case, case 2.4, we allow price discrimination. Deposits collected in year 1 keep receiving the same deposit rate while new deposits (ND_2) received a rate set in year 2. Again, the objective is to understand the nature of the maximization over two years.

The bank maximizes the present value of future profits,

$$\text{Value} = (b_1-d_1) \times D_1 + \frac{\alpha D_1 \times (b_2-d_1) + (b_2-d_1) \times ND_2}{1 + b_2}$$

$$= (1-\alpha) \times D_1 \times (b_1-d_1) + \alpha D_1 \times (b_1-d_1) + \frac{\alpha D_1(b_2-d_1) + (b_2-d_1) \times ND_2}{1 + b_2}$$

$$= (1-\alpha) \times D_1 \times (b_1-d_1) + \alpha D_1 \times (c-d_1) + \frac{\alpha D_1(c-d_1) + (b_2-d_1) \times ND_2}{1 + b_2}$$

The last relationship follows from the arbitrage that ensures that investing stable deposits in a 2-years-to-maturity bond with coupon c is equivalent to investing in a 1-year asset with roll-over at the forward rate b_2. Value is the sum of two terms: profit in year 1, which is a weighted sum of profits on volatile (1 - α) and stable (α) 1-year deposits, and the value of profit on year 2-deposits.

The volume of deposits in years 1 and 2 are given by:

$\text{Deposits}_1 = 100{,}000 \times b_1^{-1.5} \times d_1^2$
$\text{Deposits}_2 = \alpha \times D_1 + ND_2 = \alpha \times D_1 + 100{,}000 \times (1-\alpha) \times b_2^{-1.5} \times d_1^2$

The above is related to Case 2.2 with two differences: only a fraction (α) of deposit in Year 1 will transfer to Year 2 and the constant deposit rate set in Year 1 will affect the new deposits collected in Year 2 (ND_2).

Using the function "optimizer" in a spreadsheet, one can identify the deposit rate in year 1 that maximizes value. Using as an example a retention rate, α, of 90%, the optimal deposit rate in year 1 is

3.298%. Compared to the extreme case of fixed volume and fixed deposit rates discussed above, two forces are at work: the retention rate of year 1 deposits is less than 100% ($\alpha < 1$) while there is a need to keep the deposit rate high enough to attract the new volatile deposits of year 2. Again, one could identify a fund transfer price that allows optimization over one year. The FTP equivalent is 4.948%, lower than the fixed coupon of 4.971% and higher than a weighted average (90% x 2-year coupon rate, 10% x 1-year rate) of 4.87%. Multi-period maximization is again warranted in this case.

In the final case, we allow discriminatory pricing, with the flexible deposit rate set in year 2 affecting only the new deposits.

Case 2.4: Linear-additive volume of deposits and discriminatory pricing

The deposit rate set in year 1 applies to the stable deposits that remain in year 2. New deposits collected in year 2 receive the rate d_2. The bank is said to apply discriminatory pricing between the "old" and "new" deposits ND_2. To circumvent laws that prohibit price discrimination, "revenue management" consulting companies advise the creation of new products targeted at a specific segment, the new depositors. This brings us closer to the case of a fixed interest rate and fixed volume of deposits. But there is a difference, as only a fraction (α) of the deposit collected in year 1 will transfer to year 2. The objective is to see whether the ad hoc rule of a weighted average of short- and long-term interest rate can be applied or not.

The bank maximizes the present value of future profits as follows:

$$\text{Value} = (b_1-d_1) \times D_1 + \frac{\alpha \times D_1 \times (b_2-d_1) + (b_2-d_2) \times ND_2}{1+b_2}$$

$$= (1-\alpha) \times D_1 \times (b_1-d_1) + \alpha\left[D_1 \times (b_1-d_1) + \frac{D_1 \times (b_2-d_1)}{1+b_2}\right] + \frac{(b_2-d_2) \times ND_2}{1+b_2}$$

$$= (1-\alpha) \times D_1 \times (b_1-d_1) + \alpha \times D_1 \times (c-d_1) + \frac{\alpha \times D_1 (c-d_1) + (b_2-d_2) \times ND_2}{1+b_2}$$

Value is the sum of two terms: profit in year 1, which is a weighted sum of profits on volatile (1 - α) and stable (α) 1-year deposits invested in respectively the 1-year asset and 2-year-fixed coupon asset, and the value of profit on year 2-deposits. As is shown in Appendix 5, the optimal interest rate for year 1 is given by the following relation:

$$d_1^{Optim} = \frac{w + \frac{\alpha \times c}{1+b_2}}{(1-\varepsilon^{-1}) \times (1 + \frac{\alpha}{1+b_2})} \text{ with } w = (1-\alpha)b_1 + \alpha c$$

$$= \frac{b_1(1+b_2) + \alpha b_2}{(1+\varepsilon^{-1}) \times (1+b_2+\alpha)}$$

At the optimum, the value of the marginal costs incurred over two years on one dollar of deposits collected in year 1 with a retention rate α in year 2 must equal the value of the marginal revenue earned over two years. The marginal revenue includes the weighted average return w earned in investing the volatile deposits in a 1-year bond and the loyal stable deposits in a 2-year bond, and the revenue earned on deposits retained in the second year.

$$d \times (1-\varepsilon^{-1}) \times (1 + \frac{\alpha}{1+b_2}) = w + \frac{\alpha c}{1+b_2}$$

In the general case, the maximization of value must be conducted over two years, and one cannot focus solely on the weighted average return w earned on volatile and stable deposits. The reason being that if the transient and loyal deposits can be invested in a weighted average of 1-year and 2-year assets, one cannot ignore the revenues and costs faced in year 2 in the value maximization.

Two special cases stand out. If $\alpha = 1$, the deposits collected in year 1 have effectively a 2-year fixed-rate maturity and the FTP becomes the two-year fixed coupon rate c. This situation is identical to that of case 2.2 with complete rigidity of both volumes of deposits and interest rates. If the yield curve is flat ($b_1 = b_2$), the two-year maximization simplifies into a one-year myopic optimization with FTP = $b_1 = b_2$. In all other situations, the single effective maturity FTP given by the optimal pricing rule is different from the two-year fixed coupon c or from a weighted average w of the 1-year- and 2-years-to-maturity rates b_1 and c.

The case is illustrated numerically with a specification for deposit supply function similar to that of case 2.3.

$$\text{Deposits}_1 = 100{,}000 \times b_1^{-1.5} \times d_1^2$$
$$\text{Deposits}_2 = \alpha \times D_1 + ND_2 = \alpha \times D_1 + 100{,}000 \times (1-\alpha) \times b_2^{-1.5} d_2^2$$

Dynamic optimization is applied. The optimal deposit rate in year 2 for the new deposits ND_2 is equal to the myopic case of 4%. The optimal deposit rate in year 1 that maximizes total value over the two years is 3.28%, less than the 3.31% fixed deposit rate case (case 2.2). As the retention of year 1-deposits is imperfect ($\alpha < 1$), there are fewer profitable deposits in year 2 and the deposit rate is reduced. The FTP equivalent to allow myopic one-period optimization is 4.92%, smaller than the 2-Year coupon rate c and different from the weighted average market rate of 4.87% with weights of 10% for the 1-year rate b1 and 90% for the two year coupon rate c.

In the case of linear-additive deposit supply function and price discrimination, one observes again that a multi-period maximization is needed. It cannot be readily converted into a single effective period maximization. Only in two cases would such a simplification be possible: extreme stickiness ($\alpha = 1$) or the case of a flat yield curve.

CONCLUSION

Demand and savings deposits are a significant source of funds for banks. These products with undefined maturity have generated heated debates on the effective maturity that should be applied to the fund transfer price used to evaluate their profitability. Furthermore, the 2016 Basel proposal for a capital regulation on interest rate risk on the banking book also raises the issue of the choice of a behavioral maturity for non-maturing deposits. This paper has presented a pedagogical yet rigorous value-based management approach to the management of deposit with undefined maturity. I have focused on three issues associated with the management of these deposits: pricing, performance evaluation, and hedging of interest rate risk. Such deposits raise the question of their effective behavioral maturity. I have identified the conditions in which the multi-period maximization problem can be converted into one with a single effective maturity, and I have evaluated the market use of a weighted average maturity obtained by breaking down the portfolio of deposits with undefined maturity into buckets with short-term volatile deposits and stable longer-term deposits.

Managing deposits with undefined maturity is a multi-period problem as there are two intertemporal issues: the volume of deposits in year 2 is related to the deposits collected in year 1 and the deposit rate can be relatively rigid. Under most assumptions analyzed here, the management issue cannot be simplified into a profit maximization over a single period. Multi-period maximization simplifies into maximization over one period in two extreme cases: constant deposit volume/deposit rate or flat yield curve with price discrimination. Since these conditions are unlikely to be met, one needs to work with a more complex multi-period optimization.

I have also shown that the maturity of assets needed to hedge the bank against interest rate risk can be different from the fund transfer price maturity used for pricing or for measuring performance. The reason for this is that the management of interest rate risk entails an analysis of the impact of interest rates on all future profits (the franchise value), while the maturity relevant for pricing is the period over which intertemporal dynamics apply.

REFERENCES

- Bank Audi, 2013, Annual report
- Basel Committee on Banking Supervision, 2015, "Interest rate risk in the banking book," consultative document, June, 1-61.
- Dermine J., 2007, "ALM in banking," in Zenios, S. A., and W. T. Ziemba (eds.), Handbook of asset and liability management, Volume 2, North Holland series Handbooks in Finance, Elsevier Science B.V.
- Dermine J., 2013, "Fund transfer pricing for deposits and loans, foundation and advanced," Journal of Financial Perspectives, Vol. 1:1, March.
- Dermine, J., 2015, Bank valuation and value-based management. Deposit and loan pricing, performance evaluation and risk management, 2nd edition, McGraw-Hill, N.Y.
- European Banking Authority, 2015, "Guidelines on the management of interest rate risk arising from non-trading activities," 1-77.
- Intriligator, M. D., 1971, Mathematical optimization and economic theory, Prentice-Hall Series in Mathematical Economics, Upper Saddle River: Prentice Hall.

APPENDICES

Appendix 1. Computation of 2-year fixed coupon rate c consistent with 1-Year rates b_1 and b_2.

