When Crises are Failures: Contested Metrics in International Finance and Development

JACQUELINE BEST

University of Ottawa

NOTE: This is a prepublication version of an article that is forthcoming in *International Political Sociology*, 2016.

What happens if we look at events like the 2007-2008 financial crisis not just as crises but also as failures? This shift is a productive one, opening up new ways of understanding the debates and policy responses that followed the recent crisis. For this was not just a crisis: it was also a spectacular failure. Moreover, it was a contested failure: the kind of failure that made key actors question the metrics through which they measured success and failure. Through a comparative analysis of recent failures in international development and international finance, this paper argues that looking at these crises as contested failures provides us with a better understanding of the epistemic underpinnings of certain crises and the politics of the responses to them. Focusing on failure also allows us to see that policymakers have become more preoccupied with the possibility of failure in the aftermath of these crises—and that they have become increasingly cautious in response. This paper concludes by exploring the implications of the increasing prevalence of "fail-safe" approaches to policy, a cautious approach to policymaking that may in fact increase the odds of future failures.

International relations scholars love a good crisis. The 2007-08 global financial crisis was no exception, producing a great wealth of analyses along with a renaissance in social theories of crisis. But what if we were to shift the focus of our analysis slightly, and look at these moments not just as crises but also as failures? I will suggest in this paper that this shift is potentially very productive, as it opens up new ways of understanding the debates and policy responses that followed the recent crisis. For this was not just a crisis: it was also a spectacular failure. Moreover, it was the kind of failure that made key actors question the metrics through which they measured success and failure. By focusing on this kind of failure, we can gain a better understanding of the epistemic underpinnings of certain crises and the politics of the responses to them.

Although the main focus of this paper is therefore the recent global financial crisis, I will analyze it with the help of a rather different case drawn from international development. The various organizations and actors involved in the governance of international development have also suffered from several crises in recent years, as the "lost decades" in Sub-Saharan Africa (UNDP 1999: 99) and the Asian financial crisis led them to redefine development success around the concept of "aid effectiveness" (World Bank 1998; OECD 2005). By examining how debates about success and failure in development circles contributed to these shifts in governance practices, I will develop a theoretical framework that I can then use to examine the recent financial crisis.

Through this comparative analysis, I will develop two main arguments. First, I will suggest that, as in the case of international development, the 2008 financial crisis was not just a crisis, but was also a particular kind of failure: it was a *contested failure*.[1] The meaning and significance of most failures are subject to debate. For contested failures, those conflicts are so profound that they lead to debates about the very metrics through which success and failure are defined. Looking at the recent crisis as a contested failure helps us to understand the public's interest during the crisis in the mundane question of how financial risks were measured, as well as to appreciate how those metrics have (and have not) changed in its aftermath. Second, I will suggest that policymakers have become more preoccupied with the possibility of failure since the 2008 crisis—recognizing the uncertainties of the financial system and all too aware of the possibility of future crises. This marks a shift from previous crises, which were followed by the return of a more confident re-establishment of expert authority.

In developing these arguments, I will be engaging with a range of different scholars in political economy and social theory. My emphasis on the metrics of success and failure is inspired by the work of Michel Callon (1998), Donald MacKenzie (2006, 2009) and Andrew Barry (2002, 2012), who have pointed to the performative and political nature of measurement processes, while my focus on provisional expertise draws on the work of Niklas Luhmann (1998a,b). I will bring these social theorists into conversation with several strands of political economic analysis, including the rich scholarship on the narration of crises (Hay 1996; Blyth 2002; Widmaier et al. 2007; Brassett and Clarke 2012; Hay 2013) and recent writings on the 2007-08 global financial crisis.[2]

I am interested in particular in addressing an apparent paradox suggested by analyses of the crisis. As several scholars have pointed out, the crisis has produced an important epistemological shift in economic thinking, forcing policymakers to take greater account of the limits of their models in the face of uncertainty and complexity, particularly through the concept of systemic risk (Baker 2013; Datz 2013); yet, as Eric Helleiner (2014a) notes, this has nonetheless been a "status quo crisis," while Martin Wolf (2010) has described the most significant global regulatory reform—Basel III—as "the mouse that did not roar." How can we reconcile what Andrew Baker describes, following Peter

---

[1] I am building here on a concept that I first developed in *Governing Failure*, where I apply it to the case of international development governance (Best 2014).

[2] There is an enormous wealth of scholarship on the financial crisis. A good selection of early analysis is included in special issues of the *Cambridge Journal of Economics* and *New Political Economy*: (Blankenburg and Palma 2009; Brassett et al. 2010).

Hall, as a "third order change" in the epistemic framework for financial governance with such limited changes in policy in response? While some scholars have argued that this incrementalism can be explained by the bureaucratic dynamics involved in implementing this new ideational framework (Moschella and Tsingou 2013), this paper suggests that policymakers' preoccupation with failure has played a key role in producing this conservative policy response.

The remainder of this paper can be seen as a multi-stage elaboration of this argument. I will begin by developing a theory of the role of contested failures in international politics. I will then briefly discuss the case of international development, tracing the changes that have taken place in its governance in recent decades and the role of contested failures in those shifts. I will go on to examine the recent financial crisis in more detail, considering what we can learn when we view it through the lens of failure. I will then consider the implications of the growing preoccupation with failure in both development and financial circles. I will conclude by reflecting on some of the paradoxical dynamics of the growing prevalence of "fail-safe" approaches to policy, whose cautious approach to policymaking may in fact increase the odds of future failures.

Theorizing Failure

Counting Crises as Failures

As Colin Hay (1996) so aptly argued, crises are always in the eye of the beholder: it is how they are narrated that determines the kinds of policy changes that they make possible. In the aftermath of the recent global financial crisis, Janet Roitman (2013) has made a similar argument about the role of crisis as a narrative device that opens up some pathways for action and forecloses others.

This paper builds on those insights by considering what happens when crises are narrated as failures: how this particular form of narrative affects the response to the crisis and the debates around it. When crises are the effect of hurricane-force winds, civil war in a neighbouring country or an economic meltdown on the other side of the planet, they are not necessarily seen as failures. They become failures when they are perceived as a product of human error, as was the response to Hurricane Katrina in 2005 or the recent financial crisis. Although failures generally have an objective reality, like crises, they must be narrated in order to have a particular social impact.

What then counts as a failure? After all, the language of failure is everywhere: personal failure, failed states, failed policies. These are each distinct forms of failure. Yet they are all defined in relation to a particular standard of success, whether it is a meaningful life, a capacity for sovereign action or a desired policy outcome. Failures are always defined in relation to what they are not. This may be a straightforward binary distinction, as in formal tests or evaluations in which you either pass or fail, or it might be a matter of degree. A state, for example, does not fail overnight: it is only after a long process of decline that it is given that (highly contested) label. Whether the determination of failure is simple or complex, to call something a failure you must rely on a particular set of *metrics*.

