Financial Innovation and the 2008 Crisis: Lessons for Theory and Policy

Michael G Jacobides, London Business School

Draft v 1, preliminary and incomplete, February 7, 2014

Abstract

For all its alleged importance as a driver of the financial crisis, we know remarkably little about financial innovation, its antecedents and its systemic implications. Anecdotal evidence on changes to business models in financial services and their role in the recent crisis is not backed up by systematic research – and nor has it led to a rethink in terms of either theory or policy. This provides an opportunity to consider the nature and role of financial innovation (in particular, the manner in which the financial sector was structured) in driving the 2008 financial crisis, and to draw theoretical implications from this analysis. Specifically, this paper looks at some relatively neglected factors that can be fruitfully documented through historical and qualitative methods: the nature of the selection environment, the agency of actors, and the influence of structure (industry architectures and business models) in the context of the financial sector. On the basis of that evidence, as well as the premise that economic behavior is driven by feedback rather than foresight, we come to a fresh set of conclusions on what drove the crisis, and an opportunity to develop theory informed by a historically based evolutionary and institutional approach. This potentially challenges some current policy views, such as the concept of “Too Big To Fail” and the direction of regulatory initiatives. It also helps us revisit the lessons we should learn from this crisis, shifting the focus from macroeconomic factors and individual malfeasance to the role of structure and feedback mechanisms. This, in turn, generates fresh suggestions for regulatory priorities.

---

1 This paper, which is a very early draft, draws liberally from a previous manuscript, co-authored with Sidney G. Winter, titled “Survival of the Reckless: The evolution of the US mortgage industry towards disaster”, presented in the 2010 American Economic Association symposium on Evolutionary Perspectives on the Crisis (with G. Dosi, R. Shiller, J. Stiglitz and M. Colander). It also draws on its two sister projects, “Rethinking the future of financial services: A structural and evolutionary perspective on regulation” (with M. Drexler and J. Rico) forthcoming, Journal of Financial Perspectives and “Why did the financial crisis happen? Structuring our historical understanding of a predictable evolutionary disaster”, forthcoming, Business History. All errors remain my own, and comments are welcome to mjacobides@london.edu.
Financial innovation has arguably been a prime force of transformation, not only in the financial sector, but in the contemporary economic infrastructure more widely. As early as the 1980s, Merton Miller claimed that “... the word ‘revolution’ is entirely appropriate for describing the changes in financial institutions and instruments that have occurred in the past twenty years” (Miller, 1986: 437). That same year, the “big bang” in London (i.e. the deregulation of brokerage and investment banking activities in the UK) marked the beginning of a period of transformation that would dwarf anything we had witnessed before. The development of securitization, the explosion of new financial products, particularly derivatives (whose notional value went from zero to $531 trillion, i.e. almost ten times world GDP, in 2008), the growth of hedge funds and the shadow banking sector (see Poszar et al, 2010), the changes in regulation (including the repeal of Glass-Steagel in the US, which favored global financial behemoths) and a long period of deregulation transformed the landscape of banking completely (see Acharya & Richardson, 2009; Llewellyn et al, 2009; Jacobides, 2014, for a review). Lending (and deposit-taking) shifted from an integrated function, performed by a single institution, to a market-mediated chain involving a plethora of actors, products, structures, and arrangements. More important yet, the way existing financial institutions made money, as a result of this innovation, changed fundamentally, co-evolving with the new institutional structure and financial innovation (Cetorelli et al, 2012).

This financial revolution had major implications for the economy overall. Most significantly, it sowed the seeds of the latest economic crisis: The collapse of the financial service sector in 2007/8 began with the unraveling of some of these innovative products and structures. The run on the Asset-Backed Commercial Paper, the realization that some tranches of Collateralized Debt Obligations (CDOs) and other Mortgage Backed Securities (MBSs) were nowhere near as safe as had been thought (or as their ratings implied), and the subsequent “run on the shadow banking sector” (including Money Market Mutual Funds, who provided much of the short-term liquidity in the system) left banks unable to access the liquidity required to fund themselves. As it turned out, they had also gorged themselves on risky securities, which, while rated as safe by credit rating agencies, proved toxic – especially those based on sub-prime mortgage loans and transformed through Collateralized Debt Obligations (CDO). The ensuing financial crisis nearly destroyed the global financial system, leading to the biggest contraction since the great depression and serious economic hardship. It also saw real anger directed at the financial service world, which had also been characterized by an increase of pay for its captains and star performers. Concerns of social injustice, banker malfeasance, and regulatory weakness, and a rather limited contrition in terms of the tools of economic analysis soon followed (Pew Foundation, 2009; Caballero, 2010; Colander, 2011).

---

2 The changes were strongly supported by nearly all academic economists, and regulation was broadly favourably predisposed to financial innovation, as well as to the changes in financial intermediation. Some dissenting voices did exist, in particular with regard to derivatives. It may be worth contrasting three well-known quotes, all recorded a few months from each other:

“Credit derivatives and other complex financial instruments have contributed to the development of a far more flexible, efficient and hence resilient financial system.” (Alan Greenspan, 2004).

“The development of credit risk transfer has a potentially important impact on the functioning of the financial system. It provides opportunity for more effective risk management, promises the relaxation of some constraints on credit availability, and allows more efficient allocation of risk to a wider range of entities.” (Bank of International Settlements, 2003)

“Derivatives are financial weapons of mass destruction, carrying dangers that, while now latent, are potentially lethal to the financial system.” (Warren Buffet, 2004)
The change in the financial services sector, immortalized by documentaries such as *Inside Job*, or dramatizations such as *Margin Call*, was profound in its impact on financial stability, income distribution, and political discourse. Given this intense interest, one might have expected a flurry of academic work looking at the drivers and implications of financial innovation and the transformation of banking – over and above the narratives of the crisis. With some notable exceptions (e.g., Acharya et al, 2009, Rajan, 2010), most of the literature has focused on the underlying macro-economic factors, including the creation of a property and asset value bubble, the workings of collective manias, and the cycles of excessive growth and collapse (Taylor, 2013; Reinhardt & Rogoff, 2009). Strangely, though, there are thin pickings to be had when we search for a pre-crisis analysis of financial innovation and business-model transformation, or indeed a post-crisis *compte-rendu* of how and why innovation in business practices led to this disaster, and why these innovations were adopted. This paper aspires to fill this gap, focusing in particular on innovations in *business models* rather than products or technology.

My focus in this paper is on the way *institutions* turn their assets, resources, and capabilities into revenues, as opposed to how they exploit simple technological innovations that affect the relative efficiency of costs, (e.g. through the use of IT), or the nature of new products and services. I argue that while *product* innovations may lead to *business-model* innovation, focusing on the latter is critical, as it allows us to consider how *institutions* are affected, and also how these changes shape the potential courses of action available to individuals working in these institutions. I also offer some theoretical tools drawing on evolutionary economics (Nelson & Winter, 1982) that can help us understand how and why innovations are adopted and firms or sectors change.

Having established this theme, I then shift to the sector level, and to the analysis of the range of potential business models that are acceptable at any one point of time. Drawing in recent work on “industry architectures” (Jacobides et al, 2006; Brusoni et al 2009), I explain why industry architectures can help us understand the dynamics of a sector such as financial services – what drives the sector’s evolution, and how and why innovations are adopted or ignored. The industry architecture perspective also provides the foundations to study the evolution of a sector such as financial services and the risk of its systemic collapse. I then consider how and why the 2008 crisis came about, and illustrate the usefulness of this behavioral approach in practice. I conclude with implications for research, both in terms of the financial sector and more broadly, and some suggestions on how to rethink regulation and the future of the financial services sector.

**Understanding Financial Innovations and the Changes in Financial Institutions**

Rather than starting with the financial crisis and its drivers, it is worth stepping back and considering the broader question of when and why innovations take place in the financial sector. After all, it was the adoption of what proved to be a disastrous set of practices or tools that brought us here (in addition to the usual dynamics of asset bubbles and credit expansion). Given the manifest importance of financial innovation, the lack of research on the topic comes as something of a surprise. In their 2004 review piece in the *Journal of Economic Literature*, aptly titled “Empirical studies of financial innovation: lots of talk, little action”, Frame & White (2004) paint a disheartening picture of the state of our knowledge – while also offering some insightful pointers on why it is that we know so little about either the drivers or the implications of financial innovation. As they note:
A broad descriptive literature that discusses recent financial innovations and that advances various hypotheses about them has proliferated. A striking feature of this literature, however, is the relative dearth of empirical studies that specifically test hypotheses or otherwise provide a quantitative analysis of financial innovation. This is surprising given the importance of the financial sector, the widespread recognition of the rapid and broad proliferation of financial innovation, and the relative abundance of similar papers for other sectors of the economy, especially manufacturing and agriculture. (2004: 116)

In our current efforts to scour the financial literature landscape, using fairly broad criteria and a long time horizon, we could find only 39 articles that provide empirical tests of hypotheses of any kind concerning financial innovation... even more striking, only two studies provide tests of the hypotheses that have been advanced in that broad descriptive literature concerning the economic/environmental conditions that encourage financial innovation. By contrast, such tests have been at the center of the empirical economics literature concerning innovation generally and have inspired “the second largest empirical literature in the field of industrial organization, exceeded only by the literature investigating the relationship between concentration and profitability. (2004: 135)

Explaining why this would be, the authors suggest that:

...empirically oriented finance economists (with a few exceptions) have usually not been trained in IO, have usually not been taught to think about the Schumpeterian hypotheses, and have had less suitable data available to them for testing those (or any other) hypotheses. In practice, the intellectual specializations that have characterized these fields have tended to create a void that has encompassed empirical studies of financial innovation. (2004: 137)

These concerns are underscored by Beck et al (2012), who suggest that “there is a striking paucity of empirical studies of determinants and consequences of financial innovation, mainly due to the lack of data”. More recently, Frame & White’s (2013) chapter in the Acharya et al (2013) volume revisits financial innovation post-crisis. They confirm that little has changed in terms of systematic empirical evidence, despite the tempest of the financial crisis and intense debate in the polity on the nature and benefits of innovation. They analyze three major “usual suspects” (Credit Default Swaps/CDSs, Structured Investment Vehicles/SIVs, and securitization), offering a cautious assessment of the value and drivers of these innovations, and conclude that “regulatory capital arbitrage appears to have played a central role in the crisis as many tarnished innovations were those that were principally designed to lower regulatory capital for a given risk.” (2013: 226).

A similar assessment is provided by Acharya & Richardson (2009a) in their important paper focusing on the role of SIVs leading up to the crisis. The authors discovered that banks were using SIVs to offload loans and shift them off their balance sheets, thus reducing the necessary regulatory capital while retaining the risk, either formally (through conduits that legally obliged them to buy back loans that went bad) or in practice (since banks would want to stand by their own SIVs, so as not to jeopardize their funding from their wholesale market in the event of a panic). Similarly, Acharya, Schnabl, & Suarez (2012) analyze asset-backed commercial paper conduits, which played a central role in the crisis, and document that commercial banks set up conduits to securitize assets while insuring the newly securitized assets using credit guarantees. The credit guarantees were structured to reduce bank capital requirements, while providing recourse to bank balance sheets for outside investors – an option that investors duly took when loans soured, leading to losses for the banks rather than themselves. In stark contrast to the rhetoric and textbook endorsement of this
innovation, the authors find that banks used this form of securitization to concentrate, rather than disperse, financial risks while reducing their capital requirements.

Viewed from afar by one concerned with the business economics of banking, this use of the SIV might look sensible, provided we take the banks’ objective as maximizing return on equity. The problem is that this use of the innovation, as part of a bank’s business model, wasn’t what the supporters of SIVs and securitization regarded as “ideal” – hence Frame & White’s use of the term “tarnished”, or the concern of Acharya & Richardson that banks would find any ways they could (and that were permitted, or even encouraged by regulators) to maximize their ROE. It is this gap – between the “textbook use” of a product or structural innovation and the way it is actually used by a bank as part of its business model – that we will focus on later in the paper.