In equilibrium, the value of a short term investment strategy with roll-over must equal the value of investing in a bond with fixed coupon c.

$$\frac{b_1}{1+b_1}+\frac{1+b_2}{(1+b_1)(1+b_2)}=\frac{c}{1+b_1}+\frac{1+c}{(1+b_1)x(1+b_2)}$$

$$(1+b_2)xb_1+1+b_2=c(1+b_2)+c+1$$

$$(1+b_2)x(1+b_1)=c(2+b_2)+1$$

$$c=\frac{(1+b_2)(1+b_1)-1}{2+b_2}=\frac{b_2+b_1x(1+b_2)}{2+b_2}$$

Appendix 2. Deposit pricing[16]

Given a supply of deposits D(d), a positive function of the deposit rate d, and denoting by b and ε the market rate and the deposit volume price-elasticity, one has: Revenue = (b - d) x D(d)

To maximize revenue, one has:

$$\frac{\partial \text{Revenue}}{\partial d}=-D+(b-d)D'=0$$

$$-\frac{Dxd}{D'xd}+(b-d)=0$$

$$-d\,x\,\varepsilon^{-1}+b-d=0$$

$$d\,x(1+\varepsilon^{-1})=b$$

$$d^{Optimal}=b\,x\frac{1}{1+\frac{1}{\varepsilon}}$$

In the case of a log-linear supply of deposits, the price elasticity (ε) is the exponent of the deposit rare variable:

$$Deposits_t=\alpha_0\,x\,d_t^{\alpha_1}\,x\,b_t^{-\alpha_2}$$

$$Elasticity\,(\varepsilon)=\frac{\partial \log D}{\partial \log d}=\alpha_1$$

Appendix 3. Log-linear dependence with flexible deposit rates (Case 2.1)

$$D_1=100{,}000xb_1^{-1.5}xd_1^{\,2}$$

$$D_2=300xb_2^{-1.5}xd_2^{\,2}xD_1^{\,0.5}$$

$$Value=(b_1-d_1)D_1+\frac{(b_2-d_2)D_2(d_2,D_1)}{(1+b_2)}$$

The dynamic optimization starts with maximization of profits in Year 2 and the choice of the deposit rate in Year 2

$$-D_2+(b_2-d_2)D_2'=0$$

$$-\frac{D_2xd_2}{D_2'xd_2}+(b_2-d_2)=0$$

$$-d_2x\varepsilon^{-1}+b_2-d_2=0$$

$$d_2^{\,Optimal}=b_2x\frac{1}{1+\frac{1}{\varepsilon}}$$

Having calculated the deposit rate in Year 2, one can then calculate the optimal deposit rate in Year 1

$$-D_1+(b_1-d_1)D'1+\frac{(b_2-d_2)\frac{\partial D_2}{\partial D_1}x\frac{\partial D_1}{\partial d_1}}{(1+b_2)}=0$$

$$define\,D'_2=\frac{\partial D_2}{\partial D_1}$$

$$-\frac{D_1}{D'_1}x\frac{d_1}{d_1}+(b_1-d_1)+\frac{(b_2-d_2)D'_2}{(1+b_2)}=0$$

$$d_1(1+\varepsilon^{-1})=b_1+\frac{(b_2-d_2)D'_2}{(1+b_2)}$$

$$d_1^{\,Optimal}=\left[b_1+\frac{(b_2-d_2)D'_2}{(1+b_2}\right]x(1+\varepsilon^{-1})^{-1}$$

$$with\,D'_2=0.5x300xb_2^{-1.5}xd_2^{\,2}xD_1^{\,0.5-1}$$

16 Discussed in Chapter 11 of Dermine (2015).

Appendix 4. Complete rigidity of the volume of deposits and of the deposit rate (case 2.2)

The log-linear specification for the deposit volume in year 1 is given below. The deposits volume and the interest rate set in year 2 are identical to those of year 1. This is a case of complete rigidity of both volume and deposit rate.

$$D_1 = 100{,}000 x b_1^{-1.5} x d_1^{\ 2}$$

$$D_2 = D_1$$

$$and\ d_2 = d_1$$

$$Value = (b_1 - d_1) x D_1 + \frac{(b_2 - d_1) D_1}{1 + b_2}$$

$$\frac{\partial Value}{\partial d_1} = -D_1 + (b_1 - d_1) D'_1 + \frac{-D_1 + (b_2 - d_1) D'_1}{1 + b_2} = 0$$

$$= -\frac{D_1 d_1}{D'_1 d_1} x (b_1 - d_1) + \frac{-\frac{D_1 d_1}{D'_1 d_1} + (b_2 - d_1)}{1 + b_2} = 0$$

$$= b_1 - d_1 x (1 + \varepsilon^{-1}) + \frac{b_2 - d_1 x (1 + \varepsilon^{-1})}{1 + b_2}$$

$$= b_1 + \frac{b_2}{1 + b_2} - d_1 x (1 + \varepsilon^{-1}) x (1 + \frac{1}{1 + b_2}) = 0$$

$$d_1 x (1 + \varepsilon^{-1}) = (\frac{b_1 (1 + b_2) + b_2}{1 + b_2}) x (\frac{1 + b_2}{2 + b_2})$$

$$= \frac{b_1 (1 + b_2) + b_2}{2 + b_2} = c$$

Appendix 5. Linear-Additive Volume of Deposits and Discriminatory Pricing (Case 2.4)

The deposit rate set in year 1 applies to the deposits that remain in year 2. The new deposits collected in year 2 receive the rate d_2. The present value of future profits is as follows:

$$Value = (b_1 - d_1) x D_1 + \frac{\alpha x D_1 x (b_2 - d_1) + (b_2 - d_2) x D_2}{(1 + b_2)}$$

$$= (1 - \alpha) x D_1 x (b_1 - d_1) + \alpha \left[D_1 x (b_1 - d_1) + \frac{D_1 x (b_2 - d_1)}{1 + b_2} \right] + \frac{(b_2 - d_2) x D_2}{1 + b_2}$$

$$(1 - \alpha)(b_1 - d_1 (1 + \varepsilon^{-1})) + \alpha \left[(b_1 + \frac{b_2}{1 + b_2} - d_1 (1 + \varepsilon^{-1}) x (1 + \frac{1}{1 + b_2}) \right] = 0$$

$$-d_1 x (1 + \varepsilon^{-1}) x (1 - \alpha + \alpha x (1 + \frac{1}{1 + b_2}) + (1 - \alpha) b_1 + \alpha x (b_1 + \frac{b_2}{1 + b_2}) = 0$$

$$-d_1 x (1 + \varepsilon^{-1}) x (1 + \frac{\alpha}{1 + b_2}) + (b_1 + \frac{\alpha b_2}{1 + b_2}) = 0$$

$$d_1^{\ Optim} = \frac{b_1 (1 + b_2) + \alpha b_2}{(1 + \varepsilon^{-1}) x (1 + b_2 + \alpha)}$$

An alternative derivation of the last two relations is as follows:

$$-d_1 x (1 + \varepsilon^{-1}) x (1 + \frac{\alpha}{1 + b_2}) + (b_1 + \frac{\alpha b_2}{1 + b_2}) = 0$$

$$d_1 x (1 + \varepsilon^{-1}) x (1 + \frac{\alpha}{1 + b_2}) + (1 - \alpha) b_1 + \alpha b_1 + \frac{\alpha b_2}{1 + b_2} = 0$$

$$d_1 x (1 + \varepsilon^{-1}) x (1 + \frac{\alpha}{1 + b_2}) + (1 - \alpha) b_1 + \alpha c + \frac{\alpha c}{1 + b_2} = 0$$

$$d_1 x (1 + \varepsilon^{-1}) x (1 + \frac{\alpha}{1 + b_2}) + w + \frac{\alpha c}{1 + b_2} = 0$$

$$d_1^{\ OPT} = \frac{w + \frac{\alpha c}{1 + b_2}}{(1 + \varepsilon^{-1}) x (1 + \frac{\alpha}{1 + b_2})} \quad with\ w = (1 - \alpha) x b_1 + \alpha x c$$

Delegated Portfolio Management, Benchmarking, and the Effects on Financial Markets

Deniz Igan – Deputy Chief, Macro-Financial Division, Research Department, IMF
Marcelo Pinheiro – Senior Financial Economist, Center for Economic Analysis, PCAOB[1]

HIGHLY TECHNICAL

Abstract
We analyze the implications of linking the compensation of fund managers to the return of their portfolio relative to that of a benchmark – a common solution to the agency problem in delegated portfolio management. In the presence of such relative-performance-based objectives, investors have reduced expected utility but markets are typically more informative and deeper, provided that information is free. Furthermore, in a multiple asset/market framework, we show that (i) relative performance concerns lead to financial contagion; (ii) benchmark inclusion increases price volatility; and (iii) home bias emerges as a rational outcome. When information is costly, however, information acquisition is hindered and this attenuates the effects on informativeness and depth of the market.

1 The views expressed here are those of the authors and do not necessarily represent those of the International Monetary Fund or its policy, and do not necessarily reflect the views or opinions of the Public Company Accounting Oversight Board, Board members, or members of the staff. We would like to thank Franklin Allen, Markus Brunnermeier, Giovanni Dell'Ariccia, Aureo de Paula, Gaston Gelos, and Helene Rey for helpful discussions. All remaining errors are our own.

INTRODUCTION

An interesting omission in standard portfolio choice theory is that professional portfolio managers, such as mutual, pension, and hedge fund managers, are ignored and investors are assumed to directly manage their own portfolios. This assumption depicted the equity market in the mid-1950s well but it no longer does: direct holdings of corporate equity by the household sector has fallen from 99 percent in 1945 to 47 percent in 2010.[2]

Recently, two related strands of literature have aimed to answer some of the questions raised by the issue of delegated portfolio management and the associated agency problems in modern markets. The first of these strands concentrates on the determination of optimal contracts to solve the agency problem between the portfolio manager and the investor. The optimal contract is often depicted as one that rewards performance relative to a benchmark. The other strand investigates whether relative performance concerns might lead a manager to change their fund's risk exposure and if this behavior is linked to the manager's compensation package.

In this paper, we seek to expand the insights offered in the literature along the second strand. Our focus is not on optimal contracts. A rather expansive strand of the literature focuses on optimal contracts that solve/mitigate the agency problem inherit in delegated portfolio management.[3] We do not directly analyze this problem. Instead, we take the optimal contract as given, carrying the properties commonly observed in practice [see, for instance, Elton et al. (2003)]. Note that the form of relative performance objectives can be flexible. These objectives do not have to be spelled out in formal contracts but may emerge as a result of competitive forces and/or behavioral factors (for instance, keeping up with the Joneses).[4] Investors usually base their investment decisions on fund performance and typically choose funds that have high returns compared to similar funds [see, among others, Chevalier and Ellison (1997) and Sirri and Tufano (1998)]. This behavior, combined with the usual fee structure in a fund (charged as a percentage of funds), leads to relative performance objectives for fund managers: you are only doing well in as much as you are doing better than your competitor. Indeed, in a formal model setting, Palomino (2005) shows that if investors use a relative performance rule to evaluate funds and allocate money into them, and managers receive an asset-based compensation, then this translates into relative performance objectives for fund managers. So, our assumption of relative performance contracts prevailing in the asset management industry is mostly a modeling short-cut to be interpreted as a reduced form of this more elaborate situation.