As scholars working in science and technology studies (STS) and the social studies of finance have demonstrated, metrics are powerful social devices. Both Donald MacKenzie and Michel Callon have pointed to the performative power of key economic theories and equations which, by making disparate things measurable, commensurable and exchangeable, have made a whole host of different kinds of markets possible—whether in options, swaps or CO_2 (Callon 1998; MacKenzie 2006, 2009).

For something to be measured, it has to be framed: some things are included in the calculation and others are excluded as non-relevant—as "overflows" in Callon's language (1998)—or as externalities in the vocabulary of economics. The calculations involved in assessing a state's fragility in the "Failed State Index," for example, combine qualitative and quantitative assessments in twelve different issue areas, each of which involves certain inclusions and exclusions.[3] These prior decisions about what gets counted are black-boxed—concealed from view and taken for granted—in the final index which distills over 100 different sub-indicators into a single number.[4] Although metrics are highly technical, they are thus far from apolitical. As Bruno Latour (1987: 62) has argued, a great deal of social and political work is required to make something appear to be beyond contestation.

Contested Failures

In most cases, the metrics underpinning particular definitions of success and failure are taken for granted. But sometimes, if a failure is serious enough, or if failures are repeated over and over, those metrics themselves come into question. People begin to ask whether the failure occurred partly because they did not have the right definition of success.

I am calling those failures that open up this kind of debate about what counts as success and failure *contested failures*. Because failures imply responsibility, most failures involve debate and disagreement about who is to blame and for what. But only some failures lead to this kind of meta-debate about the metrics by which we define success and failure. In politics, these contested failures force policymakers and the wider community to re-examine not just the policy problems themselves but also the measures that they use to evaluate them. These moments of debate are important. They are highly technical—focusing on the nuts and bolts of policy evaluation. Yet they are also fundamentally political—since they force us to ask what we want success to look like and whether we can really know when we have found it.

Contested failures have a great deal in common with what STS scholars call "knowledge controversies" (Callon 1998: 260-261; Barry 2012). In such controversies, the very metrics used for evaluating evidence themselves become the subject of debate—

---

[3] This index, put together by the non-governmental organization, Fund for Peace, attempts to quantify and rank the degree of failure in the world's states. The index was renamed the "Fragile State Index" in 2014.

[4] The term "to black-box" is used by Callon and Latour to describe the process through which the indeterminacy of scientific findings are translated into taken for granted "facts." (Latour 1987; 1999; Callon 1998).

as in recent debates about climate change. As Barry (2012: 328) suggests, although measurement techniques often act to depoliticize issues, any system of measurement is necessarily fragile and can therefore become the subject of new forms of conflict. Because the world always exceeds the categories through which we seek to comprehend it, "measurement and calculation can have the effect of disrupting the frame of politics" (Barry 2002: 268).[5] Contested failures, I would argue, are just such moments in which the depoliticizing power of measurement becomes, for a time, repoliticized, as people ask themselves what should be counted when judging success and failure. Such failures bring a wider public into what were previously expert debates, leading policymakers to integrate some of their broader concerns into the definition of policy success, if only for a time.

A Preoccupation with Failure

What happens once the failure has been identified, the debates about metrics are over, and the lessons learned? Organizations do not generally enjoy epistemic uncertainty, and their members therefore work hard to re-establish the authority of their expertise, closing up the black boxes once more.

This time around, however, this certainty has not fully returned. In both international development and financial circles, there is far more attention to the uncertainty and volatility of political and economic dynamics than in the past, complicating efforts to re-establish definitive policy prescriptions. In her analysis of crisis, Roitman (2013: 9) suggests that crises have become a chronic condition, no longer existing as discrete events but as an ongoing malaise. Focusing on the growing preoccupation with failure allows us to see how the pervasiveness of crisis manifests itself in epistemological terms, as a growing awareness of the limits of knowledge. It is as if, rather than fully managing to conceal all of the measurement assumptions in black boxes, policymakers have been left with grey boxes in which a few unanswered questions remain partially visible, an occasional irritant that keeps things unsettled.

This preoccupation with failure fosters what I have described elsewhere as a more provisional approach to governance—a style of governance that anticipates and seeks to prevent future failures (Best 2014). An awareness of the possibility of failure can produce two different kinds of response. On the one hand, it can lead to an open-ended, even experimental, approach to problems that recognizes and responds creatively to the limits of knowledge. A number of scholars have described the spread of a more deliberative, open-ended "experimentalist governance" in other contexts, as an adaptive response to the growing uncertainty of the policy environment.[6]

On the other hand, rather than accepting and adapting to the possibility of failure, policymakers can take a more cautious, even cynical, approach that seeks to hedge against potential failures. In pointing to this second tendency, I am drawing inspiration from the sociologist, Niklas Luhmann (1998a: 69), who suggested that modernity is

---

[5] I also discuss this dynamic, drawing on Max Weber, Sheldon Wolin and Barry, in: (Best 2014: 70-73).

[6] Charles Sabel and Jonathan Zeitlin first pointed to this dynamic in the European Union. (Sabel and Zeitlin 2008; De Búrca et al. 2014)

characterized by the rise of a kind of provisional expertise that can always be revised after the fact, allowing experts to be right even when they are wrong. If your risk assessment shows that there is a twenty percent chance of failure, and that turns out to be the ultimate result, then you are not really to blame: "If the improbable happens, one can defend oneself with the argument that one decided correctly, namely in a risk-rational manner" (1998a: 72). As I will elaborate further below, this kind of approach to managing failure tends to produce a more conservative policy response.

In the remainder of this paper, I will draw on this theoretical framework to analyze the cases of international development and finance, considering in each case whether and how crises are narrated as failures, whether they are contested, and how key policy metrics have (or have not) changed as a result. I will consider the implications of the growing preoccupation with failure in the final part of the paper.

Failure and the Transformation of Development Governance

The 1990s was not an easy decade for international development organizations, which were faced with several significant policy failures including the Asian financial crisis, which forced millions back into poverty (Ramesh 2009: 82), and the "lost decade" for development in Sub-Saharan Africa, where poverty remained unabated or even worsened (UNDP 1999: 99).[7] Both signaled the failure of development policies to achieve their goals in terms of economic growth and poverty reduction (Feldstein 1998; Stiglitz 1998). The international financial institutions' (IFIs) own internal evaluations also pointed to policy failures: the success rates of World Bank policies began to decline in the 1990s, from over 80 to 65 percent by the early 1990s (OED 1994). At the same time, IMF staff and Board members became preoccupied with the number of countries who had become chronic users of the IMF resources (Boughton 2001: 618-619).

These highly visible failures occurred in a context in which the IFIs were already under attack. Non-governmental organizations (NGOs) had been growing in influence and were among the first to label the IFIs' structural adjustment policies a failure, while borrower countries and academic critics had grown more vocal in their criticisms of IMF and World Bank conditionality, particularly after the Asian crisis.[8] At the same time, internal institutional actors were also beginning to acknowledge the failure of certain aspects of their aid strategy, particularly the reliance on ever-growing numbers of conditions (World Bank 1998; Ahmed 2001).