Much of the preceding literature on financial innovation focuses on how these new products, structures and arrangements can, in a context of a set of well-intentioned actors, improve “market efficiency”. Crucially, the focus of this research is primarily on the products or structures, and secondarily on the markets, and not on the financial institutions and the way they make money – their business models. Tufano (2003) for instance, provides an overview of the motivations involved, concluding that if an innovation is introduced, it must be that both issuers and buyers benefit. The innovation’s impact on the business economics of individual institutions or its interaction with agency and the stability of the system are not actively considered. Llewellyn et al (2009), in an edited volume on financial innovations, provide a passionate thesis in support of particular innovations, making mainly theoretical arguments for why they would provide better liquidity, improve price finding, expand credit, etc. Again, they do not consider how these innovations work in practice, and how they shape the behavior of existing or new industry participants, potentially affecting the stability of the financial sector. To put it in the words of Laeven (2013: 234), “the literature on financial innovation primarily focuses on the good or the dark sides of innovation but not on the incentives that the owners of financial institutions have to exploit such innovation”. I would add that most of the research is on “the good” side; and that the incentives of employees, and in particular traders, to exploit such innovations, are also not considered. Neither are organizational and behavioral factors.

Disciplinary biases and preferences in economics may have played a part in this. Both key policy makers and regulators (such as Rubin, Summers and Greenspan) but also the vast majority of financial economists showed bias in favor of financial innovations, since they seemed to be leading to better functioning markets – at least until 2008. With the institutional realities of banks or other industry participants being abstracted away in much contemporary economic and financial analysis, it is easy to understand the enthusiasm of properly trained economists. In an edited volume on financial innovations, written after the crisis started and published after the 2008 collapse, the introductory chapter remarks that:

By increasing the range of financial instruments, the financial system [moves] closer to the Arrow-Debreu (1954) ideal where all transactors can ensure for themselves delivery of goods and services in all future contingencies, and the system comes closer to approximating the number of “states of nature”... therefore financial innovation has the capacity to significantly enhance the efficiency of the financial system for the core of its functions (Llewellyn et al, 2009: 26-27).
It is indeed true that, from a theoretical vantage point, these innovations feel more market-like, and, as such, efficient as far as the aesthetics of professional economists go; the word “ideal” in the quote above illustrates the point. The problem is that institutions and their money-making mechanics were not considered; as a result, the changes in the business logic of the financial service sector were not explicitly considered.

Allen (2010), in another review of financial innovation, defends the principle and practice of financial innovation by suggesting there are several innovations that could be beneficial, so that caution should not drive us to disregard their benefits. His argument rests partly on international comparisons of different systems that are more or less dynamic, or that have greater or lesser propensity for innovation. Of course, any such comparison is fiendishly complicated, since the stability of each country’s financial system is affected by a wealth of factors outside financial innovation – growth and political dynamics, prudential and supervisory capabilities, or the soundness of banks as institutions. Still, the most developed analysis of the role of financial innovation does come from such international longitudinal comparisons (see, e.g., Taylor, 2013, or Pagano, 2013 – who, on the basis of aggregate data, suggests a non-linear relationship between innovation and financial stability.) This research, useful as it might be in establishing basic facts, does not get into the details of how innovations shape (and potentially destabilize) financial systems.

Some research has come closer to that objective, but approaches it from the perspective of the new tools or products rather than the institutions that use them. To give a rough analogy, the research has been mostly about guns and ammunition, rather than who uses them, how they are deployed, and whether battles are won or lost as a result. Stulz (2010) has provided an overview of the role of CDSs, arguing that they cannot be blamed for the financial crisis. Taking a dimmer view of innovation but still focusing on products, Henderson & Pearson (2011) find that the pricing of equity derivatives cannot be justified in any fundamental way, other than taking advantage of uninformed investors.

For all the analysis of new products, there hasn’t been much research on how the institutions in financial services change due to these innovations, or even on what drives such changes. Using the broadest definition possible in this regard, Frame & White note that “In the course of our research, we uncovered seven papers examining three different organizational innovations. Unlike the generally positive welfare benefits that emerge from the studies of product and process innovations, the studies of organizational innovations find mixed results.” (2004: 133). Research on business models has been more limited. As we saw, Acharya & Richardson (2009), and Acharya et al (2013) have documented the use of SIVs and securitization by banks to reduce regulatory capital rather than restrict risk, but most of the research has abstracted away from the messy institutional realities of the way the sector was structured – its roles, rules, links and relationships – and the way organizations within it made money, or managed and compensated their staff.

The emergence of the shadow banking sector, though, was hard to ignore, as volumes in derivatives swelled and new types of institutions (especially hedge funds) grew and clearly became more entwined in the global financial system. Initially, the term “shadow banking system” encompassed everything that many regulators and academics thought was outside the realm of traditional banks (a belief that would prove to be mistaken). There was a view, still held today, that shadow banking activities could be separated from the regulated part of the financial sector. But problems in the shadow banking sector itself (issues with asset-backed securities, problems with money-market
funds, the restriction of wholesale funding) soon demonstrated that the shadow banking sector could not be separated from the formal banking sector, and that it was the formal banking sector that had been transformed in the way that it made money. To put it in more managerial terms, the business models for banks had changed, and the industry architecture of financial services had been transformed.

**Structural Change and New Business Models in Financial Services**

The first comprehensive report on the structure of the shadow banking sector was the 2010 report by the Federal Reserve Bank of New York (Poszar et al, 2010), which provides a 90-page analysis of how the sector works, the players within it and how they are connected. This report also includes the infamous “map” of the shadow banking system, included here as Figure 1. Remarkably, it was not until well after the crisis that this hugely important part of the financial service sector was mapped in this way, bringing its complexity and interconnectedness into the light for the first time.

Include Figure 1 around here

The story of the emergence of this research is worth retelling. Poszar (personal communication) recounts how he first took an interest in the shadow banking system in 2007, when working in Moody’s Economy.com. In the Jackson Hole meeting of 2007, PIMCO’s Paul McCulley’s stated that there was a “run on the shadow banking system” that could lead to financial destabilization. Poszar’s interest was piqued. With the encouragement of his boss, Mark Zandi, he drafted a brief overview of shadow banking and sketched out the map that later became Figure 1. He then got in touch with Bill Dudley, Head of Markets at the Federal Reserve Bank of New York, who hired him to follow the shadow banking system. In 2009, as Poszar recounts (personal communication):

During the IMF meetings in Turkey, my fiancée was gone and having a bit of extra time, I prepared the big map showing all the participants in the ecosystem. No one had asked me to do this; and I had no expectation. Just intellectual curiosity. The map was 3 by 4 feet, and it was a pretty densely printed map, too! We put it up on the wall of the briefing room. Then people started getting interested. People in Research started talking to me, so pretty soon I moved there, and started working on the structural issues for them… went up to the markets room with the map and pinned it up in the briefing room (where the market briefings to the Board of Governors and Fed operations are conducted from) and then sent an email to the Markets Group along the lines of: "I've done this map, I pinned it up on the wall and that our 13(3) facilities amount to a 360-degree backstop of the shadow banking system. We do not monitor 1/10th of what's on the map on a daily basis, we should think deep about that".

After a few days of awkward silence requests started to flood in from research, bank supervision and other parts of the bank (even Mark Carney at the Bank of Canada secured himself a copy and had a private walk through of the map on his request at the BoC in Ottawa) for additional copies. I spent the next few weeks explaining to senior staff how the system worked, where it broke and how our facilities (unbeknownst to everyone at the time) amounted to a full backstop of the shadow banking system. The head of research, Jamie McAndrews suggested that I should be seconded into research to write the map into a monograph. The result was the New York Fed’s 90 page Staff Report on "Shadow Banking".
In retrospect, it is remarkable that it took a young outsider to spot these changes. It is even more remarkable that there was such effusive interest in his work, and such a lack of understanding of structure in the sector. This same lack of focus on structure emerged from my own conversations with senior industry executives during and immediately after the crisis. During a series of meetings between UK regulators (FSA and Bank of England), MPs, Peers, and industry executives in in the Houses of Parliament in 2010 and 2011, it became clear to me that views of the system were very fragmented, even among those who tried to understand it as a whole. The same issue had emerged from a series of meetings in 2009 co-sponsored by the World Economic Forum, a major bank and a global consulting firm, summarized in Jacobides (2010).

As the role of the shadow banking sector came into sharper focus, it became clear that the regulators, through their support to Money Markets and their 13(3) lending facilities (which were meant to support institutions outside the formal banking system), were changing the sector. From late 2008 onwards, the Central Bank in the US (the Fed) was transformed, in Mehrling’s (2011) words, from “the banker of last resort” to “the trader of last resort”. In its quest to support economic activity, and as a result of the transformation and subsequent collapse of the sector, the Fed was changing its role, and transforming its own balance sheet and mode of interaction with other financial institutions. This is illustrated by Figure 2, which shows the Fed’s balance sheet composition, pre- and post-crisis.

Insert Figure 2 around here

More relevant yet to our purpose was the research carried out by the New York Fed into the nature of banks – both in terms of their relationship to shadow banking, and the composition of their revenues. This research, only published in 2012, in a special issue of the NY Fed’s research bulletin (Cetorelli et al, 2012), is remarkable in being avowedly innovative in terms of looking at institutions, rather than markets. It explains that most shadow banking activities have been supported and often driven by banks. As they note in the introduction, “banks have indeed adapted naturally to the changing model of intermediation, redefining their ‘production function’ while continuing to provide the type of services needed for intermediation to occur...” (2012: 3). In other words, as clearly emerges from this research, many shadow banking activities happen through banks, and with their participation. Banks have adapted by offering credit enhancement, guarantees, liquidity, origination, and all sorts of other activities, adapting their business models and drawing on new sources of revenues. And, at the corporate level, Bank Holding Companies (BHCs) have systematically increased their non-bank activities and share of revenue, so that “BHCs have allotted nonbank subsidiaries an increasingly important role in their activities, consistent with the view of adaptation through organizational changes.” (2012: 3). Figure 3 provides an illustration of the evolution of the total assets of BHCs, and the split between bank and non-bank companies, showing the significant shift in terms of the activity in BHCs.

Insert Figure 3 around here

The research by the New York Fed team showed some important patterns. Cetorelli & Oeristiani (2012) find that banks are dominant along several roles in the shadow banking world, and in particular in the securitization chain. They issue 56% of the loans, are a trustee for 95% of them, an underwriter of 92% of the loans, and they service 48% of them. Mandel, Morgan & Wei (2012) find
that they play a critical role as guarantors (a feature discussed by Acharya & Richardson, an a fact that led to Bear Stern’s demise, among others). Bord & Santos (2012) find that they are important
input feeders into the securitization chain. In other words, shadow banking contains a lot of banking
activities. Banks play a much more important role in securitized lending than their balance sheet
suggests. In other words, their business model in terms of how they generate revenues has shifted
and is dependent on the new securitized, market-mediated ecosystem, in ways that traditional
analyses of banking do not normally expect it to. Through their multiple roles in interfacing with
other participants in the shadow banking world, banks have become, in Cetorelli’s (2013: 250)
words, “private central bankers to important components of shadow banking activity throughout the
years of its growth”. If we shift our attention from banks to their more diversified parents, ie the
BHC, Copeland et al (2012) show the results of the change in the business model of banks, and of the
architecture of the (regulated) financial sector, visible through the increasing share or fee-based
activity.