With the optimal contract reflecting relative performance concerns given, we analyze the effects of the adoption of relative performance pay for fund managers. Differently from the previous literature, we move away from questions about risk exposure and instead explore other interesting implications for market microstructure. First, we show that if fund managers are rewarded based on a relative performance measure, there can be deleterious effects for investors. More precisely, investors will have a lower expected utility. However, markets will typically be more informative and deeper, provided that information is free. Second, we observe an increase in the correlation between markets. That is, we show that the higher the importance of relative performance, the higher the correlation between two separate markets. And, we show that a stock that is included in the benchmark experiences an increase in its price volatility. Finally, information acquisition may be hindered by these types of contracts. This last effect may, ex ante, reduce the informativeness and depth of the market. So, even though conditional on information being acquired markets function better, less information acquisition may actually emerge and act in a countervailing way.

In the model, there are two investment funds in a market for an asset. Each fund manager observes a signal concerning the profitability of the risky asset. Markets function like in Kyle (1985), with the difference that we have two insiders, signals are noisy and, more importantly, each manager's payoff depends on their performance, as well as on the performance of the other manager. Fund owners reward the managers, the fund investors, and keep the rest to themselves. We analyze standard measures of market efficiency, and how the fund investors are affected by the presence of this relative performance objective function. We also analyze how the managers' own well-being is affected by accepting such contracts.

We extend the model to allow for two separate markets for two assets, and show that the correlation across markets is increasing in the measure of relative performance. In other words, contagion effects may appear, or be exacerbated by the presence of such contracts. In a similar vein, with two assets but both traded in the same market, we show that inclusion of a stock in the benchmark leads its price to experience higher price volatility. And, in another

2 Based on Flow of Funds data from the Federal Reserve Board (Table B.100.e). Note that these figures actually understate the "influence" of financial intermediaries on the individual portfolio decisions since, according to Investment Company Institute, 67 percent of individual investors seek services of financial advisors when making their decisions.

3 This agency problem has been discussed in a number of papers, e.g., Starks (1987), Heinkel and Stoughton (1994), Admati and Pfleiderer (1997), Das and Sundaram (1998a, b), Ou-Yang (2003), and Dybvig et al. (2010).

4 We use the words "objective," "contract," and "concern" almost interchangably throughout.

extension, we show how relative performance contracts may lead to a form of home bias.

Finally, we allow the choice to be informed to be endogenous, with costly signals. Specifically, we allow for two different information structures: public or private costly information. Under the first structure, either all fund managers pay the cost of acquiring information and all become informed or none of them do so. Under the second structure, each manager may decide independently. We show that, probabilistically, an increase in the importance of relative performance reduces the availability of informed equilibrium. That is, the more the managers' pay depend on relative performance, the lower the probability that either one decides to acquire information. Therefore, even though conditional on having informed agents the informativeness of prices is increasing on the importance of relative performance, we see that ex ante this effect may be attenuated or canceled by the fact that the chance of having an informed equilibrium is reduced. This countervailing effect is also present when analyzing the expected price response to trades (a measure inversely related to market depth).

Our paper relates and contributes to the literature in several ways. Some papers study the optimal portfolio strategy of a manager receiving fees that depend on relative performance [see, among others, Grinblatt and Titman (1989), Brown et al. (1996), Chen and Pennacchi (1999), Eichberger et al. (1999), and Goriaev et al. (2000)]. The general focus in these studies is on the effects on portfolio riskiness, with results showing that managers tend to over-bear risk in order to beat the benchmark, especially when they are behind in the game. Our model delivers a similar result, although in a slightly different manner (and this result is not central to our paper). Since we adopt a Kyle-type framework, what we show is that managers trade more aggressively, that is, put more weight on their signal, as the importance of relative performance increases. Since, for a given signal, managers increase their trading activity, this can be loosely interpreted as an increase in risk-taking behavior. The literature has also provided empirical evidence of this type of behavior, hence we take this as an already tested implication of our model.[5]

In a related paper, and as mentioned earlier, Palomino (2005) develops a model that indirectly generates the same type of objectives for managers and shows that, with entry and exit by funds (with exogenously determined, differential ability), these objectives may lead to survival of better-quality funds and, hence, higher expected returns for investors. We differ in our approach by assuming a fixed market structure (no entry), homogenous ability of funds with endogenous informational asymmetries and by analyzing the effects of varying the level of relative vis-à-vis absolute compensation on investors as well as on other market variables. Our results on investors' returns are in clear contrast to his. Since the modeling approaches are different, this, in and of itself, does not mean any immediate conflict. But it does point to the need of understanding the different forces at play: differential ability and information acquisition. In Palomino (2005), better-quality funds have a competitive advantage in acquiring information, and hence in equilibrium the market is populated by more informed individuals providing services to investors. In our setting, no fund has an edge and hence there is no inclination to have a more informative market. So, more concern about relative performance here leads to aggressive trading strategies and lower expected returns for investors, while in Palomino (2005) it leads to more high-quality funds entering the market (with better information) and hence investors having higher returns. However, in contradiction with empirical evidence on market efficiency [e.g., Berk and Green (2004)], Palomino's work seems to lead to a mutual fund industry that can outperform the market (passive strategy), while our model does not suffer from such a counter-intuitive implication.[6] Finally, we also discuss home bias, contagion, and addition to benchmarks, topics not discussed in Palomino (2005).

In another related paper, Cuoco and Kaniel (2011) examine the effects of delegated portfolio management on equilibrium asset prices assuming a contract for fund managers that is a generalization of the one postulated here. The main differences between the model here and their work rest on the assumed market structure and objectives. They analyze a problem under perfect competition concentrating on the potential for price effects on the stocks that are included in the benchmark.[7] They show that stocks that are included in the benchmark tend to have higher and more volatile prices than otherwise identical stocks. The analysis in this paper is under imperfect competition and, while replicating this volatility effect, aims at examining the effects of such contracts on the market microstructure, in terms of liquidity, price informativeness, cross-market correlation and information acquisition. To the best of our knowledge, the results describing these effects are specific to this paper and have not been analyzed in the same setting elsewhere.

5 See Goriaev et al. (2000) and references therein. Some studies argue that increased risk aversion leads managers to retrench and to more closely mimic their competitors' portfolios. For instance, Broner et al. (2006) empirically show that when funds' returns are below average they adjust their holdings toward the average (or benchmark) portfolio. The seeming contradiction is likely a reflection of threshold effects. Shelef (2013), for example, provides theoretical and empirical evidence that risk taking is non-monotonic: managers who are very distant from the incentive threshold take less risk than those who are less distant, but on average risk taking increases.

6 Results on price informativeness and trading aggressiveness are similar across both papers.

7 Brennan (1993) also considers the effects of such contracts on equilibrium expected returns and prices.

From a regulator's perspective, the insights presented here open a range of questions. Most notably, there can be a trade-off between the benefits of aligning the principal's interests with those of the agent and the potential costs associated with the effects on contagion, volatility and the informativeness of markets. What type of frameworks (relating to, for example, fund managers' compensation rules, investment restrictions, and disclosure requirements) could then limit contagion and excessive volatility and maintain informativeness of markets? We leave it to future research to examine these more normative questions and the optimal ways to balance this trade-off.

The rest of this paper is organized as follows. Section 2 describes the basics of the model. In Section 3, we analyze the effects of the presence of a relative performance-type contract for fund managers in a market where they have free access to information and derive the main results. In Section 4, we study the effects of relative performance on the decision to acquire information and discuss the differences that emerge from endogeneizing the information acquisition decision. All proofs are available from the authors. Section 5 concludes.

THE BASIC MODEL

Think of two mutual funds in a market for an asset.[8] Each fund manager observes a signal concerning the profitability of this risky asset.[9] Markets function like in Kyle (1985), with the difference that we have two insiders/managers and signals are noisy. Formally, the risky asset final payoff is given by

$$\tilde{v} \sim N(0,\sigma_v^2) \tag{1}$$

Each manager observes a noisy signal about the final payoff

$$\tilde{s}_i = \tilde{v} + \tilde{\varepsilon}_i,\ i = 1,2, \tag{2}$$

with $v \perp \varepsilon_i$, and where the noise terms are jointly normal with $Var(\tilde{\varepsilon}_i) = \sigma_\varepsilon^2$. Furthermore, assume that the necessary noise in the market comes from liquidity trades given by $\tilde{u} \sim N(0,\sigma_u^2)$.

Each manager's payoff depends directly on his performance relative to a benchmark. We postulate the following form for manager i's objective:

$$\max_{\alpha_i} E\{\phi[\alpha_i(v-P)-\gamma\alpha_j(v-P)]|\ s_i\}, \tag{3}$$

where P is the price of the asset, α_i denotes the portfolio weights, ϕ is just a scaling factor, and γ represents how much his payoff depends on relative performance.[10] Fund managers' concern with relative performance such as this may arise if the fund uses a "fulcrum fee" or if they are rewarded based on the size of their fund and consumers invest more in funds that perform better.[11]

Fund owners pay the managers as well as the investors. We assume that a proportion ς of fund profits goes to investors.[12] Therefore, the total expected cost for the owners of fund i is

$$\varsigma E[\alpha_i^*(v-P)] + \phi E[(\alpha_i^* - \gamma\alpha_j^*)(v-P) = (\varsigma + \phi(1-\gamma))E[\alpha_i^*(v-P)], \tag{4}$$

where α_i^* represents the solution to manager i's problem (the optimal portfolio weights). The equality follows from the fact that we concentrate on a symmetric equilibrium and that the signals are independent and identically distributed. Given this structure of contracts, a natural assumption is that $\sigma+\phi(1-\gamma)<1$. This guarantees that owners are making positive expected profits:

$$E[\alpha_i^*(v-P)] - (\varsigma + \phi(1-\gamma))E[\alpha_i^*(v-P)] > 0. \tag{5}$$

Finally, we assume that investors are risk-neutral and would like to maximize expected profits: $\varsigma E[\alpha_i(v-P)]$

Prices are determined by a risk-neutral competitive market maker that only observes aggregate market orders. As usual, the market maker is held down to a zero profit condition that translates into semi-strong efficiency of the market:

$$P = E[v|\alpha_1 + \alpha_2 + u]. \tag{6}$$

FREE INFORMATION

In this section, we assume that the fund managers are endowed with rights to the information provided by the signals. Basically, they have free access to the signal, s_i. Note that the signal can be publicly or privately provided, the key assumption here is that the information is free. Hence, the managers' decision to acquire information is trivial: they always choose to be informed.

8 This is without loss of generality. We could allow for more funds and results would be qualitatively the same.

9 This signal may or may not be costly. We analyze both cases.

10 This payoff function is a special case of the more general one analyzed in Cuoco and Kaniel (2011).

11 See, for instance, Palomino (2005).

12 Since we have a static model, the fund payoff will also be the amount of money in the fund after the realization of the random variables. Hence, paying ς to investors is the same as charging a percentage fee on the amount of funds, i.e., the fee is $1 - \varsigma$, a common practice.