This was a clear example of a contested failure. Although these actors disagreed about who was responsible for these failures, whether they blamed the neoliberal approach to aid, specific policy tools like conditionality or even the borrowers themselves, they

---

[7] A more detailed account of the dynamics in development governance in the 1990s can be found in: (Best 2014: Chs 4-8).

[8] One of the earliest and most influential critiques of structural adjustment was UNICEF's *Adjustment with a Human Face* (UNICEF 1987); the changing pressures after the Asian crisis are well-known and were raised by numerous members of staff and management in my interviews, including senior IMF (May 2007 and June 2010) and World Bank staff members (May 2007), and IMF Executive Directors (October 2006; May 2007; June 2010).

shared a conviction that these policies had failed in part because aid *success* was not defined correctly. More specifically, they agreed that policymakers had ignored the political and institutional dimensions of successful aid. The turn to "aid effectiveness" that followed was largely an effort to redefine aid success—or to cite the title of a World Bank report, to determine "What Works, What Doesn't and Why" (World Bank 1998). Influential studies by David Dollar, Jacob Svensson and Craig Burnside all argued that for aid to be effective, the recipient country had to have "sound" institutions in place (Burnside and Dollar 1997; World Bank 1998; Dollar and Svensson 1998). The policymakers who drew on these ideas began to focus more the longer-term institutional reform, rather than on short-term policy changes, and to emphasize the importance of borrowing countries taking some political leadership for reforms. [Although critical NGOs had sought more radical reforms, they too believed that greater country ownership and participation was crucial to better aid outcomes.]

Since policymakers concluded that they had defined aid success incorrectly, they sought to change their metrics for evaluating success and failure. This involved changing the explicit goals of aid, the techniques used to encourage and measure countries' compliance, and the processes for evaluating programs' success or failure. Whereas earlier aid programs had focused narrowly on fiscal, monetary and structural economic changes, the goals of the new policies were much broader: the new emphasis on institutional soundness and "good governance" gave them the green light to focus on borrowing countries' legal, political and administrative frameworks. As the policy goals changed, so did the techniques used to encourage and assess countries' compliance. Until the late 1990s, both the IMF and the World Bank relied heavily on quantitative conditions (IMF 1986; World Bank 2005). Such pass or fail conditions had worked reasonably well when the goals were narrowly economic. More complex policy objectives required more nuanced techniques for assessing compliance. In response to their critics, the IFIs sought to "streamline" conditions by reducing their number and focusing them on a narrower range of issues. As one senior former World Bank staff member put it, conditions were replaced with "a broader conversation about whether the core aims of the Bank in supporting a program, namely poverty reduction, are being met or not."[9]

Through their introduction of the Poverty Reductions Strategy Paper (PRSP) among other policy tools, the IMF and World Bank also sought to foster greater program ownership by requiring borrowing governments to consult with civil society. Yet, at the same time, informal conditions began to proliferate, including benchmarks, policy reviews and other mechanisms that were not defined as conditions by the IFIs but were seen as such by borrowers.[10] As policy goals became more complex, the instruments for encouraging and assessing compliance did too.

The processes used to evaluate the success rates of aid efforts also underwent important changes in the late 1990s and early 2000s. In the 1980s and early 1990s, program evaluation at the IMF was sporadic, while at the World Bank it was narrowly

---

[9] Interview with former senior World Bank staff member, Washington, DC, May 19, 2007.

[10] (World Bank 2005: 4-5,9; Gould 2005) The growing informality of conditions of course did not eliminate but merely transformed the power relations between lenders and borrowers. (Best 2014: Ch. 5)

focused on country compliance and the economic rate of return.[11] Over time, both organizations sought strategies for evaluating more complex policy goals. Rather than simply asking whether countries had implemented the reforms, they now tried to determine the degree of country ownership over particular programs, as well as their long-term sustainability (Boughton and Mourmouras 2002; OED 2005).

At first, in evaluating these more complex objectives, many agencies experimented with qualitative forms of evaluation: the World Bank began to rely on interview-based Participatory Rural Appraisals (Chambers 1997), for example, while the UK's Department for International Development experimented with more qualitative assessment techniques.[12] Yet, as I will discuss further below, by the early 2000s a rather different consensus metric for program evaluation had emerged: success and failure were to be assessed on the basis of the measurable results of program efforts.[13]

Results-oriented evaluations assume that it is possible to draw a causal line between policy actions and their outcomes (such as an effort by a donor to train midwives and a decline in infant mortality) in a context in which there is a multitude of external factors. Results-measurement thus seemed to hold out the promise of providing a kind of methodological fix to the challenges posed by the increasing complexity of development.

The contested failures of the 1990s produced a series of important changes in the ontological and epistemological assumptions embedded in development policy. Once aid agencies began to see political and social institutions as central to economic development, they were forced to redefine the boundaries around what counted as development—to open up this key black box and rethink their working ontology. This meant redrawing the frame around what counted in evaluating policy success and failure: political dynamics that had previously been treated as exogenous had to be brought into the frame.

Yet this move produced an epistemological dilemma: a world in which politics is central to development is also going to be more fluid and unpredictable and thus harder to evaluate than one that is narrowly economic. In response, policymakers developed a range of different evaluation techniques designed to measure these more complex objectives: the straightforward pass/fail conditions of the past had to give way to more complex evaluations of borrower compliance. Yet neither the IFIs nor national donors were comfortable with this growing complexity in the metrics of development: they wanted to be able to point to measurable signs of aid success in order to justify the dollars spent. Over time, the focus on measuring results was seen as the principal solution to these challenges.

If we are to understand the impact of the crises that confronted development organizations in the 1990s, we need to look at how they were narrated as particular kinds of failures: contested failures that unsettled existing definitions of success and failure,

---

[11] (World Bank Archives 1979, 1982)

[12] Interview with David Booth, Overseas Development Institute, London, 9 February 2010; interview with UK government staff member, September 2006.

[13] To list just a few examples: (DFID 2007; CIDA 2010; World Bank 2011). For a useful overview, see: (Pearson et al. 2010)

precipitating important shifts in development metrics and in the epistemological and ontological assumptions that underpin them.

The 2007-2008 Global Financial Crisis as a Failure

The 2008 financial crisis was not only perceived as a failure, but its magnitude was such that it provoked wide-ranging debates about how regulators and firms could have missed the signs of impending financial disaster. Initially at least, these debates were broad-based, including not only politicians, regulators and financial market actors but also the general public, some of whom found themselves unable to pay their mortgages or withdraw money from their banks, while others were incensed at the bailouts that they were being asked to pay for. This widespread interest in finance was unusual, given that the technical complexity of the sector generally insulates it from public pressure (Helleiner 2014b).