The upshot of all this is a transformation of the sector, with the intermediation “roles are performed
in a new way, such that it becomes economically viable, and perhaps more efficient, for different
entities to specialize in providing different services... with regulated banking institutions ...crucially
involved in every step of the credit intermediation chain.” (Cetorelli, 2013: 253). These changes,
though, seem lost on many commentators who assume that shadow banking is a separate, often evil
and threatening activity, removed from banks. They seem to disregard the changes in banks’
business models, probably because such an institution-based view is not a normal approach in a
discipline more interested in markets. And while some of the existing research can offer insight into
the business models of banks (Acharya & Robertson, 2009; Acharya et al, 2013), this hasn’t been the
explicit focus to date.

During the crisis, the gap in our understanding of how banks make money- or even on how they are
structured as corporations was even more evident. Regulators lacked insight into the mechanics of
banks’ profits, or the choices they made as a result, partly as a result of their focus and training in
economics, as opposed to the business strategy of financial institutions. A senior JPMorgan Chase
executive privately confided that the NY Fed, at the height of the crisis, asked JPMC's strategy team
to teach its economists how banks make money. “They really didn’t understand the business models
– and as Geithner trusted Jamie [Dimon], we offered to help,” he noted. “Quite some time during
that period was spent educating regulators on banking business model basics and how this relates to
the shadow banking sector.”

Beyond academe, and possibly some of the regulators, these changes are somewhat better
understood and documented. Many popular books, some quite thoroughly researched, have
recorded changes in the structure and dynamics of the banking sector; they include investigative
works such as Tett (2009) or Sorkin (2009) and more focused histories such as Lowenstein (2009),
Fox (2009), McGee (2010) and Das (2011), as well as a well-researched documentary titled Inside Job
(2011). A more personal perspective is provided by memoirs-cum-analyses such as those written by
Treasury Secretary Paulson (2010) and FDIC Chairwoman Bair (2012). More detail is provided by
inquiries in the US Congress (the Financial Crisis Inquiry Commission, 2011), and in the UK (the
Banking Crisis Committee). On the legal side, more recent lawsuits against Goldman Sachs (2010)
and Standard & Poor’s (2013) have brought more evidence to light that would not otherwise be
made public. Administration reports drafted by firms such as New Century Financial, and related investigations, provide a remarkable level of detail on events in the US, the UK and other countries.

Such accounts may sometimes focus on the racy tales of greed and daring, but they do help us understand structural changes in historical context, despite their micro-focus. They also show us the way towards new ways to conceptualize the drivers of economic activity. Recent consulting reports (e.g., Oliver Wyman, 2010) also provide a more structured view on the transformation of banks’ business models, and document the transformation the sector has gone- and the changes looking ahead.

To sum up, financial innovation in the literature has been examined, but primarily in theoretical terms, and focused on products. The transformation of the architecture of the financial service sector has not been thoroughly researched – despite its massive importance as a driver of the crisis, and the wealth of anecdotal and historical evidence. Moreover, the change in banks’ business models has only been considered recently, and then only by a limited line of work. An explicit focus on business models and industry architectures could reveal new reasons why the last financial crisis came about, as well as how we might avoid the next.


Before we consider the evidence on financial innovation in business models and how this relates to the crisis, some background on the theoretical building blocks for our analysis is called for. We start with the premise that individuals are boundedly rational, and that anthropomorphizing actors such as “banks” or “regulators” and ascribing them rational behavior (forward-looking, calculative, long-term efficient) may be too heroic. Indeed, the recent crisis featured highly myopic actions that were far from rational – to the detriment of market participants. Given the fact that even the most sophisticated regulators didn’t have a firm grasp of the massive shadow banking system until after the crisis peaked, how confident are we on regulatory or business omniscience? Caution suggests we take a behavioral route. After all, as Rajan puts it:

> Each of the actors—bankers, politicians, the poor, foreign investors, economists, and central bankers—did what they thought was right. Indeed, a very real possibility is that key actors like politicians and bankers were guided unintentionally, by voting patterns and market approval respectively, into behavior that led inexorably toward the crisis. (2010: 18)

Bounded rationality doesn’t connote lack of intention to become rich, or lack of agency in trying to bring that about. It does, however, mean we pay particular attention to what drives individuals’ actions, and their competence. Clearly, actors can be highly competent within the restricted spheres of their actual activity. Evolutionary analysis focuses on the sources and limits of competence, locating them primarily in the way behavioral dispositions have been shaped by past learning and experience. Furthermore, actors are limited insofar as they focus their attention on whatever

---

3 Limited opportunities for learning imply actors of limited competence. To acknowledge that competence is based in learning and experience is to recognize that innovation, which presents new problems to actors, can be dangerous. Indeed, evolution itself is dangerous because it can make actors’ past learning obsolete. Many of the heralded “financial innovations” of the years preceding the crisis seem to have been of this character.
stimuli their organizational environment makes more salient. For an individual, the question of which division or part of the organization they belong to will shape what they think they should do, and what they attend to; it determines their skills and areas of competence. Between organizations, “what is rewarded around here” affects collective skills. A sense of “what clients want” (which, in turn, depends on who the clients are) and “how can we make money” (which also depends on the regulatory environment and the way the business model is structured) shapes trajectories of capability development aimed at resolving a specific, pragmatic, limited problem circumscribed by organizational goals and context.

This takes us to the role of structure, which determines competence and shapes agency. Across a whole sector, the options that individuals have are defined by the range of institutions that are regarded as acceptable, and prevailing norms (see Fligstein, 1990). This range is finite; at any time, there is a set of organizations with particular roles to play and particular ways of organizing and rewarding employees, or making a profit. That is to say, a limited set of business models are acceptable within the existing industry architecture.

However, architectures can be changed actors who want to change the “rules of the game”. Rules and roles in the financial sector evolve through the interplay between agents and economic, political and social institutions. Labor can be divided in several ways within a sector, so our analysis must look at the implications of these alternatives and their associated business models. As Jacobides et al (2006: 1203) argue:

An industry architecture is an abstract description of the economic agents within an economic system (in terms of economic behavior and the capabilities that support the feasible range of behaviors) and the relationships among those agents in terms of a minimal set of rules governing their arrangement, interconnections, and interdependence (the rules governing exchange among economic agents)... Architectures affect industry participants in ways that may be either anticipated and designed in, or unanticipated.

A final and crucial point about industry architectures is their reliance on rules that affect profit-making or reporting asset values – two features brought to the fore by the recent financial crisis. The Basel regulations stipulated how much capital banks had to hold against each type of loan. This made profit mechanics dependent on loan type. Moreover, that securitization was given a substantial regulatory bonus compared to holding loans on the books. So these rules were crucial in determining what paid off and what did not, and in shaping patterns of demand for loans. This, in turn, had an impact on firms’ business models and boundaries, as they responded to the signals from regulators rationally and myopically – but not necessarily dishonestly. Banks also worked within the regulations not to reduce their risk, but to pass it on to related entities (Structured Investment Vehicles) in order to maximize their own returns. While this was criticized by economists (see Acharya & Richardson, 2009b), it is precisely what we would expect profit-seeking institutions to do. As securitization grew, it also endogenously increased the pressures for fragmentation and vertical dis-integration of the sector.

The division of labor between different “parts” of a sector affects the way organizations and individuals develop skills and routines. It is well known that too much specialization can cause us to lose sight of the collective goal (Selznick, 1949). Subordinate goals displace the “big picture” because it is simpler, cognitively speaking, to identify with a limited, partial task. In the same way, we have
all encountered parochial views and cronyism, where local incentives and feedback have led people to neglect of collective efficiencies, instead focusing on their own, narrow part of the environment. Incentives and cognition interact, creating limitations that are substantial, and often inevitable.

This suggests that competence is always local; that is, it consists of performance as judged in the context of a set of tasks that characterizes a particular sub-sector. What constitutes “success” depends on the business model of the institution and what rules and regulations allow. In other words, the nature and measurement of “performance” is shaped by the architecture of the sector. And there is no reason to expect that what gets rewarded locally in the short term will be sustainable in the long term. The context for decisions and actions won’t change just because there are concerns over their viability. Better architectures will not emerge simply because the existing ones do not support the achievement of system-wide or superordinate goals, or because they are not sustainable in the long run – hence the regularity of banking crises.

Let us now focus on the operation of a business model and the nature of feedback. The term “business model” is still in flux; here, I take it to mean the expression of the value-generating activity system of a business (Teece, 2010; Zott, Amit, & Massa, 2011), and more specifically, the way in which it monetize its assets and capabilities. Business models, though, cannot operate in a social and institutional vacuum; pressures of institutional isomorphism (see DiMaggio & Powell, 1983) constrain the menu of choices. At any single point in time, there are a limited number of business models that may operate in a sector, all of them constrained by, and built on, the prevailing industry architecture(s). This, in turn, gives rise to a path of development with a given set of actors, who create systems of categorization and cognitive communities (Porac, Thomas, & Baden-Fuller, 1989) that help make sense of economic activity in the sector. Over time, new ways of monetizing an advantage emerge through a long and often arduous process that requires very substantial investments in infrastructure, as described in Baldwin (2008), Jacobides (2005), or North (1996). New markets change the architecture and impose new criteria for engaging in each part of the production process, creating a new collective outcome.

In financial services, the way firms monetize loans today bears little resemblance to the situation 20 years ago, even where the underlying product is unchanged. Securitization created not only new markets, but also new ways for new industry participants to monetize the benefits from a loan. Along with the new division of labor came a new set of rules, a new selection mechanism, and a new definition of “competence.” And while this innovation sounded exciting and forward-looking at the time, and was also supported by regulation, it was instrumental to the crisis.

Next, we turn to the nature and role of “feedback”. Feedback obviously operates at multiple levels of analysis. First, at the level of a corporation, the stock-market reaction to potential choices in terms of how a firm operates, funds itself, or directs its major business units, will shape its actions and choices. Views expressed by analysts, the media, and outsiders are critical for publically held, joint-stock companies. (The transformation of investment banks from partnerships into listed companies significantly altered how feedback operated, and what pressures were heeded; see, e.g., Morrison & Wilhelm, 2008.) At the level of a business unit, the nature of the business model and the prevailing industry architecture will affect feedback for particular courses of action. If a firm focuses on leveraging its solid ratings to issue Credit Default Swaps (CDSs), for instance (to take AIG’s example), and if the architecture is conducive to CDS growth, then feedback will suggest that the
business unit should produce more CDSs. Within the organization, particular divisions may receive feedback (based on transfer pricing mechanisms, key performance indicators, and performance criteria) that defines “success” and dictates what a unit can and cannot do (e.g., what is the cost of capital and how it is computed, what are the metrics used, etc). Finally, individuals know what actions are allowed, encouraged, rewarded, or recognized (in terms of bonus, career progression, expected behavior etc), and this will determine their individual choices, which, if they’re lucky or they can convince their organizations, will also drive corporate actions.

These different levels of feedback are not independent from each other. Career dynamics, for instance, transcend individual organizations. So an ambitious young banker who knows that a bad bet and a potential exit from Goldman Sachs may lead to a cushy job at a lower-tier investment bank or corporate will be tempted to partly ignore the feedback she receives at Goldman. Individual choices may take into account opportunities throughout the industry architecture, and feedback may operate in subtle ways.

However it may operate, though, I posit (and will aim to demonstrate in the next section) that feedback, not foresight, drives economic behavior (see Jacobides & Winter, 2010). This represents a subtle but critical shift from existing theory’s emphasis on forward-looking analysis, where even behavioral mistakes (such as “animal spirits” and excitement) concern views about the future rather feedback on past behavior.

Feedback may operate over different time horizons. While rational agents might be expected to balance short- with long-term feedback, the time when feedback emerges, and its short-term bias, tends to be an important element. Time horizons may actually help explain the dynamic instabilities we observe in financial services. The core issue is that there are very few ways to dismiss silly ideas in financial services, either at prototype stage or during actual operation. An instrument carrying unknown risk will only reveal its danger later, just as a new way of originating a loan won’t be tested until there is a cyclical hardship. More subtly, some financial services innovations may make sense in one context (e.g., the use of brokers and warehousing specialists to produce conservative, conforming loans for Fannie Mae and Freddie Mac before standards were lowered) but not in another (sub-prime mortgages where the same arrangements might become problematic).