We divide the section into two subsections. In the first, we restrict attention to one market and present the results on investor's utility and market parameters' behavior as a function of the relative performance contract. In the second subsection, we allow for the existence of two separate markets and analyze the issue of cross-market contagion.

Single market

With this basic set-up at hand, we have the following proposition.

Proposition 1: a symmetric linear equilibrium of the depicted market is defined

$$P = \lambda(\alpha_1 + \alpha_2 + u),\ \alpha_i = \beta\tilde{s_i} \tag{7}$$

Where

$$\beta = \sqrt{\frac{\sigma_u^2}{2(1-\gamma)\sigma_v^2+2\sigma_\varepsilon^2}} \text{ and } \lambda = \frac{2\sigma_v^2\sqrt{\frac{\sigma_u^2}{2(1-\gamma)\sigma_v^2+2\sigma_\varepsilon^2}}}{(2\sigma_v^2+\sigma_\varepsilon^2)\frac{\sigma_u^2}{(1-\gamma)\sigma_v^2+\sigma_\varepsilon^2}+\sigma_u^2} \tag{8}$$

Proof: contact authors.

Using these prices and portfolio shares, we can now calculate the ex-ante expected profit of fund i's investors, as a function of γ.

Corollary 1: ex-ante, investors expect to receive

$$E[\pi(\gamma)] = \varsigma E[\alpha_i(v\text{-}P) = \varsigma\sigma_v^2\sqrt{\frac{\sigma_u^2}{2}}\ \left[\frac{\sqrt{\sigma_\varepsilon^2+(1-\gamma)\sigma_v^2}}{2(\sigma_v^2+\sigma_\varepsilon^2)+(1-\gamma)\sigma_v^2}\right] \tag{9}$$

from their investments into fund $i \in \{1, 2\}$.

And, we can easily prove the following result.

Corollary 2: investors are worse off when the relative performance pay component is higher.

Proof: contact authors.

Furthermore, in the terminology of Kyle (1985), we know that the depth of the market can be represented as $1/\lambda$ and that price informativeness is a decreasing function of Var[v|P]. Then, as another corollary to Proposition 1, we have:

Corollary 3: if the importance of relative performance on fund managers pay is increased:

1. Liquidity of the markets increases;
2. The information content of prices increases;
3. Adverse selection costs faced by noise (liquidity) traders decrease;
4. Trading aggressiveness of fund managers increases.

Proof: contact authors.

Intuitively, the more a manager cares about relative performance, the more aggressively they trade on a given signal to beat the benchmark while in anticipation that their competitor will do the same and push the benchmark up. As a result, prices more closely reflect the available information set. This in turn reveals information to noise traders and reduce their costs,[13] allowing more liquid markets. The payoff to investors,[14] however, is lower because there is less payoff from informed trading.

To summarize, the adoption of relative performance contracts has positive and negative effects. Investors are worse off but markets function better, under free information acquisition. In the next section, we show that some of these apparent, positive effects of relative performance may disappear once we endogeneize the decision to become informed. But, before we move on, we present the results for multiple assets/markets, analyzing the possibility of cross-market contagion and volatility effects.

Multiple markets

Here, we extend the model to allow for multiple assets and markets. First, we analyze the effects of fulcrum fees on cross-market correlations. Then we look at the effect of including a stock in the benchmark.

Cross-market correlation

Suppose there are two separate markets for two distinct assets. Each market maintains the structure discussed in Section 2, and they are separated by the assumption that each pair of funds trades in only one of them. We assume, without loss of generality, that funds 1 and 2 trade on asset/market 1, and that funds 3 and 4 trade on asset/market 2. Each fund is benchmarked only against the other fund that trades in the same market, and each market has a potentially different γ. Let γ_m be the relative performance parameter

13 Note that this is a zero-sum market and the gains obtained through informed trading (total profits of both funds) is the losses of noise traders.

14 We show later that fund managers are worse off as well.

for each market m=1,2.[15] Furthermore, each market maker observes only the aggregate order flow of his own market. This is an important assumption because we assume that the assets' payoffs are correlated, therefore order flow in one market offers information about the payoff of the asset traded in the other market. So, there are two risky assets with final payoffs given by

$$\tilde{v}_m \sim N(0, \sigma^2_m) \tag{10}$$

and signals given by

$$\tilde{s}_{im} + \tilde{v}_m + \tilde{\varepsilon}_{im}, \; im = 11, 21, 32, 42 \tag{11}$$

where m index markets/assets and i index managers, with the understanding that managers 1 and 2 trade in the market for asset 1, and managers 3 and 4 trade on asset 2. We also assume that signals have the same precision, so that $Var(\tilde{\varepsilon}_{im}) = \sigma^2_{\varepsilon m}$, and liquidity trades are given by $\tilde{u}_m \sim N(0, \sigma^2_{um})$. Finally, we assume that all random variables are normally distributed and independent of each other, with the exception of $\tilde{v}_1$ and $\tilde{v}_2$, where we assume that $Cov(\tilde{v}_1, \tilde{v}_2) = \sigma^v_{12} > 0$. This corresponds to the assumption mentioned above that assets' payoffs are correlated.

Given this structure, it is clear that the issue at hand is not one of contagion per se. We do not investigate the question of appearance of correlation across assets that are otherwise uncorrelated. The objective is to analyze the marginal effect of relative performance contracts on cross-market correlation. Even if $\gamma_m = 0, \forall m$, we would have correlated prices, because payoffs are correlated. With the correlation recognized, we show how the presence of such contracts may enhance this correlation. That said, also note that order flow in one market reveals information in the other market but the market maker in each market can observe only the aggregate order flow in their own market. Hence, we shut down a trivial channel through which contagion between the two markets can be amplified and show how relative performance concern can increase such contagion even when the two markets operate separately.

Proposition 2: with the described set-up for each market, prices and portfolio shares given by

$$P_m = \lambda_m(\alpha_{1m} + \alpha_{2m} + u_{1m}) \text{ and } \alpha_{im} = \beta_m \tilde{s}_{im}, \tag{12}$$

Where

$$\beta_m = \sqrt{\frac{\sigma^2_{um}}{2(1-\gamma_m)\sigma^2_m + 2\sigma^2_{\varepsilon m}}} \text{ and } \lambda_m = \frac{2\sigma^2_m \sqrt{\frac{\sigma^2_{um}}{2(1-\gamma_m)\sigma^2_m + 2\sigma^2_{\varepsilon m}}}}{(2\sigma^2_m + \sigma^2_{\varepsilon m})\frac{\sigma^2_{um}}{(1-\gamma_m)\sigma^2_m + \sigma^2_{\varepsilon m}} + \sigma^2_{um}} \tag{13}$$

characterize a linear symmetric equilibrium in each market m. Given this characterization, we can calculate the correlation between P1 and P2 to be equal to

$$\rho = \frac{4\beta_1 \beta_2 \sigma^v_{12}}{\sqrt{(4\beta_1^2\sigma_1^2 + 2\beta_1^2\sigma^2_{\varepsilon 1} + \sigma^2_{u1})(4\beta_2^2\sigma_2^2 + 2\beta_2^2\sigma^2_{\varepsilon 2} + \sigma^2_{u2})}} \tag{14}$$

And,

$$\partial\rho/\partial\gamma_m > 0, \forall m. \tag{15}$$

Proof: contact authors.

This proposition shows that the presence of relative performance contracts may enhance correlation between markets. We observe more correlated prices whenever at least one of the markets has funds with relative performance contracts. This result does not depend on the assumption that both markets have relative performance contracts. It is enough that one of the markets has this type of contract. Then, if the importance of relative performance is increased in such a market, the correlation across markets increase. The intuition for this result comes, again, from the fact that aggressiveness increases with relative performance. An increase in the aggressiveness with which traders use their signals leads prices to reflect these signals more prominently. Since the signals are functions of the correlated final values, the importance of this correlation increases, leading to more correlated prices.

Benchmark inclusion

Next, we consider a slightly modified version of the model. As before, there are two risky assets whose final payoffs are given by expression (10) but, this time, with $Cov(\tilde{v}_1, \tilde{v}_2) = \sigma^v_{12} = 0$. Furthermore, assume that we have two funds, both trading on these assets, and let each manager have perfect signals about both assets, $\tilde{\varepsilon}_{im} \equiv 0$. So, we have two independent assets both of which are traded by both funds. (Previously, we have considered one asset traded by two funds, and then two assets each of which is traded by two funds). This generalization naturally leads us to re-define manager i's objective as

15 This assumption would be justified if the chosen benchmarks reflects the characteristics or "style" of a fund. This is indeed the case in a theoretical setting [e.g., Ou-Yang (2003)], but it is also true, at least to a certain extent, in practice as most funds use a market index (such as the S&P 500 or Russell) that is adjusted on size and value/growth dimensions to the fund's characteristics [Sensoy (2009)]. Note that in an international setting the use of domestic benchmarks, i.e., performance of those trading in the same market, would also be justified by the potentially higher costs faced by investors if they choose a foreign fund over a domestic one, perhaps due to cross-border differences in regulations and institutional and legal frameworks.

$$\max_{\alpha_{il},\alpha_{i2}} E\{\varphi[\alpha_{i1}(v_1 - P_1) + \alpha_{i2}(v_2 - P_2) - \gamma_1\alpha_{j1}(v_1\text{-}P_1) - \gamma_2\alpha_{j2}(v_2\text{-}P_2)] \mid v_1, v_2\} \ . \quad (16)$$

Since we have independent assets, we can re-write this objective as

$$\max_{\alpha_{il}} E\{\varphi[\alpha_{i1}(v_1 - P_1) - \gamma_1\alpha_{j1}(v_1 - P_1)] \mid v_1\} + \max_{\alpha_{i2}} E\{\varphi[\alpha_{i2}(v_2 - P_2) - \gamma_2\alpha_{j2}(v_2 - P_2)] \mid v_2\} \ . \quad (17)$$

The parameter γ_m is meant to capture the importance of asset m in the benchmark. An example where different γ_m may arise is the case where one asset is listed in the S&P 500 and the other is not. Or, alternatively, an asset may be domestic while the other may be foreign. With independent liquidity trades, u_m, we can solve the problem for each asset separately and obtain

$$P_m = \lambda_m(\alpha_{1m} + \alpha_{2m} + u_m),\ m = 1, 2 \text{ and } \alpha_{im} + \beta_m \tilde{s_{im}},\ i = 1, 2 \quad (18)$$

where

$$\beta_m = \sqrt{\frac{\sigma_{um}^2}{2(1-\gamma_m)\sigma_m^2}} \text{ and } \lambda_m = \frac{\sqrt{\frac{2(1-\gamma_m)\sigma_m^2}{\sigma_{um}^2}}}{(3-\gamma_m)} \quad (19)$$

Hence,

$$\mathrm{Var}(P_m) = (\lambda_m)^2[4\sigma_m^2(\beta_m)^2 + \sigma_{um}^2] = \frac{4\sigma_m^2}{(3-\gamma_m)^2} + \frac{2(1-\gamma_m)^2\sigma_m^2}{(3-\gamma_m)^2} = \frac{2\sigma_m^2}{(3-\gamma_m)} \quad (20)$$

We consider a stock m to be included in the benchmark if $\gamma_m > 0$. Notice that, if $\gamma_m = 0$, then $\mathrm{Var}(P_m) = 2\sigma_m^2/3 < (2\sigma_m^2)/((3-\gamma_m))$. Therefore, when a stock is not included in the benchmark, $\gamma_m = 0$, it has a smaller price volatility than when it is included. So, inclusion raises volatility. Furthermore, the more heavily the compensation package relies on a stock (higher γ_m), the higher the variance of that stock price, i.e., $[\partial\mathrm{Var}(P_m)/\partial\gamma_m] > 0$. This result is similar to the one obtained by Cuoco and Kaniel (2011), albeit in a different market microstructure. It is also intuitive because, when a stock is included in a benchmark and/or its weight in a benchmark basket of stocks is increased, managers will pay more attention to it, which may translate into more aggressive trading on this stock. Also, managers are likely to trade more frequently in such a stock as they rebalance their portfolios with the aim to track a benchmark that includes that stock.