There was, of course, considerable disagreement about exactly who or what had failed: the main candidates for censure included the failure of regulation to keep up with financial innovations (Bernanke 2010; Rampell 2010; US GAO 2009), a culture of deregulation (UNCTAD 2009), government regulators and their policies of low interest rates and mortgage support (Alexander et al. 2009; Wallison 2009), financial institutions and credit rating agencies who drove the risk-taking culture (US Senate 2010; Sinclair 2010; Economist 2014), everyday investors and borrowers (Kotkin 2009; Wallison 2009), or the economic profession itself (Colander et al. 2009; Stiglitz 2009).

Behind all of this finger pointing was a shared sense of shock at the scale of the financial failure and a conviction that there should have been some warning signals. Clearly something was very wrong in how financial actors had evaluated the health of the system, if in just a few weeks the same investments, firms and regulations that had been deemed successes had become abject failures. In spite of divergent opinions on who was to blame for the failure of the financial system, there was considerable agreement that part of the problem lay in the technical calculations that underpinned it. In fact, as the crisis unfolded, there was a surprising amount of popular discussion of the nuts and bolts of how financial products are structured, measured and evaluated—an important clue that the crisis was becoming a contested failure.

One of the obvious sources of the crisis was the excessive level of risk that had been taken on by financial institutions, a practice that had been enabled by the lax regulatory standards of the time. Many of the discussions on policy reform thus focused on the mundane but suddenly crucial question of how financial risks were measured, and how those metrics might be revised in order to prevent future crises. These debates focused on two primary changes to risk metrics: shifting the emphasis from micro- to macro-level conceptions of risk, and changing the way that regulators managed financial institutions' risks by revising how their capital adequacy was calculated.

Although it was only after the crisis that it became common to distinguish between micro- and macro-prudential regulation, there is little doubt that the main way of thinking about and managing risk before the crisis was focused on its micro dynamics. There are many different forms of financial risk: credit risk, for example, involves the possibility of

a borrower defaulting on a debt; liquidity risk relates to the likelihood that an asset cannot be sold quickly enough to make a profit; market risk is the risk of losing money because of changes in the market value of an asset; while operational risk tries to capture the risks related to the failure of internal processes, practices and people. For regulators, a micro-prudential approach to managing risk involved not only treating different forms of risk as relatively discrete but also focusing on the risks faced by individual institutions, trying to ensure that each was financially sound, rather than assessing the risks in the system as a whole.

The advantage of this micro approach to risk is that each of these risks is far easier to measure and manage when they are treated separately. For firms, it becomes more difficulty to quantify, price and hedge against risks when they are interconnected. For regulators, it is easier to tell an individual bank to fix its balance sheet than to figure out how instability in one institution or sector of the economy is likely to affect the rest.[14]

International banking standards in place before the crisis also reinforced this micro-level approach to measuring risk by relying on banks' internal risk models and encouraging the use of Value at Risk (VaR) as a key metric. One of the most significant pieces of global regulation in place before the financial crisis was the Basel II accord, an international agreement finalized in 2004 that established global standards of capital adequacy for financial institutions (BCBS 2004). Banks have developed very high debt to equity ratios recent years to finance their investments, borrowing as much as thirty times their capital. Although high levels of leverage can produce enormous profits when things go well, they also make banks vulnerable to even a small decline in their assets.

Basel II sought to minimize such risks by setting out how much capital banks and other financial institutions had to maintain. Rather than relying on a straightforward leverage ratio—which would require a similar capital requirement for all of a bank's assets—regulators used a risk-weighted measurement, which required more capital for those assets deemed of higher risk. Basel II required banks to hold capital in order to offset potential losses from three different kinds of risk: credit, operational and market risk. In calculating credit risk, larger banks were encouraged to use their own internal models. Although regulators argued that this would allow for a more sophisticated and individualized evaluation of risk, this micro-based approach gave banks considerable room to understate their risks.

In its efforts to manage market risk, the Basel framework encouraged banks to make use of VaR, a metric that estimates a firm's risk of experiencing large losses in a particular time frame (BCBS 2004: 50-52, 120-123). VaR came under significant scrutiny during the crisis, as critics pointed out that the measure only predicted the losses that were likely to occur in normal conditions and had nothing to say about the potential for catastrophic losses (known as the problem of "tail risk") (Nocera 2009; 2014; O'Brien and Szerszen 2014). Financial actors' over-reliance on such measures meant that they did not recognize risks that they faced, even as they mounted, because their metrics literally did not allow them to see them.

[14] The disadvantage of this micro approach to risk, of course, as became quite apparent during the recent financial crisis, is that it doesn't work.

In the aftermath the crisis, policymakers began to search for new metrics capable of more accurately assessing the financial system's risks. Even though Basel II had only just come into effect as the financial crisis took off, policymakers quickly revised it. Basel III includes higher capital adequacy standards, increasing the amount of high quality capital that financial institutions are required to maintain. It also introduces a leverage ratio of 3% (allowing 33 dollars of debt for every dollar of capital) and provides a countercyclical buffer that enables regulators to increase capital requirements in an overheating economy (BCBS 2011). On the face of it, these changes appear substantial: the Tier I Common Ratio (one of the categories of capital) has been increased from 2% to 7%, and as much as 9.5% when the countercyclical buffer is added. Nonetheless, banks still retain the right to use their own internal models when calculating the risk weighting of their assets, so they are often able to avoid meeting these requirements (Wolf 2010; Admati and Hellwig 2013). The leverage ratio is designed to put a floor on such evasive efforts but, as I will discuss further in the conclusion, it remains quite low.

In addition to these quantitative changes in how regulators and banks measure financial risks, there have also been qualitative changes in the ways in which risk is conceptualized and measured, including a greater attention to its macro-dynamics. The crisis made it clear that the risks to the financial system as a whole were much greater than the sum of the risks to specific firms, sectors or countries: actions that reduced risk to an individual bank (like the sale of risky assets) could increase risks to the system as a whole if a large number of banks follow suit (producing an overall decline in the value of assets) (Persaud 2009: 2). During the crisis, financial actors argued that if regulators were going to be able to manage these kinds of risks, they would have to make them visible by redrawing the measurement frame to include macro-level, systemic risk in their calculations.[15] Advocates of systemic risk insisted on the need to conceptualize risk holistically—not as a series of individually calculable probabilities, but as a product of complex linkages that can cause unanticipated externalities, or overflows, to use Callon's language.

The failure of the global financial system was profound enough that it not only produced changes in the metrics that underpinned it but also lead to more fundamental ontological and epistemological shifts. Understanding what had failed during the crisis required not just new methods, but a new conception of the financial universe. To move beyond seeing risk as a micro-level phenomenon to conceptualizing it as a macro force requires a shift in policymakers' working ontology. Taking systemic risk seriously means seeing the global financial system as not only complex, dynamic, and highly interconnected, but also recognizing that because of this, it will inevitably be highly uncertain and prone to crisis.