Also, pursuing a change that seems acceptable based on local feedback (“produce more loans of given types”) might lead to very bad outcomes in another area (the periphery of prime loans), when conditions (house prices) change. So we might (as we did) see many innovations whose true costs in terms of risk were not visible for some time. In that sense, the financial services sector can be more “dangerous” than many others, not only because of its vital role in the economy, but also because of its particular fragility and the evolutionary risks of delayed selection. New business models, with their attendant new ways of measuring performance, can produce misleading indicators that may shape both the sector and the wider economy by determining which activities “make sense”. Changing industry architectures can throw up untested models that hold the seeds of failure, while regulators unwittingly set rules that encourage the very behaviors they want to avoid.

**Beyond feedback: Differential Growth, Imitation and Fear of Bankruptcy in Industry Change**

While feedback is important, it is not the only selection mechanism, in particular at the level of organizations and their role in a population. So, to complete our theoretical detour, some more
foundation on evolutionary economics (see Nelson & Winter, 1982) is offered here, especially as it departs from the more traditional textbook variety of economics: First, there is selection via differential growth, through the industry as an ecosystem “evolves” and “carries the genes” of effective innovation. Second, there is selection via “imitation”, or what is known in evolutionary biology as a Lamarckian process. Third, as economists know, there is selection by “fear of extinction” – a rational, forward-looking concern with the threat of bankruptcy. This last area has received a lot of attention stemming from the “Too Big To Fail” problem, and empirically may have been vastly overplayed. Likewise, the belief that the only problem with innovation has been inappropriate incentives for senior bankers merits further examination. So let us look at each of these factors in turn.

Selection via differential growth. The relative prevalence of particular practices is the result of the relative growth of the firms that use them. Profitable firms are likely to be rewarded by the capital markets, allowing them to expand the scope of application of their practices. This remains true even when the profits flow from risky activities, as long as the downsides have not yet materialized. Ownership and governance structures play a key role in this regard, as they mediate between the ability to finance growth and the policy decision to pursue it. In theory, these structures could help the firm by cautioning against expanding risky but profitable activities – in practice, however, temptation proves hard to resist. In the absence of alarming feedback, there is little reason to expect risky practices to be abandoned, because of the strong tendency for actors to “satisfice” with respect to their established dispositions (Simon 1955, 1987; Winter, 1971). Risk might be abandoned by firms whose realized results hamper their growth, but those results may come too late in a context where the feedback is delayed – such as financial services. Risk abandonment is even less likely when a number of firms are pursuing similar programs, validating each other’s behavior by bidding up the prices of risky assets.

Selection via imitation. Practices can become more prevalent simply because more firms adopt them. According to a very large empirical literature on the diffusion of innovations, practices are more likely to be adopted when they are seen as contributing to realized success for actors who currently employ it. When the practice involves the quest for profit from risky investments, the same logic of delayed feedback operates for imitation processes as for growth.

For relatively prudent firms, the siren song of temptation is amplified by a decision mechanism featured in some evolutionary models: Recognizing that its returns are inferior to its peers’, a firm often ends up adopting its peers’ policies in order to chase the “easy money”. Of course, the same mechanisms also operate constructively in the economy all the time, diffusing practices that are in fact superior. But in the build-up to the crisis, toxic practices were promoted at the expense of prudent ones.

This observation needs to be tied in to an understanding of incentives that differs from stylized models. The problem is that, for all the rhetoric about preferring stability, and even the recent calls to jail reckless bankers, the system is set to allow risk to creep in, as prudent corporate practices can be a career hazard for executives. Consider Jamie Dimon of JP Morgan Chase, still widely considered to be one of the most competent managers in the sector. If the crisis had not struck in 2007, and the meltdown of 2008 had been deferred by two years or so, it is quite conceivable that Dimon and
other similarly disposed executives would have been replaced under the pressure of “the market” for growth and additional returns.

**Selection via bankruptcy.** The idea that there is a long-run environmental constraint on firm behavior is most straightforwardly illustrated in the case of bankruptcy, which played such a prominent role in the crisis. In principle, the ability of a firm to pay its bills when due is a fundamental consideration that affects firms’ survival through the institution of bankruptcy law. Sometimes, disastrous behavioral dispositions are checked effectively; carriers of “inefficient genes” (including dispositions to accept poor or excessive investment risks) are weeded out as inefficient firms become bankrupt. However, this mechanism is by no means effective across the full range of potential “disastrous dispositions.”

In order to fulfill the constructive evolutionary role suggested by the “market test” concept, a selection mechanism should reflect the direct financial consequences generated over time by the repetitive exercise of the firm’s routines. In fact, as I will explain, bankruptcy is often detached from those consequences. Thus, whatever a firm’s inability to pay its bills in the short run may mean for its survival, its status as a verdict on the firm’s established ways of doing things is much less clear.

In the first instance, this is because cash derived from past successes does not always accumulate in some special account as a “rainy day fund” (Apple being the exception that proves the rule). Instead it is paid out, irreversibly, as dividends or bonuses (in a world of limited liability). So firms typically operate with a degree of financial vulnerability that is partly attributable to their payout policy, rather than the quality of their productive performance.

This is not a puzzling phenomenon when we look at the economic institutions of the financial world and the incentives they create. Indeed, those incentives tend to produce underfunding much more broadly – of the reserves that would permit an insurer to fulfil its contractual obligations, of pension benefits, deficient bank capital, and so forth. Incentives to “take the money and run” are a feature of the institutional system – grounded in self-interest, but also supported by concerns about what will happen to the money if you don’t take it and run.

When we examine the roots of a crisis, we must recognize that systemic hazards can be created by relatively competent and foresighted actors. This point is well recognized under the heading of “moral hazard,” but its significance in evolutionary analysis is less remarked. Firm behaviors that are objectively reasonable and prudent are not necessarily viable from an evolutionary viewpoint. Firms following sound business practices that generate fluctuating returns may succumb to short-term adversity – and that adversity may stem from competitive pressure from growing firms who are unburdened by prudence. Evolution can favor the reckless, at least in the short run. When evolution finally renders a negative verdict in survival terms, not all of the reckless (or the passive investors who supported them) will suffer in wealth terms.

The second source of detachment is that the state of a firm’s balance sheet is ultimately a reflection of expectations about the future. Here, the word “ultimately” acknowledges a contrast with what is true for most firms, most of the time. Within the intrinsic limitations of financial accounting, and subject to the qualifications about payout mentioned previously, a firm’s balance sheet ordinarily provides a valuable, though imperfect, indication of the cumulative effect of past operations. In conditions of exceptional threat or opportunity, however, balance-sheet transactions (or the
absence of anticipated ones, such as loan roll-overs) may change the picture dramatically in a short period.

The financial crisis illustrates the interplay of cumulative experience and payout practices, as recorded in firms’ balance sheets, and changing expectations about the future. Consider, for example, the somewhat contentious question of why the US government intervened to save Bear Stearns but not Lehman Brothers. According to accounts of the frantic quest for a deal to save Lehman, a major factor was the poor quality of Lehman assets (Cohan, 2009; Stewart, 2009). It seems that, even against a background of severely depressed expectations, the “hole” that new equity had to fill was revealed to be unexpectedly large, particularly because of the weakness in commercial real-estate loans. This posed a problem for the potential buyer, and also forced the Fed to confront the question of the legal limits to its discretion in accepting dubious collateral. Thus, Lehman’s “balance sheet” (at the detailed level) did provide a cumulative record of fatally flawed operations in the past, operations that ultimately “caught up with it.” That record defeated a major public and private effort to fend off the rising systemic risk by converging on new expectations and saving the firm.

The “detachment” point here is that the operation of these mechanisms on the firm’s balance sheet cannot be regarded simply as a “verdict of experience” on its operations up to that time. In principle, expectations and/or payout practices could run counter to the verdict of experience, in either direction – producing failure where experience indicated success, and success where it indicated failure. At the practical level, of course, experience to date is typically a significant influence on expectations, as the Lehman example illustrates.

The third reason for the detachment relates to the bluntness of the selection mechanism as a form of feedback on organizational performance. A firm, especially a large and diversified firm, has a mind-boggling array of practices, while financial results and bankruptcy-based selection processes operate on the single amalgamated entity. The case of AIG is instructive. A generally well-run (if under-reserved) insurance firm allowed one of its branches to run an illusorily profitable business out of London, issuing credit default swaps. This CDS business was effectively subsidized by the insurance firm by way of the latter’s good credit rating. If not for government intervention, the CDS business would have led the firm to bankruptcy; in the event, it wound up in partial public ownership. Similarly, Lehman had healthy areas of business unrelated to the risky real estate positions that brought it down. Arguably, this sort of bundling of strength and weakness in a single firm should not be a problem – provided it is possible to separate them, and distinguish the practices that contributed to failure from those that did not.

While efforts to accomplish such separation are often a feature of bankruptcy proceedings (including Lehman’s), there are major obstacles to success. First, it is often very difficult to identify dangerous practices; a practice that is healthy in one environment can become toxic in another. And second, even when it is known what is valuable and what is not, the complexity of large financial institutions makes it hard to separate them. On the administrative side, financial institutions may be too complicated to be properly understood by their managers, as the JPMorganChase “Big Whale” incident showed: a seemingly risk-free part of a bank was amassing risk that was not understood by its own management. Legal issues may also make it hard to separate the “healthy” and “unhealthy” parts of a company, if disaster strikes. And while the recent regulatory overhaul in the US tries to
force firms to do so ex ante through living wills (see Acharya et al, 2011; Herring and Carmassi, 2009), legal separation isn’t enough. Reputation muddies the waters regardless of legal status, which is why Bear Sterns took responsibility for its legally separable REIT SIVs during the crisis (which were its downfall), and why Citibank, in 2007, brought an astounding $87 billion in Structured Investment Vehicle positions onto its balance sheet (Jablecki and Machaj 2011, 223–24). Strictly speaking, Citibank wasn’t obliged to take that step — but it felt that it had to, in order to maintain its reputation (and its wholesale funding flow). Banks are more systemically intertwined and less separable than what economic theory predicts — for sound economic reasons.

In sum, if bankruptcy is conceived as the principal “enforcer” of economic selection, there are important qualifications to be made. I do not mean to overthrow the basic idea of the market test, but there is serious doubt over whether it will always support competent actors and suppress incompetent ones. More fundamentally, the structuring of the evolutionary process by limited liability implies a strong warning about the prospects for benign evolution when negative outcomes appear initially as risk, and realizations are long deferred.

Individual-level incentives, and the challenge of feedback vs foresight. The belief that the fear of bankruptcy will discipline economic actors in the short term neglects the institutional realities of actual firms. Similarly, the strongly held belief that individual incentives are what can be blamed for the crisis lacks factual support. Dick Fuld personally lost $900m as a result of Lehman Brothers’ collapse, while the managers of Bear Stearns’ two main sub-prime funds (which later took the firm with them), Ralph Cioffi and Matthew Tannin, were both heavily invested up until the collapse. It is hard to overlook the fact that both Lehman and Bear were investment banks with higher-than-usual employee stock-ownership percentage, and that Lehman had a five-year vesting period and no step-function —close to what is recommended nowadays as a safer set of compensation incentives. Larger-scale research also suggests that, if anything, there is a slight negative correlation between the strength of CEO compensation and bank failures (see Fahlenbrach & Stulz, 2009). Yet this isn’t to diminish the importance of incentives in banking overall, or to suggest that incentives and actions were not way off what society would want them to be. The questions are how exactly they were off, why they changed, and what we can learn from them.