Home bias with exogenous benchmark

Now, once again, we change the framework of our model slightly to explore yet another effect of relative performance: home bias. We assume that, instead of having two funds each indexed against the other, we have one fund indexed against a given exogenous benchmark (for instance, the S&P 500). So, we have two assets each with final value given by

$$\tilde{v_m} \sim N(0, \sigma_m^2),\ m = 1, 2 \quad (21)$$

and signals given by

$$\tilde{s_m} = \tilde{v_m} \quad (22)$$

where m index assets and there is only one fund manager. We also assume that liquidity trades are given by $\tilde{u_m} \sim N(0, \sigma_{um}^2)$ and are independent of each other and of the assets' final value.

Finally, we assume that the fund manager's expected payoff is given by

$$E\{\varphi(v\text{-}P)D^T + \chi[(v\text{-}P)D^T - \gamma(v\text{-}P)D_B^T]\} \quad (23)$$

where

$$v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix},\ P = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix},\ D = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix},\ D_B = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}, \quad (24)$$

and w_i is the weight of asset i on the exogenous benchmark (for instance, the participation of the stock on the S&P 500) and T denotes transpose and φ and χ are scalers.

By solving the model in the same way as before, we have

$$P_m = \lambda_m(\alpha_m + u_m) \text{ and } \alpha_m = \beta_m v_m + A_m, \quad (25)$$

for m = 1, 2 with

$$\beta_m = \frac{(\varphi+\chi)}{2\lambda_m},\ A_m = \frac{\chi\gamma w_m}{2},\ \lambda_m = \frac{\beta_m}{\beta_m^2\sigma_m^2 + \sigma_{um}^2} \quad (26)$$

And, if we concentrate on a situation similar to the ones discussed in the previous sections by imposing $\varphi + \chi = 1$, we have that

$$\alpha_m = \left[\sqrt{\frac{1}{2}\frac{\sigma_{um}^2}{\sigma_m^2}}\right]v_m + \frac{\chi\gamma w_m}{2} > \left[\sqrt{\frac{1}{2}\frac{\sigma_{um}^2}{\sigma_m^2}}\right]v_m = \alpha_m^{Kyle} \quad (27)$$

where the last term is the traditional demand in this setting without relative performance pay [see Kyle (1985)]. Therefore, we see that, if funds are benchmarked against an exogenous index, demand for asset m increases with w_m (and γ).

If we think about asset 1 as domestic and asset 2 as foreign and assume that domestic funds are benchmarked against domestic assets

only (w_2=0), we see that these funds will be overexposed to domestic assets: $\alpha_m > \alpha_1^{Kyle}$. This result provides an alternative explanation for the widely-documented home bias phenomenon, i.e., the observation that investors tend to favor local stocks, both in an international [e.g., Tesar and Werner (1995)] and domestic setting [e.g., Coval and Moskowitz (1999)].[16, 17] When managers are judged against a benchmark that represents a subset of available assets, they will aim to track those assets more closely and put a larger portfolio weight on those assets. The more important their performance against the benchmark, the more severe the home bias will be.

INFORMATION ACQUISITION

Next, we revert back to the single asset framework, but assume that fund managers may or may not acquire costly information. Let c denote the information acquisition cost. If no information is acquired, the manager stays out of the market. This is because uninformed trading is non-profitable. We start with a simple case where both managers have access to the information or do not – either both are informed or both are uninformed. Then we analyze the more interesting case where each manager decides independently whether to acquire information or not.

Information as a public good

If either both managers have access to the information or neither have access, calculating manager i's profit if both decide to be informed is straightforward and leads to the following.

Proposition 3: each manager expects to be paid

$$E\phi[\alpha_i - \gamma\alpha_j)(v\text{-}P)] = \phi(1\text{-}\gamma)\beta\sigma_v^2\left[\frac{\sigma_\varepsilon^2 + (1-\gamma)\sigma_v^2}{2(\sigma_v^2+\sigma_\varepsilon^2)+(1-\gamma)\sigma_v^2}\right] \quad (28)$$

And,

$$\frac{d}{d\gamma} E\phi[(\alpha_i - \gamma\alpha_j)(v\text{-}P)] < 0. \quad (29)$$

Therefore, fund managers' pay is decreasing on the relative performance parameter. Furthermore, if there exists a $\bar{\gamma}$ such that

$$E\phi[\alpha_i - \bar{\gamma}\alpha_j)(v\text{-}P)] = c \quad (30)$$

then for any $\gamma > \bar{\gamma}$ no information is acquired.

Proof: contact authors.

The above proposition shows that not only are investors worse off but also the managers themselves. More importantly, given the cost of signals, it may be the case that for high enough γ no information is acquired. While it may come as a surprise at first glance, this is an intuitive result: profits are decreasing in γ, so if there is a point where the managers are just indifferent between acquiring information or not, then, for any higher γ they will prefer to be uninformed and save c.

In summary, conditional on the information being acquired, prices are more informative (as described in Section 3), but it may be the case that no information is acquired at all, rendering the earlier point on the information content of prices moot.

To enhance our understanding of these countervailing effects, we further assume that the cost of information acquisition is randomly drawn from a distribution F(.), common to both managers. The cost is drawn once from this distribution and represents the cost for each manager. Based on the analysis so far, we know that the managers will acquire information with probability

$$F(E\phi[(\alpha_i - \gamma\alpha_j)(v\text{-}P)]), \quad (31)$$

i.e., if $\tilde{c} < (E\phi[(\alpha_i - \gamma\alpha_j)(v\text{-}P)]$.

We can then calculate the expected price response to trades, i.e., $E\lambda(\gamma)$:

$$F(E\phi[(\alpha_i - \gamma\alpha_j)(v\text{-}P)]) \frac{2\sigma_v^2\sqrt{\frac{\sigma_u^2}{2(1-\gamma)\sigma_v^2+2\sigma_\varepsilon^2}}}{(2\sigma_v^2+\sigma_\varepsilon^2)\frac{\sigma_u^2}{(1-\gamma)\sigma_v^2+\sigma_\varepsilon^2}+\sigma_u^2} + [1\text{-}F(E\phi[(\alpha_i - \gamma\alpha_j)(v\text{-}P)])] \text{ x } 0. \quad (32)$$

Hence,

$$\frac{d}{d\gamma}E\lambda(\gamma) = \lambda(\gamma)\left[\frac{d}{d\gamma}F(E\phi[(\alpha_i - \gamma\alpha_j)(v\text{-}P)])\right] + \left[\frac{d}{d\gamma}\lambda(\gamma)\right]F\left(E\phi[(\alpha_i - \gamma\alpha_j)(v\text{-}P)]\right) < 0 \quad (33)$$

where the inequality follows from the fact that

$$\frac{d}{d\gamma}F\left(E\phi[(\alpha_i - \gamma\alpha_j)(v\text{-}P)]\right) < 0 \quad (34)$$

16 Another form of home bias has been documented in that employees tend to hold excessive amounts of own company stock in their retirement saving plans [e.g., Benartzi and Thaler (2001)]. Pinheiro (2008) provides a rational explanation for this seemingly suboptimal exposure based on social interactions.

17 Most explanations for home bias rely either on impediments to trade due to institutional, legal and regulatory frameworks or on some form of residence-based assymmetric information.

as argued before, and $\frac{d}{d\gamma}\lambda(\gamma) < 0$, as proved in Corollary 3. In other words, the inequality follows because the probability that the managers acquire information decreases with the importance of relative performance contract but liquidity, provided that managers are informed, increases with the strength of relative performance incentives.

This result shows that the expected price response to order flow decreases as γ increases. This would indicate that markets are more liquid, even when there is an additional information acquisition stage. However, if we were to measure the depth of the market, the usual measure of liquidity in this framework, we would run into problems. Depth is given by 1/ λ, which is undefined (goes to infinity) if there are no informed traders in the market. Hence, as long as there is a strictly positive probability of no information collection, depth would be independent of γ. And, the market would have an infinite expected depth. Therefore, we concentrate on the former measure, expected price response to order flow, to claim that markets are more liquid with relative performance, even in the presence of information acquisition stage. Notice that this result is independent of F(.).

We can also measure the expected informational content of prices. This can be shown to be inversely related to[18]

$$\sigma_v^2[1\text{-}F(E\phi[(\alpha_i - \gamma\alpha_j)(v\text{-}P)])[\frac{4\beta^2\sigma_v^2}{4\beta^2\sigma_v^2+2\beta^2\sigma_\varepsilon^2+\sigma_u^2}]] \quad (35)$$

And, differentiating the above expression with respect to γ reveals

$$\frac{d[F(E\phi[(\alpha_i - \gamma\alpha_j)(v-P)])}{d\gamma}[\frac{-4\beta^2\sigma_v^4}{4\beta^2\sigma_v^2+2\beta^2\sigma_\varepsilon^2+\sigma_u^2}] - F(E\phi[(\alpha_i - \gamma\alpha_j)(v\text{-}P)])[\frac{\sigma_v^2}{4\beta^2\sigma_v^2+2\beta^2\sigma_\varepsilon^2+\sigma_u^2}]^2\, 8\beta\sigma_u^2\frac{d\beta}{d\gamma} \quad (36)$$

The first term is the product of two negative terms, hence it is positive. The last term is the product of two terms with opposite signs, hence it is negative. The overall sign of the expression cannot be determined. Interpreted in an intuitive manner, the first term represents the fact that, as relative performance incentives get stronger, there is a smaller chance of obtaining more informative prices and less information acquisition. The second term represents the effect that, if information is acquired, prices will be more informative. The overall effect on the information contect of prices is indeterminate.