This ontological shift, moreover, has epistemological implications. It is not easy to measure or predict the complex unknowns that make up this kind of financial world. The pursuit of metrics for systemic risk has in fact become something of a holy grail among regulators: a recent survey prepared for the United States' Office for Financial Research

[15] The concept of systemic risk became widespread enough to feature as the chief topic of a US Senate Hearing before the Committee on Homeland Security (US Senate 2009). Baker provides an excellent account of the key figures involved in moving this "macroprudential" agenda forward. (Baker 2013)

lists thirty-one different measures of systemic risk being developed around the world (Bisias et al. 2012).

What Has Changed: a Growing Preoccupation with Failure

The international economic system has encountered contested failures in the past. The stagflation crises of the 1970s were highly contested failures that ultimately produced a significant shift in economic metrics, as monetarist and new-classical inspired policies came to replace discredited Keynesian metrics (Helleiner 1994; Blyth 2002). The same thing happened in international development circles, as the poverty-targeting of the 1960s and 1970s, once championed by World Bank President Robert McNamara, was proclaimed a failure in the early 1980s and replaced with the a return to a "trickle-down" approach (Best 2013). Yet, in both cases, once the turmoil of contestation passed, these new metrics were embraced with confidence. That kind of confidence is far less evident this time around.

When an organization is aware of the limits of its knowledge and concerned about possible failures, it can respond either creatively or cautiously. Recognizing the limits of knowledge can be liberating—encouraging creative solutions that acknowledge their limits—or it can be constraining—producing cautious behaviour. Although we can see both trends in international development and finance, what is notable in both cases is how cautious and conservative policy responses have become, as policymakers have grown preoccupied with the problem of failure.

After prolonged debates about the sources of development failure, aid organizations were more aware of the uncertainty of the environment that they worked in and of the difficulties of finding a quick fix. In responding to these challenges, they have demonstrated both creativity and caution. In the early days of the aid effectiveness agenda, more ambitious and complex goals—such as fostering country ownership and improving the quality of governance—led to more creative and open-ended techniques: borrowing governments and civil society actors were given a bigger role in defining the programs' objectives, while conditions became less formal. Agency staff also experimented with more complex forms of evaluation with the aim of developing a more nuanced picture of policy successes and failures. A decade and a half later, however, these policies have lost momentum, as many organizational staff found the participatory techniques too cumbersome, and as borrowers and NGOs became disappointed with the thin forms of political engagement that the IFIs' conception of country ownership involved.[16]

At the same time as these more open-ended and experimentalist policies lost favour, results measurement has become ever-more prominent. By measuring the results of their policies, donor governments and international organizations hope to "prove" their effectiveness. This trend has become more pronounced in after the recent financial crisis, as policymakers search for some kind of certainty in an increasingly volatile world.

[16] Interviews with senior IMF staff members, June 14, 2010 and with a former World Bank Vice-President, May 9, 2007. Also: (Molenaers and Renard 2003; Piron and Evans 2004; Kovach 2005)

Results-measurement promises a way of knowing something certain in a world in flux, drawing a clear causal line between spending on school textbooks, for example, and measurable increases in test scores.[17]

Although results-based approaches appear confident, they instead reflect a cautious preoccupation with failure. In contrast with the ambition of early aid effectiveness initiatives, which recognized the complexity of development, results-based approaches are very narrow. In order to be able to demonstrate results, you have to be able to measure them; since subtler improvements in governance or social cohesion are difficult to quantify, results-based approaches gravitate towards simpler policy interventions that can be more easily counted. The results matrixes that staff members prepare also require them to identify risks, encouraging them to keep an eye out for potential failures.

In recent years, as budgets have been squeezed and more conservative governments have gained power in many donor states, results-based approaches have become increasingly cynical: aimed at meeting donors' needs to placate voters rather than at responding to concerns on the ground. A growing number of organizational staff are aware of the problematic nature of the causal claims they are making.[18] Many recognize just how difficult it is to attribute particular outcomes to the policy interventions of one donor in a context in which there is a multitude of influences. Results have also shown themselves to be infinitely malleable when political circumstances require, with few donors or recipients admitting to poor results.[19] Results-based measurement is thus becoming a classic example of provisional expertise as Luhmann described it, allowing donors to hedge against the possibility of failure even as they claim success.

Since the 2008 crisis, financial policymakers have also become preoccupied with the persistence of extreme uncertainty and the possibility of another major failure. In the immediate aftermath of the crisis, US Federal Reserve Chairman, Ben Bernanke (2010), pointed to "the profound uncertainty associated with the 'unknown unknowns' during the crisis" and argued for more sustained research into the behavior of economic actors in conditions of extreme uncertainty. Although in the two decades leading up to the crisis, central bankers had championed a policy approach based on simple, quantitative rules, even they have been forced to rethink monetary and financial policy in a highly uncertain economy. The Governor of the Bank of Canada, Stephen Poloz (2014), has recently pointed to the challenges that this more volatile environment poses for traditional policy tools:

---

[17] Education is one of the most popular areas for results-based approaches to aid, particularly the "cash on delivery" approaches that promise to pay by results. (DFID 2013)

[18] I was struck in my interviews by just how widespread this ambivalence has become. Interviews with World Bank staff, June 15 and 17, 2010, with senior IMF staff members, June 11 and 14, 2010, IMF Executive Board member, June 17, 2010, and with OECD-DAC staff members, May 2 and 4, 2011. For a widely-read critique from a former head of USAID, see: (Natsios 2010)

[19] For a critique of the Canadian government's fudging of its results data, see: (Brown and Sinclair 2011). For an analysis of the ways in which Mali, as a borrowing country, also tried to massage the data, see: (Bergamaschi 2011)

> The sort of post-crisis uncertainty that central banks are dealing with today is more profound than that which is typically subjected to rigorous analysis and does not lend itself easily to formal modelling.

As policymakers have sought to come to terms with this more unpredictable and failure-prone environment, they have adopted a more provisional style of governance. The most significant efforts to reform financial theory and practice have revolved around systemic risk, including efforts to measure systemic risk, on the one hand, and to manage them, on the other. While the first of these policy initiatives has tended to be experimental in character, the second has been conservative.

Because systemic risks are more difficult to measure and manage than micro-level risks, a considerable amount of institutional effort is being put behind the search for these new metrics.[20] The task, nonetheless, remains daunting. In fact, the authors of the US government survey of thirty-one different methods for measuring systemic risk suggest that because of the complexity of the phenomenon they are seeking to understand, there will never be a single best metric. Instead, they take an experimentalist approach, making the software for all thirty-one metrics available online in open-source format and encouraging different economists to test and expand on them. They go on to suggest "we remain agnostic at this stage about what is knowable" (Bisias et al. 2012: 4).

Some central bankers are also being more open about the limits of their own metrics. Poloz (2014) has indicated that the Bank of Canada will no longer engage in forward guidance, in which the Bank communicates its intentions over the longer-term, because it does not want to provide the markets with a false sense of security about the knowability of the economy. Instead, he suggests, Bank communications will provide market actors with a better picture of the uncertainties and risks that the Bank is struggling to understand and manage. Although the US Federal Reserve has not backed away from forward guidance, its current Chair, Janet Yellen, has also been more ambiguous in communicating a range of possible actions on interest rates, depending on the ultimate direction of the US economy (Fleming 2015).[21]

While efforts to measure and communicate systemic risk have thus been somewhat experimental, efforts to manage it remain conservative. One of the main mechanisms that policymakers have developed to contain systemic risks has been to identify domestically and globally systemically important financial institutions (SIFIs and G-SIFIs) (FSB 2011).