Changes in Industry Architecture, Business Models, and the Financial Crisis

With the theoretical background set, we can now look at how financial innovation affected business models and industry architectures, and how this encouraged the practices that led to the crisis. Two points need explanation. First, why did lending standards fell so egregiously, so that vast sums were lent to people who should never have received them, for properties and projects that were not viable as collateral? And second, why did banks find themselves owning securitized or structured products that ultimately brought them to their knees, jeopardizing economic stability, since they should have known their risks, having been instrumental in their production?

To answer these questions, I will focus first on one of the central causes of the financial crisis: inappropriate sub-prime lending and the unexpected collapse of securitized products. But, to explain how these worked, we should first consider mortgage lending and trace the process of change in this
industry. Briefly, the sector shifted from a vertically integrated structure, where banks or savings & loans associations originated, held, and kept the loans, to one where the loan was originated by a mortgage broker, sold to a mortgage banker, and sent on to a securitizer to be tranched and sold to the capital markets. A number of events enabled this, including the government intervention as a “market creator” with its agencies Ginnie Mae, Fannie Mae, and Freddie Mac, followed by the changing economics of capital supply and demand. To this, however, we must add the Basel banking regulations, which only asked for banks to hold half the reserve, meaning they could make their equity go further with securitized loans on their books.

Let us now move to more proximate culprits: Private-Label Mortgage Backed Securities (PLMBs). These were bundles of loans owned by a SIV (on the asset side). The SIV would then sell “tranches” of ownership of these loans to investors – the riskiest most cheaply, and then the safer ones at higher prices. Given that a share in a pool of loans is less risky than an individual loan (as there is diversification of the default risk), and given that structuring loans in the SIV allows to re-engineer which tranche gets “hit” first, this allowed, presumably, for gains from trade to be had, as well as allowing a higher-risk set of loans to be tranching into low-risk, medium-risk, and high-risk components. The question here is why this system – lauded before the crisis, but criticized for being “tarnished” or “inappropriately used” after it (Acharya & Richardson, 2009; Allen, 2010; Frame & White, 2013) – led to such disastrous outcomes. Much of the answer, it appears, lies in understanding the demand for loans. Why were investors, including banks, so keen to be involved in PLMBs? Another popular instrument was the Collateralized Debt Obligation, which is similar to MBS but composed of tranches of MBS rather than individual loans – i.e. already-securitized loans, then repackaged into a CDO. Figure 4 provides a guide to these structures for the non-expert reader.

Include Figure 4 about here

Why was there demand for these loans? The canonical theoretical answers include increased liquidity and the possibility of trading the loan as needed, the possibility of keeping the exact risk profile each institution needed (and as such the latent gains from specialization and tranching), and flexibility (see Ashcroft & Schuermann, 2008; Llewellyn et al, 2009). But additional answers have to involve both fees for structuring them and, most of all, banking regulation and the gains from regulatory arbitrage. The reason was that Basel regulation provided relief for banks that held securitized loans: meaning that they needed to tie down less capital for the same volume of loans, thus increasing their ROE. The very same residential mortgage loan, once securitized, required half the capital to be held in reserve. This meant securitization could print money for the banks. In addition to this, some global differences in what standards banks had to follow created yet more gains from trade. In the US, banks adhered to Basel I, the initial set of rules on banking supervision. By the beginning of the century, though, their European counterparts switched to the more exacting Basel II rule; and this meant that highly rated securities got additional relief, requiring even less capital to be held. In Europe, Basel II meant that it made more sense for a bank to hold an AAA security on its books than to lend funds directly; and in the US, since there was no special relief for

4 This process is discussed elsewhere (Jacobides, 2005), and I will assume that it is familiar territory; the interested reader may also want to consult Tett (2009) for a vivid discussion of CDOs and PLMBs. For a more academic discussion, see Barth, 2010; on an overview, Zandi, 2008; on securitization, Ashcraft & Schuermann, 2008.
AAA securitized tranches, the other parts of the mortgage loan would suit just fine.\(^5\) As a result, chopping up loans and distributing different tranches in different sides of the Atlantic made eminent sense for the banks, which is precisely why so many continental European and UK banks found themselves holding US-originated tranches of securities- which had the extra interest in being Dollar denominated, which might have helped a European banks’ currency diversification. Against this regulatory background, it is no surprise that demand for PLMBSs in the United States exploded – doubling in 2001, and racking up growth rates of almost 50% in each of the following three years – as banks and their CIOs shifted funds from loans to AAA PLMBSs (regarded as safe on the strength of their ratings).

Figure 5a provides an example of banks’ calculations of the capital that had to be held for a $1 million loan. As we can see, a whole (AAA) mortgage would require 4% of the capital lent – i.e. $40,000. A securitized (AAA) mortgage (before the Recourse Rule) would require $20,000. Under Basel II, or the Recourse Rule, the capital reduced to $16,000. If the bank also bought a CDS (from AIG, for example, as they often did), the capital could go to exactly zero. This last feature explains the tremendous demand for CDS from AIG- which cumulatively exceeded $500 Billion. It was this financial smoke and mirrors, fully facilitated by regulators, that led to the raging demand for PLMBS, but also for all of the synthetic products that were based on them- and in particular, the Collateralized Debt Obligations, which were built using MBS, or ABS in general, and as such needed their “raw material”.

Include Figure 5a and 5b around here

At the same time, on the investment side, pension funds and fund managers the world over became increasingly more “commercial”, seeking higher returns in increasingly global markets – provided their investments had the “right” risk ratings. Regulators of pension funds, perhaps the most important source of capital to be managed, endorsed the ratings system, and maintained that a particular percentage of the investable sums should be AAA. This meant that demand for AAA securities was high. And this encouraged CDOs as well. The reason is simple: A pool of MBS had a maximum percentage of tranches that could be deemed AAA. Yet, for the riskier tranches of this, and similar, MBS, structurers (seeking fees) could create a CDO, made of, say, BBB MBS, which would have a AAA tranche, and then a BBB tranche, etc, leading to yet more securities labelled AAA.

From the perspective of investors who needed returns, this was manna from heaven. Figure 5b shows an example from the Ohio Police & Fire, a pension fund in the American Midwest that was prudently restricted to AAA securities. That gave it several options. It could invest in GSE (Agency), i.e. Fannie Mae or Freddie Mac, which had a near-zero yield over the Treasury. It could invest in a corporate bond, paying around 11 bps (basis points, ie percentages of a percentage point) over the Treasury (ie, the presumably risk-free debt issued by the US government.) Or it could take an AAA tranche of a PLMBS, which offered 18 bps over the Treasury. Those seven-hundredths of a percent proved decisive, and between 2005 and 2006 the composition of the Ohio Police & Fire portfolio was pushed to PLMBA. A similar comparison could be made with AAA tranches of CDOs, which offered yet again a slightly higher yield- but, as it turned out, much higher risk, for reasons I will outline

\(^5\) US banks though, complaining of their relative disadvantage, managed to get the Fed to agree to the Recourse Rule, which gave holders of AAA mortgage-backed securities the right to hold less capital than they otherwise would.
below. But this was not consequential when the investments were made. The fund’s managers were after yield – subject to ratings, which they equated with risk assessment – with the blessings of pensions regulators. And even if they (managers of Ohio Police & Fire Pensions) hadn’t come up with the idea of these new products to invest in, they copied what they saw as an appropriate behavior. The craze for derivatives and securitization was in full swing.

These developments brought credit rating agencies (CRAs), the de facto arbiters of risk, to the fore. Originally part of publishing houses, CRAs started out as compilers of information. They would check data on potential borrowers or issuers of debentures and report, confidentially, to investors. Such were the humble roots of the “ratings” industry, with three non-regulated rating firms (Moody’s, Standard & Poor’s, and Fitch) assessing anything from mortgage loans to sovereign risk.

Ratings agencies were paid by the investor until the 1970s. During that decade, photocopiers became widely used, and illegal duplication of their reports put a substantial dent in ratings agencies’ revenues. So they changed to charging issuers instead – essentially, changing their business model. As more complicated financial products emerged, the ratings agencies’ purview broadened, and became instrumental in the development of securitization. As Partnoy (1999) notes, the US government gave the three CRAs “regulatory license”, which came to extend de facto beyond the US. CRAs were granted quasi-regulatory authority over the investment decisions of various fiduciaries, including the rating of securitized debentures. Now, as long as rating agencies’ policies were sufficiently permissive, new markets could easily take off, with the indirect blessing of pension-fund and bank regulators. Loan originators were naturally eager to “baptize” some of their loans as AAA, as this gave them access to broader pools of capital. And ratings agencies were all too happy to gain new business and increase their profits. They spent less and less time reviewing the increasingly complicated instruments and competed for issuers’ business – all the while enjoying an oligopoly of just three Nationally Recognized CRAs that was protected by the government.

A cool headed analysis of the crisis points to ratings agencies’ incompetence, even though there has been little work on what can happen moving forward- and even there, some of the suggestions look downright puzzling, as I will explain in the concluding section. The poaching of their staff by those who structured the debt (a tactic to “game” the system and secure the best possible rating), and the ability to find ways to get the rating a structurer wanted were commonly held secrets. By any rational account, ratings around the epicenter of the crisis were deeply flawed. Ratings agencies had rated Lehman as solid investment grade just before it went bust; Fannie Mae and Freddie Mac were AAA the week before they were declared insolvent; and Bear Sterns was also solid investment grade four days before its collapse. To give a sense of magnitude of the downgrade (and, as ratings were included in the regulatory calculations for capital, of the impact for banks), it is worth considering the example Friedman & Kraus (2011: 103) report: “triple-A CDO bonds from the 2007 vintage were, on average, downgraded to CCC by 2008 (Barnett-Hart 2009, 23–24), producing a 4,900 percent increase in the capital required for these bonds.”

If ratings were instrumental in creating the demand for these financial innovations, and for the change on how participants made money (their business model), another part of this related to how profit was booked. In this regard, we have to consider the role of Mark to Market, accepted as a practice in 1993 as part of the Financial Accounting Standards (FAS) 115 and refined (in a manner that would precipitate the market collapse) with FAS 157. Marking to Market, as we will see,
affected the crisis dynamics by forcing fire-sales, and by (ex post facto) perhaps unduly reducing the “booked price” of assets. At the outset, though, it also allowed paper profits to be booked as if they were real, making the fanciful pricing of these instruments an organizational reality. Now, the units that created these products could recognize the profit they created, delighting the institutions that saw their profits and balance sheets buoyed as a result. This innovation reinforced the cycle of feedback and action, in both its positive (growth) phase and in its later, negative spiral. While motivated by a sensible reading of economic theory and a desire for transparency, Mark to Market changed the organizational incentives, pushing actors to generate more derivatives, book profits, and demonstrate benefits.

As securitized structures took hold, and capital market actors learned how to describe and understand them, ever more complicated structures emerged, with new firms seeking new ways to make money. Rating agencies would assess the risk of each new financial instrument, and each tranche of each instrument, in return for revenue. The only thing that could stand in the way of permissive assessments was a concern with “potential reputation issues”. As far as immediate feedback was concerned, competing agencies were clearly expected to show “constructive behavior” that would be welcomed by the firms footing the bill.

Given the ability to tranche sub-prime mortgages and call them “AAA”, and the dearth of quality “prime” mortgages, the market for securitized sub-prime mortgages quickly took off. An environment of low interest rates and low yields, plus the desire for growth, led to a fight for higher rates of return. “AAA” PLMBS/ CDO tranches offered somewhat (not substantially!) higher returns than “AAA” national debt, making it irrational for bankers and (most) funds to refrain from investing in them. They also helped reduce regulatory capital, which had a real economic impact for the bank. Yet these presumed AAA structures were becoming riskier and riskier.

Whether those structuring the debts knew the whole enterprise was unsustainable is probably unknowable. What is crucial to note is that the initial positive feedback on these new business models drove the behavior of the sector as a whole. (Interestingly, much of the feedback system, and many of the rules for participants such as ratings agencies, have yet to change.)