In order to determine the sign of this expression under certain circumstances, we further simplify our model. The assumptions and results are spelled out in the following proposition.

Proposition 4: Assume that the distribution of costs is uniform on $[0, \bar{c}]$, with

$$\bar{c} > \phi\frac{\sqrt{(1-\gamma)^3}}{3-\gamma}\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}} \quad (37)$$

And, furthermore, set $\tilde{\varepsilon}_i \equiv 0$. Then, the expected informational content of prices is decreasing in γ, i.e., expression (30) is positive, so that the expected conditional variance of prices increases with γ.

Proof: contact authors.

This proposition shows that, with some additional restrictions, the result that price informativeness increases with γ is reversed. The effect of less information acquisition dominates the effect that, conditional on information acquisition, prices are more informative. So, higher γ leads to less informative prices ex-ante.

In summary, either with free information or costly public information, our model delivers that markets are deeper, the more important the relative performance objective is (lower $E\lambda(\gamma)$). However, costly information, even public, may lead to less informative prices (e.g., under uniform distribution of costs). The intuition for the first result is clear: the depth of the market is inversely related to the information asymmetry problem, so less information acquisition cannot, in the current framework, decrease depth. The second result is a direct consequence of the decrease in the probability of information acquisition.

Information as a private good

We now move to a more interesting informational structure. Suppose that each manager may or may not be informed. In other words, there are three possible outcomes: two informed managers, one informed and one uninformed, and both uninformed. For this part of the analysis, we again consider the case where agents have perfect information if they decide to be informed, i.e., we set $\tilde{\varepsilon}_i \equiv 0$. First, suppose that both managers are informed. Then each has an expected profit of

18 Note that expression (35) is not Var(v|P). Rather, it is the expected value of the conditional variance of v, conditional on whether or not there is information acquisition. Since P now is a mixed random variable and no longer normal, calculation of Var(v|P) would be very elaborate. And, we adopt this short-cut measure just to grasp the idea of the countervailing effects at play. The true value would be obtained by adding to expression (35) the following term:

$F(E\varphi[(\alpha_i - \gamma\alpha_j)(v-P)])[1-F(E\varphi[(\alpha_i - \gamma\alpha_j)(v-P)])]\ \frac{2\lambda\beta\sigma_v^2[\lambda(\beta\tilde{s_1}+\beta\tilde{s_2}+\tilde{u})]}{\lambda^2(4\beta^2\sigma_v^2+2\beta^2\sigma_\varepsilon^2)}$ making it even messier, and, now, dependent of the realization of random variables. Note that this term has zero expected value. Therefore, expression (35) is just the expected value of the conditional variance.

$$E\phi[(\alpha_i - \gamma\alpha_j)(v\text{-}P) = \phi\frac{\sqrt{(1-\gamma)^3}}{3-\gamma}\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}} \quad (38)$$

If only one is informed, we are in the exact framework of Kyle (1985), γ is irrelevant, and we know that the informed fund manager makes

$$E\phi[\alpha_i\,(v\text{-}P)] = \phi\frac{\sqrt{\sigma_u^2\sigma_v^2}}{2} \quad (39)$$

while the uninformed has zero expected profit. Finally, in the case of two uninformed managers, both have zero expected profits.

An informed manager's expected profit when her counterpart is uninformed exceeds her profit when her counterpart is informed, i.e.,

$$E\phi[\alpha_i\,(v\text{-}P)] > E\phi\,[(\alpha_i - \gamma\alpha_j)(v\text{-}P)], \quad (40)$$

if and only if $\frac{\sqrt{(1-\gamma)^3}}{3-\gamma}\sqrt{2} < 1, \forall\gamma$. Notice that

$$\frac{\partial}{\partial\gamma}[\frac{\sqrt{(1-\gamma)^3}}{3-\gamma}\sqrt{2}] < 0, \text{ and } [\frac{\sqrt{(1-\gamma)^3}}{3-\gamma}\sqrt{2}]\,|_{\gamma=0} = \frac{\sqrt{2}}{3} < 1.$$

Therefore, the required inequality holds for all γ, as expected, since managers' profits are decreasing in γ. The following proposition summarizes the insight from this reasoning.

Proposition 5: in the information acquisition stage, there are three possible equilibrium outcomes, depending on the realization of c:

(i) If $c > \phi\frac{\sqrt{\sigma_u^2\sigma_v^2}}{2}$, the market cannot support any informed traders, and the managers are both better off not acquiring information and staying out;

(ii) If $\phi\frac{\sqrt{\sigma_u^2\sigma_v^2}}{2} > c > \phi\frac{\sqrt{(1-\gamma)^3}}{3-\gamma}\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}}$, then the market can support one informed agent, but not two;

(iii) Finally, if $\phi\frac{\sqrt{(1-\gamma)^3}}{3-\gamma}\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}} > c$, the market supports two informed agents.

Now, the measure of the expected informational content of prices is inversely related to

$$F[\phi\frac{\sqrt{(1-\gamma)^3}}{3-\gamma}\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}}][\frac{\sigma_u^2\sigma_v^2}{4\beta^2\sigma_v^2+\sigma_u^2} - \frac{\sigma_v^2}{2}] - F\,[\phi\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}}]\frac{\sigma_v^2}{2} + \sigma_v^2, \quad (41)$$

or,

$$F[\phi\frac{\sqrt{(1-\gamma)^3}}{3-\gamma}\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}}][\frac{\frac{\sigma_u^2\sigma_v^2}{2} - 2\beta^2\sigma_v^4}{4\beta^2\sigma_v^2+\sigma_u^2}] - F\,[\phi\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}}]\frac{\sigma_v^2}{2} + \sigma_v^2, \quad (42)$$

And, again the derivative

$$[\frac{\frac{\sigma_u^2\sigma_v^2}{2} - 2\beta^2\sigma_v^4}{4\beta^2\sigma_v^2+\sigma_u^2}]\frac{d}{d\gamma}F\,[\phi\sqrt{\frac{(1-\gamma)^3}{3-\gamma}}\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}}] + F[\phi\sqrt{\frac{(1-\gamma)^3}{3-\gamma}}\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}}]\frac{d}{d\gamma}[\frac{\frac{\sigma_u^2\sigma_v^2}{2} - 2\beta^2\sigma_v^4}{4\beta^2\sigma_v^2+\sigma_u^2}], \quad (43)$$

cannot be signed, since the first term is positive (product of two negative terms), and the second is negative (product of terms with opposing signs).

We can similarly analyze the expected price response to trades, i.e., $E\lambda(\gamma)$:

$$F\,[\phi\frac{\sqrt{(1-\gamma)^3}}{3-\gamma}\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}}]\frac{2\sigma_v^2\sqrt{\frac{\sigma_u^2}{2(1-\gamma)\sigma_v^2}}}{\frac{\sigma_u^2}{2(1-\gamma)}+\sigma_u^2} + [F\,[\phi\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}}] - F[\phi\sqrt{\frac{(1-\gamma)^3}{3-\gamma}}\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}}]\frac{\sigma_v^2}{2\sigma_u^2}, \quad (44)$$

Or

$$F\,[\phi\frac{\sqrt{(1-\gamma)^3}}{3-\gamma}\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}}][\frac{2\sigma_v^2\sqrt{\frac{\sigma_u^2}{2(1-\gamma)\sigma_v^2}}}{\frac{(3-\gamma)\sigma_u^2}{(1-\gamma)}} - \frac{\sigma_v^2}{2\sigma_u^2}] + F[\phi\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}}]\frac{\sigma_v^2}{2\sigma_u^2}, \quad (45)$$

and calculating its derivative as

$$[\frac{4\sigma_v^2\sqrt{\frac{\sigma_u^2}{2(1-\gamma)\sigma_v^2}} - \frac{(3-\gamma)\sigma_v^2}{(1-\gamma)}}{\sigma_u^2\frac{(3-\gamma)^2}{(1-\gamma)}}]\frac{d\,F\,[\frac{\phi\sqrt{(1-\gamma)^3}\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}}}{3-\gamma}]}{d\gamma} + F[\frac{\phi\sqrt{(1-\gamma)^3}\sqrt{\frac{\sigma_u^2\sigma_v^2}{2}}}{3-\gamma}]\frac{d[\frac{4\sigma_v^2\sqrt{\frac{\sigma_u^2}{2(1-\gamma)\sigma_v^2}} - \frac{(3-\gamma)\sigma_v^2}{(1-\gamma)}}{\sigma_u^2\frac{(3-\gamma)^2}{(1-\gamma)}}]}{d\gamma} \quad (46)$$

But, now we cannot sign this expression, as we do not know the sign of $[\frac{4\sigma_v^2\sqrt{\frac{\sigma_u^2}{2(1-\gamma)\sigma_v^2}} - \frac{(3-\gamma)\sigma_v^2}{(1-\gamma)}}{\sigma_u^2\frac{(3-\gamma)^2}{(1-\gamma)}}]$.

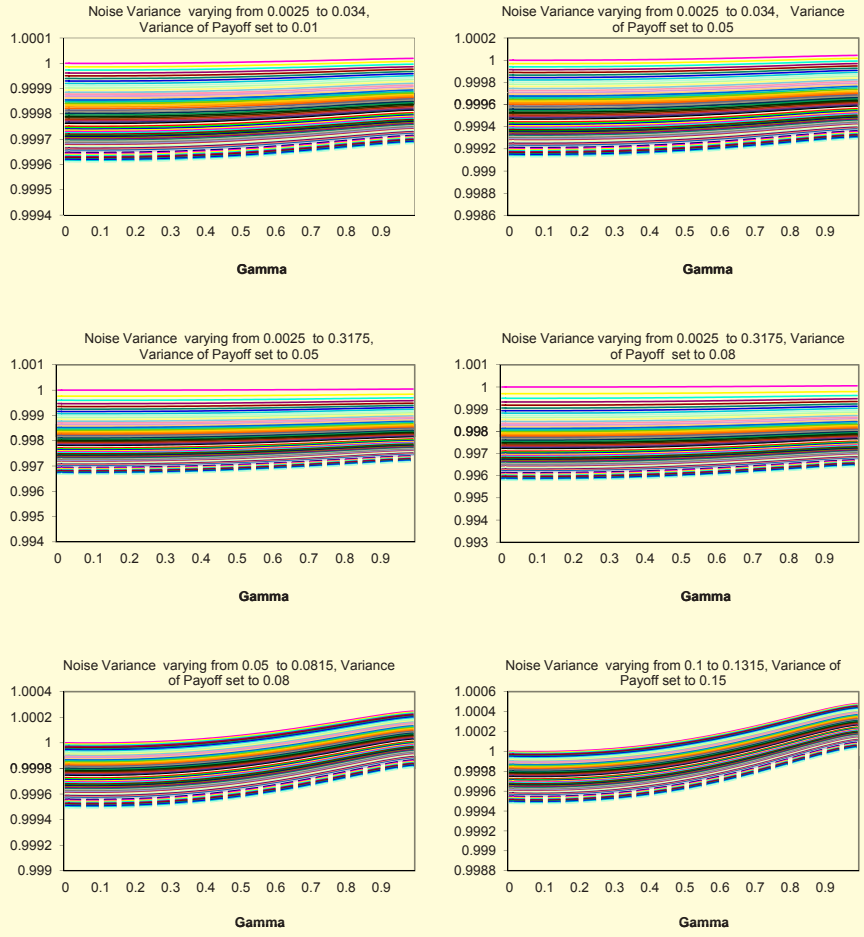
So when we add discretionary information acquisition as described here, we cannot determine what happens to the expected liquidity of the market as relative performance objectives become more important. The term in square brackets is the difference between price responses in a market with two informed traders and price responses in a market with one informed trader, i.e., $\lambda_{2Trader}(\gamma) - \lambda_{1Trader}$. This expression depends on the parameter values and, often, it changes sign as γ varies.