---

[20] The US Dodd-Frank Act created both the Office of Financial Research and the Financial Stability Oversight Council, which has the mandate to measure and manage system risk. The UK government created the Financial Policy Committee of Bank of England, while the European Union established the European Systemic Risk Board in 2010.

[21] In a *Financial Times* blog, Martin Sandbu notes that the ambiguity of the minutes from the July 2015 meeting of the Federal Reserve's interest-rate setting committee reveals "a recognition by the policymakers of how unusually hard it is to conduct monetary policy in today's conditions" (Sandbu 2015). IMF staff have also written a recent blog post noting that most monetary policy decisions are made "in the dark" because of the long lags between real-time estimates and the ultimate effects of key variables like "output gaps" (Grigoli et al. 2015).

These financial institutions, which are deemed large and interconnected enough to have damaging effects on the financial system if they run into trouble, must meet higher standards of information disclosure and capital adequacy. This policy is a clear example of a provisional approach to financial governance: it is explicitly focused on the risks of financial failure, or more specifically, the problem of institutions that are "too big to fail." The additional demands made of SIFI and G-SIFI banks are designed to reduce the chance of their failure. Yet the policy also recognizes that failures may still occur, seeking to limit their impact by asking financial institutions to provide a "living will" setting out how they would be wound down. Although this preoccupation with failure appears to have produced some significant innovations, when we dig a bit deeper, what is surprising is just how little has really changed. As Lucy Goodhart (2014) notes, citing a member of the US regulatory community, focusing on systemically important institutions is "'a micro-prudential solution to a macro-prudential problem'": it still assumes that it is possible to manage system-wide dynamics by focusing on individual institutions.

The Basel III provisions also remain cautious and retain a strong micro-level bias, particularly by relying on banks' internal risk models. As one *Economist* blog put it "Since it did not change this risk-weighting, Basel III effectively doubles down on Basel II"—increasing its bet that individual banks can effectively measure and manage their own risks (Economist 2010). Moreover, the Basel III leverage ratio of 3% remains surprisingly low, while only a few national jurisdictions have been more aggressive.[22] As Martin Wolf notes, "this amount of equity is far below levels markets would impose if investors did not continue to expect governments to bail out creditors in a crisis, as historical experience shows" (Wolf 2010). Like Martin Hellwig and others, Wolf suggests that a ratio well over ten percent, and possibly as high as the twenty or thirty percent that was once the norm, is necessary to protect the financial system (Hellwig 2010: 11; Admati and Hellwig 2013). In spite of the fact that policymakers appear open to the challenges posed by efforts to conceptualize and measure risk holistically, their efforts to translate these insights into policy remain conservative to date. Yet precisely because their efforts are all geared towards taking systemic risk more seriously, policymakers can say (to borrow Luhmann's words) that they have acted in a "risk-rational manner" (Luhmann 1998a: 72).

Conclusion

While the more traditional emphasis on crises provides useful insights into the strategic deployment of particular narratives of crisis, an attention to contested failures allows us to see how moments of crisis can provoke debates that are at once highly technical and quite profound, producing fundamental shifts in how policy success and failure are defined and managed.

Combining STS scholars' insights into the political character of metrics with the rich tradition of crisis analysis in IPE thus provides us with a whole new set of tools for understanding recent changes in the global governance of finance and development.

[22] Although the US has adopted its own higher ratios of 5 and 6%, the UK ultimately decided on a much lower ratio than expected (4.05%), while the EU has simply adopted the 3% ratio set out in Basel III.

Callon and Latour famously introduced the powerful concept of the black box to demonstrate how effective social actors can be in tidying away the unresolved questions and uncertainties that haunt even the most widely accepted of scientific claims. What is particularly interesting in these cases is how the black boxes containing financial and development expertise have been replaced by grey boxes, in which some of those residual unknowns remain just visible, like shadows seen through a translucent window.

The chief result of policymakers' continued awareness of such unknowns, I have argued, is a growing preoccupation with the possibility of failure. In both finance and development, policymakers have responded to this possibility in two very different ways: with a more open-ended and experimental approach and a more cautious and cynical one. In both cases, however, the experimental approach has been overwhelmed by more conservative strategies that allow policymakers to make only modest policy changes while also hedging against the possibility of future failures.

We can now make sense of the puzzle that I identified at the outset: how the recent financial crisis could generate both a significant shift in policymakers' epistemological frameworks and only status-quo policy changes. Recognizing the problem of failure need not lead to significant policy changes as long as policymakers follow Luhmann's lead and ensure that their expert claims retain an escape hatch in case of future failures. Bernanke himself (2010) beautifully articulated this strategy when he said, in defense of economists' failure to predict the crisis:

> Economic models are useful only in the context for which they are designed. Most of the time, including during recessions, serious financial instability is not an issue. The standard models were designed for these non-crisis periods. . .

In other words, the models work, until they don't.[23] A footnote that limits a model to certain normal conditions provides the necessary hedge against the possibility of failure. Moving beyond this kind of cautious approach to failure thus will require not just changing policymaking culture but also transforming the very discipline of economics—two changes that, in spite of their champions, do not appear to be in the cards at present (Colander 2009; Stiglitz 2009).

Ironically, this move towards more cautious policy responses may actually produce more failures in the longer-term: a development program that is limited to those initiatives whose outcomes can be quantified will avoid tackling many crucial challenges; and a system of financial regulation that continues to act as though systemic uncertainties are reducible to micro-level risks will be more prone failure. Yet such policies are safer from the policymakers' point of view, not only because they do not require politically difficult changes, but also because they can be justified in ways that can be measured. Ironically, while these policies may be fail-safe from an operational point of view, allowing experts to avoid responsibility if things go wrong, they are far from risk averse as far as outcomes are concerned—and in fact are very likely creating a world in which failure is not only always anticipated but also endlessly repeated.

---

[23] As Wes Widmaier pointed out when he brought this passage to my attention, Bernanke's argument bears a great deal of resemblance to Kramer's logic in the "levels" episode of Seinfeld.

Acknowledgements

This paper was first presented at a panel on "Governing failure, crisis and emergency" at the International Studies Association conference in New Orleans in March 2015. Thanks are due to my co-organizer for the panel, Paul Langley, as well as Tony Porter for his thoughtful comments as discussant. I would also like to thank the anonymous reviewers for their very helpful suggestions, Wes Widmaier for some electronic conversations that helped to spark the idea for this paper, and Marie Langevin for her superb research assistance. This paper was researched with the financial assistance of the Social Sciences and Humanities Research Council of Canada. It was written at the Academic Leadership Centre at the University of Ottawa in their Writing in Company of Others program.