Include Figure 6 around here

Figure 6 provides a real example, from a Goldman Sachs structure, of operations during the go-go years. On the left are the risk attributes of the pool of loans – and it looks very risky indeed, with 83% having low scores in terms of their creditworthiness, less than a 660 FICO score (even assuming the information was accurately reported), most of the loans concentrated in two risky, bubble areas (Florida and California), 88% being “non-conventional”, i.e. not fixed-rate, known to be riskier; 57% took the loan motivated to “release equity” from their houses, etc. Despite this, 97% of the slices were deemed to be investment grade, and 79% were rated AAA. How? Quite simply, because of the raters’ model. And the model came from the assumption that defaults would not be correlated, so if there was a 10% probability of default for each of the, say, 10 loans that constitute the loan pool, and if all defaults “hit” the lower tranche (set to be 10%) then higher-up tranches would have a default likelihood of less than 1%. Of course, as it turned out, correlations were very high, and defaults in such structures were an order of magnitude higher than the ratings agencies anticipated.
The fact that some of the loans produced on the basis of dodgy assessments by the rating agencies were eventually risky was no secret. If one looks at the evidence of what investors and firms knew about the securities that were being issued, information wasn’t suppressed, and surely not to the extent that many commentators would have us believe. As Foote et al (2012) note,

To nonexperts, one of the most confusing things about the mortgage securities market is that issuers were quite careful to document the extent to which they did not document a borrower’s income and assets. Loans were typically given a four-letter code that informed investors whether the information about income (I) and assets (A) were either verified (V), stated (S) or not collected at all (N). For example, the code SIVA meant stated income-verified assets. The crucial point here is that investors knowingly bought low-doc/no-doc loans. In fact, we now know that lenders provided loans to borrowers with damaged credit without documenting their incomes not because of any after-the-fact forensic investigation, but rather because lenders broadcasted this information to prospective investors.

Investors had access not only to important data, but also to tools that allowed them to use these data to price securities. The MBS and CDOs that contained the mortgages (or the mortgage risk) appeared complex on the surface, but they were in fact straightforward to model. Most investors used a program called Intex that coded all of the rules from a prospectus for the allocation of cash flows to different tranches of a deal. To forecast the performance of a deal, an investor would input into Intex a scenario for the performance of the underlying loans. Intex would then deliver cash flows, taking into account all of the complex features of the deal, including so-called overcollateralization accounts and the treatment of interest income earned on loans that were paid off in the middle of a month.

Cordell, Huang, and Williams (2011) shows that using Intex, one could accurately measure the losses and value of ABS CDOs in real time throughout the crisis. (2012: 14-15)

Drawing on a 2005 analysis by Lehman about its own products, they also note that:

Lehman analysts were not alone in understanding the strong relationship between house prices and losses on subprime loans. As Gerardi et al. (2008) show, analysts at other banks reached similar conclusions and were similarly accurate in their forecasts conditional on house price appreciation outcomes. JPMorgan analysts used MSA-level variation in losses on 2003 subprime originations to produce remarkably accurate predictions about losses (Flanagan et al. 2006a). A UBS slide presentation about subprime securities in the fall of 2005 was subtitled, “It’s (Almost) All About Home Prices” (Zimmerman 2005). [UBS was the first bank to be hit by the AAA tranches of sub-prime mortgage CDOs its own unit had structured.] The Lehman analysis, and others like it, are crucial documents for anyone hoping to understand why investors lined up to buy securities backed by subprime loans... the analysis shows that investors knew about the significant risk inherent in subprime deals. (2012: 16)

...the use of such loans exploded from 2002 to 2006. The growth of reduced-documentation lending is often presented as Exhibit A in narratives of how the declining standards of mortgage lenders caused the housing crisis. What this growth really shows is the declining standards of investors. These loans were clearly marked as “stated income, stated assets”

---

6 The NINA loan is the basis for the apocryphal “NINJA” loan that is often used as an example of excesses in the boom-era mortgage market. NINJA supposedly stood for “no-income, no job, no assets,” but no such loan ever existed. Also, the NINA code, which did exist, did not signify a loan to a borrower with no income. Rather, the code signified that the lender had no information about the borrower’s income.
loans, so investors knew what they were getting. In particular, investors knew that borrowers were likely to have inflated their incomes and assets. Yet investors purchased the loans anyway because they expected these loans to be profitable. (2012: 28)

This analysis suggests that the truth was there for those who cared to look. From a review of other qualitative evidence, though, it seems that this knowledge was dispersed very unevenly. The infamous “banker from Dortmund” – i.e. unsophisticated investor hungry for yield – has been shown to play a role by uncritically consuming risky but highly rated securities. Real-life “Dortmund bankers” included the German Vereinsbanken and Landesbanken. But regardless of how well informed the users of AAA-rated securities were, it is crucial to stress that ratings – through the use of Basel regulations, as well as prudential regulations for insurers – became the defining criteria of how returns would be generated. So people’s foresight and understanding of risk – or lack thereof – was ultimately less compelling than the immediate feedback they received.

If we dig deeper into the roots of the crisis, we soon unearth further evidence of how an innovation can lead to disaster in the absence of short-term negative feedback. The worst damage, by some margin, was sustained by AAA tranches of CDOs, as opposed to AAA tranches of asset-backed securities (ABSs) such as MBSs. Fundamentally, both instruments are similar, as we can see from Figure 7, which shows how CDOs were based on MBSs (as illustrated in Figure 4), and provides concrete examples, in terms of the tranches and ratings, for 2007. But, as Figure 8 shows, the impairments in CDOs and MBSS are quite markedly different – CDO tranches (of the same CRA rating) having suffered much more

Insert Figures 7 and 8 around here

As Figures 4 and 7 show, ABSs are securities, tranched in terms of their risk, and collateralized by mortgage loans (MBSs) or other assets more generally. CDOs of ABSs are structured vehicles, tranching in a similar fashion. The underlying concept in both an ABS and a CDO is the same, and their presumed benefits (fees for those who structure, reduction of risk through risk correlation, and tranching for the investors, presumably) were also similar. The only difference is the “base material” from which they were made – but in practical terms, this difference was very consequential, as we can see in Figure 8 (see Foote et al, 2012). To forecast ABS performance, analysts, both in issuers and in rating agencies, could model default probabilities of individual loans and aggregate them up. These predictions were, to use the technical jargon, “structural”. Yet with the CDOs coming to the stage, the “raw material” was different: tranches of ABSs. And, more important, Wall Street’s CDO analysts, and raters, were organizationally independent of the researchers analyzing asset pools, and mortgage pools in particular. As Foote et al (2012) note:

The CDO analysts did not devise structural models for the individual BBB-rated tranches in their CDOs. Instead, they essentially skipped ahead to the step of asking how correlated BBB defaults were likely to be. To do this, the CDO analysts looked at past financial market data, including the prices of default insurance on individual BBB tranches.... The different outlooks could have stemmed from the backgrounds of the two groups of analysts. CDOs were originally constructed from various corporate bonds, for which historical correlations have been excellent guides to future performance, even during the recent crisis. CDO analysts probably assumed that the same type of historical analysis would also work well for CDOs made up of subprime mortgage bonds. By contrast, mortgage analysts were trained to model mortgages individually, and they had the data and the tools to do so.
In hindsight, it is hard to see how two groups of analysts could work in close proximity at the same financial institution and not notice the colossal dissonance implied by their respective analyses. For example, during the peak of the mortgage boom, mortgage analysts at UBS published reports showing that even a small decline in house prices would lead to losses that would wipe out the BBB-rated securities of subprime deals (Zimmerman 2005). At the same time, UBS was both an issuer of and a major investor in ABS CDOs, which would be nearly worthless if this decline occurred. Why didn’t the mortgage analysts tell their coworkers how sensitive the CDOs would be to a price decline? This question goes to the heart of why the financial crisis occurred. The answer may well involve the information and incentive structures present inside Wall Street firms. Employees who could recognize the iceberg looming in front of the ship may not have been listened to, or they may not have had the right incentives to speak up. If so, then the information and incentive problems giving rise to the crisis would not have existed between mortgage industry insiders and outsiders, as the inside job story suggests. Rather, these problems would have existed between different floors of the same Wall Street firm. (2012: 23-25).

Going back to the role of ratings agencies, the passage above illustrates the role of inertia, as well as the way business practices take hold in the absence of adverse feedback. It also illustrates the risks of looking at “one firm” as a rational actor, making choices, motivated by its medium-term future. What we see is the familiar picture of organizations as groups of people with differing agendas, vantage points, and skills, driven by their positions in the organization; each following what they think is appropriate and is locally rewarded; and ensconced in routines and processes that determine their actual menu of options (March & Simon, 1958; Nelson & Winter, 1982). It also suggests that at the top of the organization lies not a shrewd and far-sighted strategy but a fairly simple set of short-term priorities that is driven by local demands (boards, analysts, the stock market, and external perceptions).

At the aggregate level, the lax assessments of rating agencies, the increasingly loose investment policies, and the optimistic views on valuation took hold because there was no feedback to stop them, and because local pressures were shaping the corporate context. In the mid-2000s, for instance, McGraw-Hill, the parent company of Standard & Poor’s, was under significant pressure: the internet was denting revenues, particularly hurting BusinessWeek publications. As recounted by a leading former executive, the corporate parent’s CEO had set clear priorities for revenue generation, with the express view “not to be brought any bad news”. It was therefore no surprise that the CEO of Standard & Poor’s boasted in 2006 of nearly trebling revenue per employee – meaning fewer people were doing less diligence and getting paid more by the issuers. Also, as we have seen, the US (and, by extension, international capital markets) had given regulatory license to only three raters, meaning that a cozy monopoly had formed, and raters were not held accountable for their views. Indeed, the recent effort by the US Administration to reverse that, suing Standard & Poor’s, has been vehemently decried by orthodox economists in the conservative camp, who for some reason regard the current arrangement as satisfactory, and Standard & Poor’s as blameless.

These feedback pressures at the top level obviously trickled down, painting a clear picture of “what to go for” at the level of those making the ratings. Evidence is emerging from the internal email trail of ratings agency executives, corroborated by the testimony of the leaders of ratings agencies under oath (see Taibbi, 2013). To give but a brief example, consider the case of Stephen McCabe, an analyst with S&P. In a 2005 email relating to a deal (for Cheyne) that S&P were rating, he suggested (to an investor!) that “looking at the numbers it is quite obvious that we have just stuck our
preverbal [sic] finger in the air!" To a colleague, he wrote "Remember the dream of being able to
defend the model with sound empirical research? If we are just going to make it up in order to rate
deals, then quants are of precious little value." More consequentially, Frank Parisi, Standard & Poor’s
chief credit officer for structured finance, said that the model that his company used to rate
residential mortgage-backed securities in 2005 and 2006 was only marginally more accurate than "if
you just simply flipped a coin". Richard Gugliada, head of S&P’s CDO operations, was later
questioned under oath about why this was happening, leading to the following exchange when the
government side asked why the company was slow to implement updates to its model for evaluating
CDOs:

Q: Is it fair to say that Standard & Poor’s goal of preserving an increasing market share and
profits from ratings fees influence the development of the updates to the CDO evaluator?
A: In part, correct.
Q: The main reason to avoid a reduction in the noninvestment grade ratings business was to
preserve S&P’s market share in that category, correct?
A: Correct.