Unfortunately, the inconclusiveness of both results remains even if we assume that costs are uniformly distributed, as before. The only thing that can undoubtedly be stated is that information acquisition is ex-ante hindered by the presence of relative performance objectives.

If we increase γ, then some values of c that support equilibria with two informed agents may no longer do so. Increase in γ leads to an increase in the length of the cost interval where only one informed agent can be supported in equilibrium, and a decrease in the length of the cost interval where two informed agents can co-exist.

To shed some more light into this discussion, we resort to numerical analysis, under the assumption that costs are uniformly distributed on the unit interval. We concentrate on the typical behavior of the market microstructure elements that obtain indeterminate derivative signs in the analytical solutions.

First, we look at the informational content of prices. Figure 1 presents the typical relationship between the expected conditional



Source: Authors' calculations.
Note: We picture the expected conditional variance of prices, which is inversely related to the informational content of prices, as a function of relative performance concerns, denoted by γ. In all simulations, costs are assumed to be uniformly distributed on the unit interval and ϕ is set to 0.1 (using alternative values for ϕ does not alter the behavior). The variance of payoffs, σ_v^2, is set to 0.01, with results again being robust to other parameter values chosen. Each curve corresponds to a different value of the noise variance, σ_u^2, within the range indicated in increments of 0.005. The calculated values of the expected conditional variance of prices are then rescaled to equal 1 when γ is 0 and the noise variance is at the minimum of its set range.

Figure 1

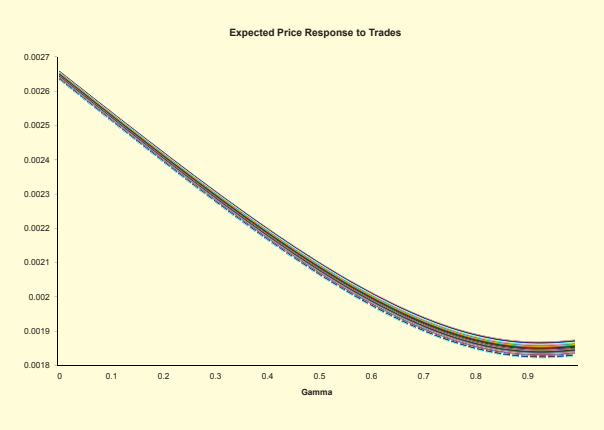


Figure 2

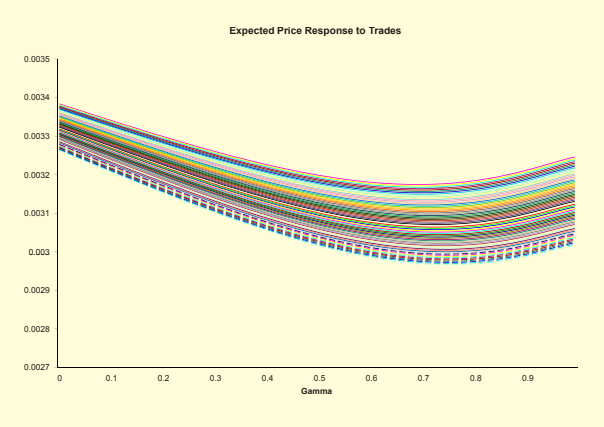


Figure 3

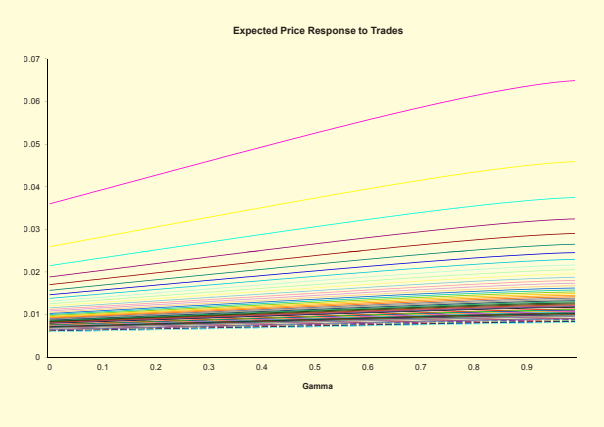


Figure 4

variance of prices (inverse of expected information content) and the measure of relative performance, γ. As with the case of public information, we obtain that this variance is an increasing function of γ, hence, as before, prices become less informative as relative performance objectives increase in importance. After a large number of simulations, the conclusion emerges to be that this result is pervasive and does not seem to depend on the parameter values. So, costly information, private or public, is enough to reverse the positive effects of relative performance on price informativeness.

Second, we take a closer look at the expected price responses to trades analyzing $E\lambda(\gamma)$, a measure inversely related to depth. Note that the behavior of this function depends on the difference $\lambda_{2Trader}(\gamma) - \lambda_{1Trader}$. If this difference is positive, less information acquisition is always good: moving from two informed traders to one informed trader, or one to zero, increases depth. This is a situation similar to the one under public information. Now, if this difference is negative, then less information acquisition may be bad: moving from two to one informed trader decreases depth but moving from one to zero informed traders still increases depth. Notice that as $\gamma \to 1$, $\lambda_{2Trader}(\gamma) \to 0$ so that for high enough γ the difference is negative.

Furthermore, whenever $2\sqrt{2} < \frac{3-\gamma}{\sqrt{(1-\gamma)}}\sqrt{\frac{\sigma_v^2}{\sigma_u^2}}$, the difference is again negative. So, apart from γ, the ratio $\frac{\sigma_v^2}{\sigma_u^2}$ plays an important role in determining the sign of this expression. In summary, the behavior of the expected price responses to trades as a function of γ can take three possible shapes. We visually depict these three possible situations below. In Figure 2, we have an always decreasing price response to trades, indicating an increase in depth. Figure 3 shows a case where the price response is decreasing at the beginning and reverses as γ grows, implying that less information acquisition becomes bad because $\lambda_{2Trader}(\gamma) - \lambda_{1Trader} < 0$. Finally, Figure 4 depicts a case where the price response is always increasing, reflecting the fact that the effect of less information acquisition is dominating and negative.

The shape depicted in Figure 2 is obtained whenever we have a low value for $\frac{\sigma_v^2}{\sigma_u^2}$, so that $\lambda_{2Trader}(\gamma) - \lambda_{1Trader} > 0$ and less information helps the depth of the market. Toward the end, the relationship flattens since $\gamma \to 1$. The shape shown in Figure 3 is obtained for higher values of the ratio $\frac{\sigma_v^2}{\sigma_u^2}$. The reversal sets in because, as γ increases, less information acquisition hurts the depth of the market (going from two to one trader is bad), and this effect dominates when γ is high enough. The shape illustrated in Figure 4 is obtained when we

start with a very high value of $\frac{\sigma_v^2}{\sigma_u^2}$ so that moving from two to one trader is always bad, i.e., it decreases depth.

The numerical results for the price informativeness confirm the analytical results obtained with public information. Relative performance decreases price informativeness when information is costly. However, the results for the depth of the market can be quite different depending on parameter values. The main insight here is that information acquisition can help increase market depth, even though it increases information asymmetry, and there exists parameter values for which information acquisition would only decrease market depth.

To summarize, our theoretical results predict that relative performance always increases market depth except under private costly information and some additional restriction on the parameters. However, we can only say with confidence that relative performance increases price informativeness under free information. Under costly information, this result may be easily reversed and it is always reversed under private costly information.

CONCLUSION

In this paper, we analyze the potential effects of the adoption of relative performance pay for fund managers. We show that, if fund managers are rewarded based on a relative performance measure, there can be deleterious effects for investors translating into a lower expected utility. However, markets will typically be more informative and deeper, if information is free.

When we endogeneize the acquisition of information, we see that incentives to acquire information may be hindered by relative performance contracts. This last effect may, ex ante, reduce the informativeness and depth of the market. So, even though conditional on information being acquired markets function "better," less information acquisition acts in a countervailing way. In other words, the result that prices are more informative holds only for the case of free information, it is reversed when information is costly. The result that markets are deeper (more liquid) also holds for the case of free information and for public costly information and may or may not hold in the case of private costly information.

These findings as a whole point to the need for a more nuanced understanding of the contracts that aim to solve the agency problem between managers and investors. Especially from a regulator's perspective, there can be a trade-off between the benefits of aligning the principal's interests with those of the agent and the potential costs associated with the effects on contagion, volatility, and the informativeness of markets. How this trade-off can be dealt with is a direction for future research.