References

(2010). Basel III: Third Time's the Charm? *The Economist*. September 13.

(2014). Credit Where Credit's Due: The Ratings Industry has Bounced Back from the Financial Crisis. *The Economist*. April 19.

Admati, Anat and Martin Hellwig (2013). *The Bankers' New Clothes*. Princeton: Princeton University Press.

Ahmed, Masood (2001). Refocusing IMF Conditionality. *Finance and Development* 38(4): 40-43.

Alexander, James, Michael Beenstock, Philip Booth and Eamonn Butler (2009). Government Failure Caused the Financial Crisis. Institute of Economic Affairs. *http://www.iea.org.uk/blog/government-failure-caused-the-financial-crisis* (Accessed October 12, 2014).

Baker, Andrew (2013). "The New Political Economy of the Macroprudential Ideational Shift." *New Political Economy* 18(1): 112-139.

Barnett, Michael and Martha Finnemore (1999). "The Politics, Power, and Pathologies of International Organizations." *International Organization* 53(4): 699-732.

Barry, Andrew (2002). "The Anti-political Economy." *Economy and Society* 31(2): 268-284.

Barry, Andrew (2012). "Political situations: knowledge controversies in transnational governance." *Critical Policy Studies* 6(3): 324-336.

BCBS (2004). *International Convergence of Capital Measurement and Capital Standards: A Revised Framework (Basel II)*. Basel, Switzerland, Basel Committtee on Banking Supervision, Bank for International Settlements.

BCBS (2011). *Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems*. Basel, Switzerland, Basel Committee on Banking Supervision, Bank for International Settlements.

Bergamaschi, Isaline. (2011). *Le Consensus Post-Washington au Mali: Trajectoires, Usages et Processus d'Appropriation*. Doctorat, Institut d'Études Politiques de Paris.

Bernanke, Ben. (2010). Causes of the Recent Financal and Economic Crisis: Speech before the Financial Crisis Inquiry Commission, Washington, DC, Board of Governors of the Federal Reserve System.

Bernanke, Ben. (2010). Implications of the Financial Crisis for Economics: Speech at the Bendheim Center for Finance, Princeton University, Board of Governors of the Federal Reserve System.

Best, Jacqueline. (2012). "Ambiguity and Uncertainty in International Organizations: A History of Debating IMF Conditionality." *International Studies Quarterly* 56(4): 674-688.

Best, Jacqueline. (2013). "Redefining Poverty as Risk and Vulnerability: Shifting Strategies of Liberal Economic Governance." *Third World Quarterly* 34(2): 109-129.

Best, Jacqueline. (2014). *Governing Failure: Provisional Expertise and the Transformation of Global Development Finance*. Cambridge, Cambridge University Press.

Bisias, Dimitrios, Mark Flood, Andrew Lo, and Stavros Valavanis (2012). "A Survey of Systemic Risk Analytics." *Office of Financial Research Working Paper* (0001).

Blankenburg, Stephanie and José Gabriel Palma (2009). "Introduction: the global financial crisis." *Cambridge Journal of Economics* 33(4): 531-538.

Blyth, Mark. (2002). *Great Transformations: The Rise and Decline of Embedded Liberalism*. Cambridge, Cambridge University Press.

Boughton, James. (2001). *Silent Revolution. The International Monetary Fund 1979–1989*. Washington, D.C., International Monetary Fund.

Boughton, James and Alexandrow Mourmouras (2002). "Is Policy Ownership an Operational Concept?" *IMF Working Paper* WP/02/72.

Brassett, James and Chris Clarke (2012). "Performing the Sub-Prime Crisis: Trauma and the Financial Event." *International Political Sociology* 6(1): 4-20.

Brassett, James, Lena Rethel and Matthew Watson (2010). "The Political Economy of the Subprime Crisis: The Economics, Politics and Ethics of Response." *New Political Economy* 15(1): 1-7.

Brown, Stephen and John Sinclair (2011). Transparency and Canadian Foreign Aid. *Huffington Post.* August 1.

Burnside, Craig and David Dollar (1997). "Aid, Policies and Growth." *World Bank Policy Research Working Paper*(1777).

Callon, Michel. (1998). An Essay on Framing and Overflowing: Economic Externalities Revisited by Sociology. *The Laws of the Markets*. M. Callon. Oxford, Blackwell**:** 244-269.

Callon, Michel. (1998). Introduction: the Embeddness of Economic Markets in Economics. *The Laws of the Markets*. M. Callon. Oxford, Blackwell**:** 1-57.

Callon, Michel., Ed. (1998). *The Laws of the Markets*. Oxford, Blackwell.

Chambers, Robert. (1997). *Whose Reality Counts? Putting the First Last*. London, ITDG Publishing.

CIDA (2010). *Development for Results 2010-2011: At the Heart of Canada's Efforts for a Better World*, Canadian International Development Agency.

Colander, D., et al (2009). "The Financial Crisis and the Systemic Failure of Academic Economics." *Kiel Working Papers* (1489).

Datz, Giselle. (2013). "The Narrative of Complexity in the Crisis of Finance: Epistemological Challenge and Macroprudential Policy Response." *New Political Economy* 18(4): 459-479.

De Búrca, Gráinne, Robert Keohane and Charles Sabel. (2014). "Global Experimentalist Governance." *British Journal of Political Science* 44(3): 477-486.

DFID (2007). Results Action Plan. London, Department for International Development.

DFID (2013). DFID's Approach to Payment by Results. London, UK Department for International Development.

Dollar, David and Jakob Svensson (1998). What Explains the Success or Failure of Structural Adjustment Programs? *World Bank Development Research Group*. Washington, DC, World Bank.

Feldstein, Martin (1998). "Refocusing the IMF." *Foreign Affairs*(March/April): 20-33.

Fleming, Sam. (2015). Janet Yellen defends US central bank independence. *Financial Times*. London.

FSB (2011). Policy Measures to Address Systemically Important Financial Institutions. Basel, Financial Stability Board.

GAO (2009). Financial Regulation : A Framework for Crafting and Assessing Proposals to Modernize the Outdated U.S. Financial Regulatory System, United States Government Accountability Office.

Goodhart, Lucy. (2014). "Brave New World? Macro-prudential policy and the new political economy of the federal reserve." *Review of International Political Economy*.

Gould, Jeremy, Ed. (2005). *The New Conditionality: The Politics of Poverty Reduction Strategies*. London, Zed Books.

Grigoli, Francesco, Alexander Herman, Andrew Swiston, and Gabriel Di Bella. (2015). "Making Monetary Policy Decisions in the Dark." *IMFdirect* http://blog-imfdirect.imf.org/2015/08/12/making-monetary-policy-decisions-in-the-dark/ Accessed August 26, 2015.

Hay, Colin. (1996). "Narrating Crisis: The Discursive Construction of the 'Winter of Discontent'." *Sociology* 30(2): 253-277.