This example brings us back to the main issue with rationality, feedback, and economic activities.
Imagine that some in Standard & Poor’s or Goldman Sachs saw that this model didn’t make sense.
What could they do? If they were in S&P, they could resign – and some did. But where could they
work if all three competitors had similar mores? And if they weren’t (with US government blessing!) allowed to put together a competing ratings agency? If they were investors, and if they had
unbelievably deep pockets, as hedge-fund manager Paulson had, and a keen sense of timing,
perhaps they could bet against it. But most agents do not have a few billion of their own to gamble.
So what would they do? They would have to continue, even while knowing that the house of cards
might come crashing down tomorrow. Their actions wouldn’t be driven by limited cognition and intent, or the survival risks of their institution in the short term, but by feedback – feedback that was
structured by the immediate pressures for survival that firms faced, or the profit mechanics imposed
by regulations such as Basel. For these individuals, or indeed for these businesses, it is feedback, not foresight, that drives economic actions.

Making sense of the crisis, the role of financial innovation, and business-model change

We can now come full circle to our explanation of the crisis, linking it to the vicious circle of causality
between changes in business models, shifts of industry architecture and financial innovation. The
regulation-induced demand from the issuers of PLMBSs and CDOs, and the local incentives inherent
in the established industry architecture, pushed the system to continue structuring and selling
financial products. So hedge funds and banks creating ABSs and CDOs created pressure to brokers
and the originating arms of banks to issue still more loans – the “raw material” that could be sliced
and diced. Originators responded by relaxing their underwriting standards, extending credit to all
comers and often becoming pure predators. Meanwhile, risk was creeping in.

Originators are a good example of firms driven by feedback rather than foresight. Countrywide, for
example, made profits of $11 billion before being absorbed as it failed. Eight of the top 10 sub-prime
mortgage originators in 2006 were gone by 2009. These firms received no government support and
were never going to be bailed out. In truth, though, the threat of bankruptcy would not have
“disciplined” them early on, nor would the possibility of a bailout (had it existed) have affected their
actions at any stage. Viewed from the present, future bankruptcy is actually a very distant and largely ineffective driver of economic action — despite economists’ fascination with it. The real incentives of the boundedly rational actors we observed were very different and focused on the short term; they were legitimate, logical responses to the architecture that was built with the forbearance of regulators.

This takes us to a very different interpretation of events than the one offered in some of the existing literature. Acharya & Richardson (2009), for instance, argue that

> If bad mortgages sold to investors hoodwinked by AAA ratings were all there was to it, those investors would have absorbed their losses and the financial system would have moved forward. The crash would have been no different, in principle, than the bursting of the tech bubble in 2000. In our view, what made the current crisis so much worse than the crash of 2000 was the behavior of many of the large, complex financial institutions (LCFIs) — the universal banks, investment banks, insurance companies, and (in rare cases) even hedge funds — that dominate the financial industry. These LCFIs ignored their own business model of securitization and chose not to transfer the credit risk to other investors.”

While it is easy agree with some of their observations, my own reading of the same facts differs: I would submit that these large, complex financial institutions didn’t “ignore” their own business model of securitization, choosing not to transfer the risk. First, these institutions didn’t work as one, human-like, rational actor, interested in their long-term survival. Division of labor within the company, and the adoption of rules in different parts of it, meant that some (on the investment side) were tasked with growth; and some (on the risk side) were tasked with ensuring that the pressures to grow wouldn’t jeopardize the firm overall. The problem is that while the part of the firm that produced CDOs might have known that there were some issues with these securities, the risk officers, much as the rest of the industry, had accepted that ratings would suffice as surrogates of risk. And regulations allowed banks to make more money by investing in AAA (or “super-senior”) tranches and make more money without any visible downside. The forces of feedback at the corporate level clearly pushed for “better use of the balance sheet”; those involved in structuring were clearly incentivized to produce more; and those in risk management would only be alerted if something unambiguously indicated trouble ahead. The fact that everyone knew there was something off with those securities wasn’t enough to create the feedback required to change the practice.

Consider UBS, a Swiss bank that was once quite conservative. It found itself close to bankruptcy as a result of its bet on “supersenior” (presumably safe) AAA tranches of CDOs that its own structuring team would produce – a case discussed by Clementi et al (2009: 198-200). UBS’s Investment Bank would issue CDOs, and make fees ranging from 30 to 150 basis points in order to warehouse (ie, buy and package the loan), structure, and then sell them. But, starting in 2006, UBS didn’t sell all of its CDOs; it kept the “best” tranches (the supersenior) ones on its own balance sheet. And while Clementi et al point out that any production would be good for the CDO structuring team (and also the P&L of the Investment Bank), it doesn’t explain why they started holding on to them. To get a better understanding, consider the reasoning provided by Huw Jenkins, the CEO of UBS’ Investment Bank in 2005–2007 and one of the first senior victims of the crisis. As he (in personal communication),):
In 2005 I was hired with a mandate to grow what was seen as a sleepy Swiss bank, and transform it into a dynamic, effective financial institution. This involved both building teams in investment banking, and moving in areas such as Brazil, but also being more effective in terms of using our balance sheet. One of the opportunities we saw was to invest in supersenior tranches of CDOs. We did think they were safe; and, regardless, we sent it over to risk. There’s 1,000 people working there, with PhDs, and they’re supposed to be experts. We asked them, and they were happy with it. So what we did, was keep to the mandate – which was driven by analyst and press pressures at the time. We kept the supersenior tranches – it looked like a sensible investment, and we got no objection from our risk teams.

As we know, CDO AAA tranches were seriously mis-appraised by raters, for reasons already covered. But this gives us a sense of how decisions were made, in UBS and other banks, where divisionalization and the local nature of the feedback process was instrumental in shaping collective decision. If we anthropomorphize banks, as Acharya & Richardson implicitly did, we might be too optimistic about organizational coherence, or about the omniscience of the people at the top. And it does appear that many of the corporate leaders were caught by surprise, so that the idea that leaders of banks knew and willfully didn’t transfer risk simply does not stack up with the evidence. It is worth re-iterating the examples of Bear Sterns (and their leaders’ continued investment, including fresh personal funds, right up to the final collapse); of Dick Fuld and his $900 million loss with Lehman’s collapse; and any number of post-crisis accounts from seasoned veterans, including Hank Paulson, a former CEO of Goldman Sachs, suggest that the surprise was genuine. Both Bear Sterns, and in particular Lehman Brothers had bonus arrangements with clawbacks and long vesting periods, more so than other firms; and this is part of the broader pattern we know empirically, but dismiss casually, that no link can be found between the nature of bonus schemes and bank failure (see Fahlenbrach & Stulz, 2009.) This casts some serious doubt on the mainstream reading of the drivers of the bank collapse we observed, in particular ascribing the blame to greedy and reckless bankers intentionally deceiving society. It does appear that the banking system did just as we should have expected with the architecture that regulators put in place.

That said, there was plenty of reckless agency around- as it is meant to be in the context of financial services. The. One telling example is AIG, and its Financial Products division, the business unit selling CDSs (which, due to regulatory structure, it did not need to provision against). AIG/FP was busy selling protection, through CDSs, and booking much of it as profit. Leveraging AIG’s AAA rating, its CDSs allowed firms with sub-par loans to convert them into AAA, as AIG would stand behind them. It also allowed AIG to underwrite the senior tranches of CDOs (which eventually did collapse) without any shadow cost of capital for AIG, as their default was considered “statistically unlikely”. This was a pure money-making machine for those working in the unit – but also for AIG, in the short-run. AIG/FP, employing a few people in the City of London, was responsible for 17% of the firm’s operating income, and underwrote half a trillion (!) dollars’ worth of CDSs. Joe Cassano, the head of AIG/FP, pocketed over $300 million; the firm had little reason to spoil the party. Its structure, deficiencies in governance and stock market pressures and, crucially, the rules in the sector, encouraged it to sit back and enjoy the revenue ride.

The fact that both individuals and, on aggregate, organizations can be greedy, especially in financial services, is not novel. What is important is to understand how organizations allow such desire for wealth to be turned into a bet that might bring the company down. The question is how individuals
and organizations changed the rules of the game at the level of the firm (e.g., through compensation and authority) and at the level of the sector (through supervision and regulation).

The recent history of financial services shows actors trying to shape the environment they inhabit, making short-term gains at the cost of long-term viability. Yet this is neither unique, nor strange. As we noted earlier, firms don’t just compete within boundaries (as most economic models suggest), but also try to shape the rules by which they play. This idea does not sit well with mainstream economics, which focuses on either “responding to incentives” or “reckless behavior” – but here we have something that is neither of those, but rather an effort to shape rules and regulations to make more profit. The exemption of derivatives from regulation, in the Financial Modernization Bill in the US in 1999; the exemption of Investment Banks from supervision (relating to such activities) in 2004, through the use of internal risk models; the changes in Basel I with the Recourse Rule – all of these destabilized the system not because of criminal recklessness, but by creating an “evolutionary time-bomb”. The problem was that regulators didn’t appreciate that point, being in thrall to a “market-focused” ideology, and failing to see how financial innovation altered the business model and the incentives and feedback of main players in the market. They also failed to foresee the interaction of the very regulations that would end up amplifying the crisis – Basel and Mark to Market.7

Seen from this perspective, the crisis was both completely understandable and entirely endogenous. It was not primarily the result of “irrational exuberance” or monetary largesse, but of changing organizational incentives that elude the standard analysis of economic accounts of the crisis. Indeed, the credit expansion that led to the asset bubble may have originated in the banking sector. This view differs from the usual analysis of crises (see, e.g., Minsky, 1986; Kindleberger, 1978, 2005) in that it suggests that the structure of the banking sector, and business models in financial services, can cause both the misallocation of credit and an increase in the total level of credit. This is the necessary condition for an asset bubble to form, which then generates further pressure for the financial service sector to lend.

In terms of financial innovation, this implies that there is an important feedback loop between new products and relationships and business models in the banking sector. As new products emerge, banks change their business models, and the architecture of the sector evolves to accommodate them; and as these change, institutions advance yet more financial innovations that are consistent with their own needs and ability to grow and profit. None of this guarantees that the changes will be beneficial for society. Regulators should consider innovations in the broadest sense: in terms of business models and how financial service firms make money – which, as Cetorelli et al (2012)

7 Finally, we should note the role of regulators, and their own constraints and feedback. As we discuss elsewhere (see Jacobides et al, 2014), prior to the financial crisis, the global financial regulatory system had evolved into a collection of agencies that largely specialized in a particular geography and vertical line of business. Such a system would indeed make sense in a traditional financial system with limited spillovers between verticals or geographies. However, such a structure would not work well if the system was both dynamic and interconnected. Furthermore, the political context within which regulation has happened is fraught with adverse incentives. Politicians, wanting to show themselves spendthrift and decisive in times of crises, have often cut back on the budgets and pay levels of regulators much as the demand for supervision and the skill required to supervise has increased, with qualified supervisors aiming to become consultants later on in their careers so as to finally be able to afford a good living. Also, political pressures to “show toughness” to a universally unpopular sector, amplified by the bureaucratic reflexes of governmental organizations has been to propose a regulations, whose end results (or even aims) are not always clear.
demonstrate, have changed substantially. This should be a key priority, not just so we can understand past events, but so we can move forward.

**Implications for Policy and Regulation**

The ideas in this paper challenge mainstream views that regard *foresight* as the main issue. The focus on structure, along with the view that human boundedness plays a crucial role, provides the basis for a fresh reading of the evidence and a powerful caution against some of the policies being considered.

Firstly, “Too Big To Fail” (TBTF), a central point in the current policy agenda, seems misplaced at best. Do we really believe that the prospect of failure, distant as it is, shapes balance-sheet composition or credit practices in an organization? In the previous section I provided a detailed discussion of why adaptation due to the fear of bankruptcy does not work in the real world. At the margin, dealing with TBTF might be a good idea – but the hope that it will provide true discipline to the sector is closer to faith than science. Therefore, it seems that regulatory energy would be better directed towards other areas; a more behaviorally friendly view should yield more effective candidates.