REFERENCES

- Admati, A. R., and P. Pfleiderer, 1997, "Does it all add up? Benchmarks and the compensation of active portfolio managers," Journal of Business 70, 323–350
- Bagnoli, M., and S. Watts, 2000, "Chasing hot funds: the effects of relative performance on portfolio choice," Financial Management 29:3, 31–50
- Benartzi, S., and R. Thaler, 2001, "Naive diversification strategies in retirement saving plans," American Economic Review 91, 79–98
- Berk, J. B., and R. C. Green, 2004, "Mutual fund flows and performance in rational markets," Journal of Political Economy 112:6, 1269–1295
- Broner, F. A., R. G. Gelos, and C. M. Reinhart, 2006, "When in peril, retrench: testing the portfolio channel of contagion," Journal of International Economics 69:1, 203–230
- Brennan, M. J.,1993, "Agency and asset pricing," working paper, UCLA
- Brown, K., W. Harlow, and L. Starks, 1996, "Of tournaments and temptations: an analysis of managerial incentives in the mutual fund industry," Journal of Finance 51, 85–110
- Chen, H.-L., and G. Pennacchi, 2009, "Does prior performance affect a mutual fund's choice of risk? Theory and further empirical evidence," Journal of Financial and Quantitative Analysis 44:4, 745–775
- Chevalier, J., and G. Ellison, 1997, "Risk taking by mutual funds as a response to incentives," Journal of Political Economy 105, 1167–1200
- Coval, J., and T. Moskowitz, 1999, "Home bias at home: local equity preference in domestic portfolios," Journal of Finance 54:5, 2045–2074
- Cuoco, D., and R. Kaniel, 2011, "Equilibrium prices in the presence of delegated portfolio management," Journal of Financial Economics 101, 264–269
- Das, S., and R. Sundaram, 1998a, "On the regulation of mutual fund fee structures," NBER Working Paper 6639
- Das, S., and R. Sundaram, 1998b, "Fee speech: adverse selection and the regulation of mutual fund fees," NBER Working Paper 6644
- Dybvig, P., H. Farnsworth, and J. Carpenter, 2010, "Portfolio performance and agency," Review of Financial Studies 23:1, 1-23
- Eichberger, J., S. Grant, and S. P. King, 1999, "On relative performance contracts and fund manager's incentives," European Economic Review 43, 135–161
- Elton, E. J., M. J. Gruber, and C. R. Blake, 2003, "Incentive fees and mutual funds," Journal of Finance 58:2, 779–804
- Goriaev, A., F. Palomino, and A. Prat, 2000, " Mutual fund tournament: risk taking incentives induced by ranking objectives," Working Paper CentER, Tilburg University, No. 2000–94
- Grinblatt, M., and S. Titman, 1989, "Adverse risk incentives and the design of performance based contracts," Management Science 35, 807–822
- Heinkel, R., and N. M. Stoughton, 1994, "The dynamics of portfolio management contracts," Review of Financial Studies 7, 351–387
- Koski, J., and J. Pontiff, 1999, "How are derivatives used: evidence from the mutual fund industry," Journal of Finance 54:2, 791–816
- Ippolito, R. A., 1992, "Consumer reaction to measures of poor quality: evidence from the mutual fund industry," Journal of Law and Economics 35, 45–70
- Kyle, P., 1985, "Continuous auction and insider trading," Econometrica 53:6, 1315–1336
- Lettau, M., 1997, "Explaining the facts with adaptative agents: the case of mutual fund flows," Journal of Economic Dynamics and Control 21, 1117–1147

- Ou-Yang, H., 2003, "Optimal contracts in a continuous-time delegated portfolio management Problem," Review of Financial Studies 16, 173–208
- Palomino, F., 2005, "Relative performance objectives in financial markets," Journal of Financial Intermediation 14, 351–375
- Pinheiro, M., 2008, "Loyalty, peer group effects, and 401(k)," Quarterly Review of Economics and Finance 48:1, 94–122
- Sensoy, B., 2009, "Performance evaluation and self-designated benchmark indexes in the mutual fund industry," Journal of Financial Economics 92:1, 25–39
- Shelef, O., 2013, "Incentive thresholds, risk taking, and performance: evidence from hedge funds," working paper, Haas School of Business
- Sirri, E. R., and P. Tufano, 1998, "Costly search and mutual fund flows," Journal of Finance 53, 1589–1622
- Starks, L. T., 1987, "Performance incentive fees: an agency theoretic approach," Journal of Financial and Quantitative Analysis 22, 17–32
- Taylor, J., 2003, "Risk-taking behavior in mutual fund tournaments," Journal of Economic Behavior and Organization 50:3, 373–83
- Tesar, L., and I. Werner, 1995, "Home bias and high turnover," Journal of International Money and Finance 14:4, 467–492

FINANCIAL COMPUTING & ANALYTICS

STUDENTSHIPS

Four-Year Masters & PhD

for Final Year Undergraduates and Masters Students

As leading banks and funds become more scientific, the demand for excellent PhD students in **computer science, mathematics, statistics, economics, finance** and **physics** is soaring.

In the first major collaboration between the financial services industry and academia, **University College London, London School of Economics,** and **Imperial College London** have established a national PhD training centre in Financial Computing & Analytics with £8m backing from the UK Government and support from twenty leading financial institutions. The Centre covers financial IT, computational finance, financial engineering and business analytics.

The PhD programme is four years with each student following a masters programme in the first year. During years two to four students work on applied research, with support from industry advisors. Financial computing and analytics encompasses a wide range of research areas including mathematical modeling in finance, computational finance, financial IT, quantitative risk management and financial engineering. PhD research areas include stochastic processes, quantitative risk models, financial econometrics, software engineering for financial applications, computational statistics and machine learning, network, high performance computing and statistical signal processing.

The PhD Centre can provide full or fees-only scholarships for UK/EU students, and will endeavour to assist non-UK students in obtaining financial support.

UCL LSE Imperial College London

financialcomputing.org

INDUSTRY PARTNERS

Financial:

Barclays
Bank of America
Bank of England
BNP Paribas
Citi
Credit Suisse
Deutsche Bank
HSBC
LloydsTSB
Merrill Lynch
Morgan Stanley
Nomura
RBS
Thomson Reuters
UBS

Analytics:

BUPA
dunnhumby
SAS
Tesco

MORE INFORMATION

Prof. Philip Treleaven
Centre Director
p.treleaven@ucl.ac.uk

Yonita Carter
Centre Manager
y.carter@ucl.ac.uk

+44 20 7679 0359

Guidelines for Manuscript Submissions

Capco Journal of Financial Transformation

GUIDELINES FOR AUTHORS

In order to aid our readership, we have established some guidelines to ensure that published papers meet the highest standards of thought leadership and practicality. The articles should, therefore, meet the following criteria:

- Does this article make a significant contribution to this field of research?
- Can the ideas presented in the article be applied to current business models and financial problems? If not, is there a roadmap on how to get there?
- Can your assertions be supported by empirical data?
- Is your article purely abstract? If so, does it picture a world that could exist in the future?
- Can your propositions be backed by a source of authority?

All submitted articles need to pass peer review in order to be accepted for publication, and we aim at communicating our decision to authors within three months of submission.

Subjects of interest

The topics that are of interest to our readership include:

- Risk management in finance and insurance
- Operational risk and strategic issues
- Financial regulation
- Financial innovation and new sources of liquidity
- Capital acquisition and asset allocation
- Valuation and hedging for financial portfolios
- Financial modeling
- Algorithmic trading and optimal execution
- Impact of e-finance on financial markets and institutions
- Financial technology
- Cyber-finance, including peer-to-peer lending, cryptocurrencies and mobile payments
- Financial inclusion, unbanked world
- Impact of macroeconomic policies on the financial industry
- Organizational behavior and structure
- Competitive landscape
- Structural readjustment
- Leadership
- Marketing and branding

Request for Papers

Deadline September 30, 2016

MANUSCRIPT GUIDELINES

Manuscripts should be submitted by e-mail to **institute@capco.com**.

All submissions must be in U.S. English. Manuscripts should be no longer than 10,000 words, including all footnotes, references, charts and tables. The manuscript title should be no more than 90 characters (including spaces).

Manuscripts must be in Microsoft Word, using Times New Roman font, size 10. Accompanying the Word document should be a PDF version that accurately reproduces all formulae, graphs and illustrations. When graphs are used, the data should also be provided in a Microsoft Excel spreadsheet format. Tables should be in Excel or Word, artwork should be vector-based or in high resolution.

Note that the Journal style does not include numbered sections. To represent section hierarchy levels, we recommend formatting subtitles in bold 14 pt, 12 pt and 10 pt. Bold is used only in subtitles, tables and graphs. Formatting in italics or underline will not be reproduced.

The first page must provide a short abstract, as well as the full names, titles, organizational affiliations, JEL codes, and contact details of the authors, including phone number and e-mail address.

Footnotes should be kept to a minimum and be numbered consecutively throughout the text with superscript Arabic numerals.

Citations in the text should be formatted in the following ways.

- Author (year) if part of sentence: Smith (2001) mentions that…
- [Author (year)] if inserted in the sentence: This common approach [e.g. Smith (2001), Jones (2003)] is the one we used.
- [Author (year), page number(s)] if using a direct quote: "in the financial world today" [Smith (2001), 25].

References at the end of the paper should use the following formats.

Books
Copeland, T., T. Koller and J. Murrin, 1994, Valuation: Measuring and Managing the Value of Companies, New York: John Wiley & Sons

e-Books
Copeland, T., T. Koller and J. Murrin, 1994, Valuation: Measuring and Managing the Value of Companies, New York: John Wiley & Sons. Kindle edition

Contributions to collective works
Ritter, J. R., 1997, Initial Public Offerings, in Logue, D. and J. Seward (eds.), Warren Gorham & Lamont Handbook of Modern Finance, Cincinnati: South-Western College Publishing

Journals and periodicals
Aggarwal, R. and S. Dahiya, 2006, "Demutualization and Cross-Country Merger of Exchanges," Journal of Financial Transformation, 18: 143-150

Website pages and articles, and e-journals
Allen & Overy, 2014, Capital Requirement Directive IV Framework, 2014, Introduction to Regulatory Capital and Liquidity. Briefing Paper 1. Available at: http://www.allenovery.com/ publications/en-gb/Pages/Regulatory-capital.aspx

Government publications
Bank for International Settlements (BIS), 2011, Quarterly Review, Basel

Newspaper articles
Cumming, F., 1999, "Tax-Free Savings Push," Sunday Mail, 4 April, p. 1

Unpublished material
Gillan, S. and L. Starks, 1995, Relationship Investing and Shareholder Activism by Institutional Investors. Working Paper, University of Texas

Layout, production and coordination: Cypres – Daniel Brandt, Kris Van de Vijver and Pieter Vereertbrugghen
Graphics: DuKemp
Photography: Alex Salinas

De Kleetlaan 6, B-1831 Machelen

Imperial College Business School | Centre for Global Finance and Technology

Centre for Global Finance and Technology

The Centre for Global Finance and Technology at Imperial College Business School will serve as a hub for multidisciplinary research, business education and global outreach, bringing together leading academics to investigate the impact of technology on finance, business and society.

This interdisciplinary, quantitative research will then feed into new courses and executive education programmes at the Business School and help foster a new generation of fintech experts as well as re-educate existing talent in new financial technologies.

The Centre will also work on providing intellectual guidance to key policymakers and regulators.

"I look forward to the ground-breaking research we will undertake at this new centre, and the challenges and opportunities posed by this new area of research." – Andrei Kirilenko, Director of the Centre for Global Finance and Technology

Find out more here:
imperial.ac.uk/business-school/research/finance/centre-for-global-finance-and-technology/

CAPCO

BANGALORE
BRATISLAVA
BRUSSELS
CHICAGO
DALLAS
DÜSSELDORF
EDINBURGH
FRANKFURT
GENEVA
HONG KONG
HOUSTON
JOHANNESBURG
KUALA LUMPUR
LONDON
NEW YORK
ORLANDO
PARIS
SINGAPORE
TORONTO
VIENNA
ZÜRICH