Hay, Colin. (2013). "Treating the Symptom Not the Condition: Crisis Definition, Deficit Reduction and the Search for a New British Growth Model." *British Journal of Politics & International Relations* 15(1): 23-37.

Helleiner, Eric. (1994). *States and the Reemergence of Global Finance: From Bretton Woods to the 1990s*. Ithaca, NY, Cornell University Press.

Helleiner, Eric. (2014a). Out from the shadows: governing OTC derivatives after the 2007–08 financial crisis. *The Return of the Public in Global Governance*. J. Best and A. Gheciu. Cambridge, Cambridge University Press.

Helleiner, Eric. (2014b). *The Status Quo Crisis : Global Financial Governance after the 2007-08 Financial Meltdown*. Oxford ; New York, Oxford University Press.

Hellwig, Martin. (2010). "Capital Regulation after the Crisis: Business as Usual?" *Preprints of the Max Planck Institute for Research on Collective Goods*.

IMF Archives (1986). EBS/86/211. Program Design and Performance Criteria. Washington, D.C., International Monetary Fund.

Kotkin, Joel. (2009). Did Homeowners Cause The Great Recession? *Forbes*. June 30.

Kovach, Hetty. (2005). A new devolpment agenda? Sierra Leone's First PRSP, Eurodad.

Latour, Bruno. (1987). *Science In Action*. Milton Keynes, Open University.

Latour, Bruno. (1999). *Pandora's Hope: Essays on the Reality of Science Studies*. Cambridge, Mass, Harvard University Press.

Luhmann, Niklas. (1998a). Describing the Future. *Observations on Modernity*. Stanford, Stanford University Press: 63-74.

Luhmann, Niklas. (1998b). The Ecology of Ignorance. *Observations on Modernity*. Stanford, Stanford University Press**:** 75-111.

MacKenzie, Donald. (2006). *An Engine, Not a Camera: How Financial Models Shape Markets*. Cambridge, Mass, MIT Press.

MacKenzie, Donald. (2009). "Making things the same: Gases, emission rights and the politics of carbon markets." *Accounting Organizations and Society* 34(3-4): 440-455.

Molenaers, Nadia and Robrecht Renard (2003). "The World Bank, Participation and PRSP: The Bolivian Case Revisited." *The European Journal of Development Research* 15(2): 133-161.

Moschella, Manuelle and Eleni Tsingou, Eds. (2013). *Great Expectations, Slow Transformations: Incremental Change in Post-Crisis Regulation*. Colchester, ECPR Press.

Natsios, Andrew. (2010). The Clash of the Counter-bureaucracy and Development. Washington, DC, Center for Global Development.

Nocera, Joe. (2009). Risk management: what let to the financial meltdown. *The New York Times*. January 2.

O'Brien, James and Pawel Szerszen (2014). "An Evaluation of Bank VaR Measures for Market Risk During and Before the Financial Crisis." *Finance and Economics Discussion Series, Federal Reserve Board* (2014-21).

OECD (2005). Paris Declaration on Aid Effectiveness. Paris, Organisation for Cooperation and Development.

OED (1994). Annual Review of Evaluation Results 1993. Operation Evaluation Department, Washington, DC, World Bank.

OED (2005). 2004 Annual Review of Development Effectiveness: The World Bank's Contributions to Poverty Reduction. Washington, DC, World Bank.

Pearson, Mark, Martin Johnson, and Robin Ellison. (2010). Review of Major Results Based Aid (RBA) and Results Based Financing (RBF) Schemes: Final Report. London, Department for International Development.

Persaud, Avinash. (2009). "Macroprudential Regulation: Fixing Fundamental Market and Regulatory Failures." *Crisis Response, World Bank*.

Piron, Laure-Helene, and Alison Evans (2004). Politics and the PRSP Approach: Synthesis Paper. London, Overseas Development Institute.

Poloz, Stephen. (2014). "Integrating Uncertainty and Monetary Policy-Making: A Practitioner's Perspective." *Bank of Canada Discussion Paper* 2014(6).

Power, Michael. (1996). "Making Things Auditable." *Accounting, Organizations and Society* 21(2/3): 289-315.

Ramesh, M. (2009). "Economic Crisis and its Social Impacts: Lessons from the 1997 Asian Economic Crisis." *Global Social Policy* 9: 79-99.

Rampell, Catherine. (2010). Lax Oversight Caused Crisis, Bernanke Says. *The New York Times*. January 4.

Roitman, Janet. (2013) *Anti-crisis*. Chapel Hill: Duke University Press.

Sabel, Charles and Jonathan Zeitlin (2008). "Learning from difference: The new architecture of experimentalist governance in the EU." *European Law Journal* 14(3): 271-327.

Sandbu, Martin. (2015). "Will They or Won't They?" *Free Lunch - Financial Times* August 20. http://www.ft.com/intl/martin-sandbu-free-lunch. Accessed August 21, 2015.

Sinclair, Timothy. (2010). "Round Up the Usual Suspects: Blame and the Subprime Crisis." *New Political Economy* 15(1): 91-107.

Stiglitz, Joseph. (1998). Restoring the Asian Miracle. *The Wall Street Journal*.

Stiglitz, Joseph. (2009). "The Anatomy of a Murder: Who Killed America's Economy?" *Critical Review* 21(2-3): 329-339.

UNCTAD (2009). The Global Economic Crisis: Systemic Failures and Multilateral Remedies. Geneva, United National Conference on Trade and Development.

UNDP (1999). Human Development Report 1999. New York, United Nations Development Programme.

UNICEF (1987). Adjustment with a Human Frace. New York, United Nations.

US Senate (2009). Where Were the Watchdogs? Systemic Risk and the Breakdown of Financial Governance. *United States Senate Committee on Homeland Security and Government Affairs*. Washington, DC.

US Senate (2010). Wall Street and the Financial Crisis: The Role of Credit Rating Agencies. *United States Senate Committee for Homeland Security and Governmental Affairs, Permanent Subcommittee on Investigations*. Washington, DC.

Wallison, Peter. (2009). "Cause and Effect: Government Policies and the Financial Crisis." *Critical Review* 21(2-3): 365-376.

Widmaier, Wesley, Mark Blyth and Leonard Seabrooke (2007). "Exogenous Shocks or Endogenous Constructions? The Meanings of Wars and Crises." *International Studies Quarterly* 51(4): 747-759.

Wolf, Martin. (2010). Basel: the mouse that did not roar. *Financial Times*. September 14.

World Bank. (1998). *Assessing Aid: What Works, What Doesn't, and Why*. New York, Oxford University Press.

World Bank. (2005). Review of World Bank Conditionality. Washington, DC.

World Bank. (2011). World Bank for Results. Washington, DC.

World Bank Archives (1979) OED. Doc. 26815. World Bank Standards and Procedures. Washington, DC, World Bank Operations Evaluation Department.

World Bank Archives (1982) OED. Doc. 4101. Eighth Annual Review of Project Performance Audit Results. Washington, DC, World Bank Operations Evaluation Department.