Second, structural features – such as what Basel II and III encourage firms to keep on the balance sheet – have a huge impact on firms’ actions. Yet little work looks at how these regulations will push firms to change their balance sheets, or push their non-bank partners to provide the instruments that banks may want. Of course, part of the current rules may serve regulators’ political masters. Basel II (and III) still see sovereign debt as risk-free, with the travesty being that rational banks in Greece, Spain, Italy, and France take on sovereign debt rather than providing loans to businesses – since it makes sense from a feedback perspective. The European banking sector has returned to the time of the Medicis, with banks bankrolling governments for their ventures. We urgently need a better understanding of how regulations often unwittingly affect the business models and actions of participating firms.

Third, financial innovations should be seen through the prism of their potential application. Thinking about how an innovation might fare in the context of a general equilibrium model, without actual institutions, information asymmetry, bounded rationality and feedback-based behavior will not help. Rather, we need to consider how the innovation will affect *particular institutions and the way they make money*. We should also consider, as Calomiris (2013: 27-8) points out when considering new regulations, “how, within [any] proposed reform, market participants will find it difficult to circumvent the reform, and why regulators and supervisors will have personal incentives to enforce it in a manner that will make it effective.”

Fourth, the role of rating agencies has yet to be addressed, and the business model of banks, including their role in shadow banking, has yet to be mapped and properly understood. It beggars belief that some of the suggestions on how to improve rating agencies include “an ‘advertiser pays’ model whereby the rating information is provided with advertisement attached, and the advertising covers the cost of the informational content”, or “a joint venture model whereby... major investment banks would form a collectively financed joint venture to produce ratings” (Richardson & White, 2009: 111) and that “it seems inevitable that [more intrusive regulation] would reduce the flexibility of the rating industry and could discourage innovation and creativity in business models and in
developing better ways of assessing the default probabilities of bonds” (ibid, 2009: 105). The problem has been that we have seen too much creativity in such a nodal part of the financial services sector, and that the results of the existing business model have not been thought through. The analysis in this paper suggests that this reasoning entirely misses the point, and is downright dangerous, even as a knee-jerk reaction against regulation. What we probably need, both for credit risk agencies and participants in the financial services sector, is better monitoring and quite possibly a restriction in terms of the potential business model.

In that regard, the mechanics of the payment model require more thought. Current proposals to mandate a rotation of raters, or pressure to “open up” the previously closed oligopoly of raters may, on the margin, be useful, but they fail to address the rootcrux of the issue by rethinking what drives feedback (and as such behavior and effectiveness) for CRAs. A bolder solution is likely to require a rethink of CRAs’ business modelwithin, including them maintaining some liability for the views expressed. Calomiris (2011) suggests that there be the possibility of legal damages for egregious negligence. Given the pivotal role of rating agencies, and their being enshrined in regulation through Basel III and Solvency II, we need to either remove their quasi-regulatory license or treat them in a way analogous to auditors. Thought also needs to go into the governance and organizational model of credit rating agencies, which shapes local feedback. The idea that these can be publically listed companies, or worse still, divisions of non-financial corporates with pressures for margins and contribution, seems to be asking for trouble, given recent history. Requiring the CRAs to morph into a limited partnership, which would tie the CRA executives’ fortunes to the performance of the CRA, could help improve CRA ethos and efficacy. For the financial system to work, given the fact that capital markets are so central, we need to consider the challenges with CRAs organizational, governance and business models. Alarmingly, precious little has happened in this direction.

Fifth, we should consider the institutional realities of regulators, and their ability to do what we would expect them to. Regulators, like any other organization, can suffer from goal displacement, i.e. shifting from the original objective to a narrower, or even different one. They, too, have particular feedback to respond to, and we must devise a way of making sure that they operate more effectively – especially as the variety of regulators creates opportunities for yet more new business models through regulatory arbitrage. Whether this involves a “supra-regulatory body” (see Barth et al, 2012) to ensure that the guardians of finance are doing their job, or a structured way to review the performance of regulators, we need to be mindful of the organizational realities of regulation and the risks of their own biases and limitations.

Further suggestions on the regulatory implications of this evolutionary and structural view are provided in greater detail elsewhere (see Jacobides et al, 2014). Briefly, the first priority should be to create a robust map of the industry architecture in financial services; take the same principles underlying the work of Poszar et al (2010), to document rules, roles, and relationships, and assess business models in the sector. Particular attention should be paid to business models that generate

---

8 Yet even auditors may need to have their role stepped up. As a recent report in the House of Lords in the UK showed, auditors should share a part of the blame for the financial debacle; the report (House of Lords, 2011,ch. 6) is sobering. The recommendations for tightening up the role of auditors appear well founded- even if not likely to be all implemented.
particularly high returns or growth rates in the short term, to assess their impact on the stability of the system.

The re-orientation of regulatory attention from financial parameters to the dynamics of business models will surely not be easy, and it will require significant up-skilling of regulators. In particular, we need to complement regulators based in economics with those able to understand and track how banks make money, focused on their business models, and how these collectively constitute the industry’s architecture. The one-way street from regulation to banking, career-wise, would best be reversed, so that regulators co-opt poachers as gamekeepers. Pay levels would probably have to be adjusted to make such a shift plausible, but it should yield a solid return on investment. The use of derivatives and speculators as signals to focus regulatory attention, the regular review of regulation to consider its practical usefulness, the focus on scenario planning, and ongoing impact assessment rather than just principles-based regulation are all suggestions worth considering.

Overall, until we shift our focus from the abstractions of economic theory to real-world business structures, we will continue to jeopardize our economies — and our understanding of them. We need to treat financial services as a sector, and consider what its dynamics and evolutionary properties may be, and focus on the institutions within it. In other words, we need to focus on business models and industry architectures, and consider the mechanisms of feedback to be able to be effective.

This approach might help balance the current, potentially misplaced faith on individual malfeasance being the end-all of regulation. Recent initiatives, such as the UK’s Committee on Banking Standards Report of June 2013, have hinted that jail terms for reckless bankers might be needed to address the problem. They surely will address the problem of voter satisfaction from the politicians who have endorsed, and will publicize, this report. On the margin, they may prompt some people to listen to their consciences and leave banking — allowing other, more optimistic (or shrewd) operators to take their place. But, our analysis suggests, alone these wont fix the systemic problem that led, and will soon lead anew, to the wrong kind of loans and the wrong level of credit to be extended to the economy. The image of a rational, calculating banker who needs to be motivated with a stick and a carrot, as opposed to a decomposed, locally rational and globally inefficient system that needs fixing, will come to haunt us again — and sooner than we think. The next bubble, enabled by loose monetary policy, may already have started.

**Implications for Research and Theory**

This analysis provides an illustration of how, at the level of a sector, the existing rules, roles and relationships, i.e. the industry architecture, can change, with potentially devastating effects. It shows how firms are trying to change the architecture to their advantage; how their efforts are myopic and local, often focusing on short-term benefits or narrow areas of competence and interest. Most important, the paper explains the operation of feedback, and suggests that it is feedback, as opposed to foresight, that drives economic behavior. It also shows how feedback interacts with other mechanisms of selection (imitation, differential growth, the threat of extinction).

In terms of the analysis of the crisis, this paper offers a very different reading of the historical record from the one found in much existing research. Methodologically, it focuses on structures, industry architectures, business models, and feedback as significant drivers of success or failure in the financial sector (also, see Jacobides, 2014). At the level of theory, the main contribution is to show
how feedback, as opposed to foresight, drives economic behavior. It also extends existing work on business models to consider their systemic implications, an issue that has evaded attention so far. This can and should be generalized well beyond the confines of financial services. From telecommunication and media to healthcare and education, a bewildering variety of business models has emerged, transforming previously sleepy, integrated, and simple sectors into dynamic ecosystems with rapidly changing architectures and business models. The question arises, how should we think of them? And how should we regulate them? Which is the best way to balance innovation with stability? One of the silver linings of the 2008 crisis would be to use it as an example of the worst-case scenario, and ensure something similar does not happen in other settings. The need for an applied evolutionary, structural analysis appears to transcend sectors and settings. As such, a rich research agenda lies ahead.
References


Barth JR. 2010. The Rise and Fall of the U.S. Mortgage and Credit Markets. New York: Wiley


Figure 1: The map of the different entities and links in the Financial Service Sector

Source: Poszar et al, 2010; original size, 3 x 4 feet
Figure 2: The evolution of the Fed’s Balance Sheet

Figure 3: Growth in Assets of Bank and Nonbank Subsidiaries of Bank Holding Companies and of Other Financial Intermediaries

Board of Governors of the Federal Reserve System, Flow of Funds Accounts and Consolidated Financial Statements of Bank Holding Companies (FR Y-9C data)

Source: Cetorelli, et al. (2012)
Figure 4: A graphical illustration of ABS (here, RMBS) and CDO structures for the non-expert reader
Figure 5a Regulatory Arbitrage, Capital Held, and PLMBS for AAA Loans or MBS

<table>
<thead>
<tr>
<th>Mortgage</th>
<th>AAA-rated non-Agency MBS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Originator (e.g., WaMu)</td>
<td>MBS issuer (e.g., JPM)</td>
</tr>
<tr>
<td>Basel 1</td>
<td>Basel 2</td>
</tr>
<tr>
<td>Capital regime</td>
<td>Capital regime</td>
</tr>
<tr>
<td>Basel 1</td>
<td>Basel 2</td>
</tr>
<tr>
<td>Risk-weighting</td>
<td>50%</td>
</tr>
<tr>
<td>Capital required (per $1MM)</td>
<td>$40,000</td>
</tr>
</tbody>
</table>
| Note: Assumes 8% Core Tier 1 Capital. Source: JPMorganChase

Figure 5b: Returns and Portfolio Composition for Insurers for PLMBS AAA tranches

Example: Ohio Police & Fire – MBS portfolio

<table>
<thead>
<tr>
<th>100% = Non-Agency</th>
<th>12/31/05</th>
<th>12/31/06</th>
</tr>
</thead>
<tbody>
<tr>
<td>33%</td>
<td>$735MM</td>
<td>$799MM</td>
</tr>
<tr>
<td>67%</td>
<td>12/31/05</td>
<td>12/31/06</td>
</tr>
<tr>
<td>7%</td>
<td>93%</td>
<td></td>
</tr>
</tbody>
</table>

Comparative AAA spreads (bps, June 2006)

<table>
<thead>
<tr>
<th>Non-Agency (ABX)</th>
<th>Corporate (CDS derived)</th>
<th>GSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>11</td>
<td>Near risk-free rate</td>
</tr>
</tbody>
</table>

Source: Understanding the Securitization of Subprime Mortgage Credit, FRBNY, 2008
Figure 6: An example of a Structured Deal: A Goldman Sachs PLMBS Issuance

Goldman Sachs – GSAMP Trust 2006-NC2
(% of mortgages by value; 100% = $881.5MM)

Weighted average coupon at origination (all tranches) = LIBOR + 23 bps

- Mainly non-traditional (% non-fixed rate): 88%
- Borrowers had poor credit (% FICO <660): 83%
- Concentrated in limited geographies (% in CA & FL): 49%
- Motivated by desire to access equity (% cash-out ReFis): 57%
- Leaving limited equity in home (% CLTV >80%): 38%
- Other tranches (% of MBS principal): BBB 3%, A 4%, AA 5%, AAA 9%

97% investment grade
Figure 7: An example of ABS (PLMBS) based on Subprimes, and subsequent CDO structures, with Credit Ratings Tranche Breakdown

Figure 8: Impairment rates of similarly rated tranches of MBS and CDO drawing on MBS

Source: Moody’s Structured Finance Default Risk Services

Note: There are 3823 tranches of Aaa MBS amounting to $203 billion in total value. There are 2403 tranches of Baa MBS amounting to $26 billion in total value.

2006 Vintage MBSs

Source: Foote et al (2012), p 56

2006 Vintage CDOs

Source: Foote et al (2012), p 56