The “Other” Imbalance and the Financial Crisis

Ricardo J. Caballero1
December 1, 2009

Abstract

One of the main economic villains before the crisis was the presence of large “global imbalances.” The concern was that the U.S. would experience a sudden stop of capital flows, which unavoidably would drag the world economy into a deep recession. However, when the crisis finally did come, the mechanism did not at all resemble the feared sudden stop. Quite the opposite, during the crisis net capital inflows to the U.S. were a stabilizing rather than a destabilizing source. I argue instead that the root imbalance was of a different kind: The entire world had an insatiable demand for safe debt instruments that put an enormous pressure on the U.S. financial system and its incentives. The world economy experienced a massive safe-assets imbalance, which is not likely to go away any time soon and, in fact, it has been exacerbated by the crisis. This general equilibrium feature is missing from much of the policy reform debate, which often points in directions that could exacerbate the shortage problem by limiting the financial sectors’ ability to engineer safe assets. Instead, I argue for private-public solutions that would preserve the good parts of the securitization industry while removing the system risk from the banks’ balance sheets.

Keywords: Global imbalances, financial crisis, safe assets shortage, securitization, systemic fragility, panic, complexity, Knightian uncertainty, contingent insurance

JEL Codes: E32, E44, E58, F30, G01, G20

---

1 MIT and NBER. Prepared for the Paolo Baffi Lecture to be delivered at the Bank of Italy on December 10th, 2009. I thank Francesco Giavazzi and Arvind Krishnamurthy for their comments, and Fernando Duarte, Jonathan Goldberg, and Matthew Huang for outstanding research assistance. First draft: November 24, 2009.
I. Introduction

One of the main economic villains before the crisis was the presence of large “global imbalances,” which refer to the massive and persistent current account deficits experienced by the U.S. and financed by the periphery. The IMF, then in a desperate search for a new mandate that would justify its existence, had singled out these imbalances as a paramount risk for the global economy. That concern was shared by many around the world and intellectually grounded on the devastating crises often experienced by emerging market economies that run chronic current account deficits. The main trigger of these crises is the abrupt macroeconomic adjustment needed to deal with a sudden reversal in the net capital inflows that supported the previous expansion and current account deficits (the so called “sudden stops”). The global concern was that the U.S. would experience a similar fate, which unavoidably would drag the world economy into a deep recession.

However, when the crisis finally did come, the mechanism did not at all resemble the feared sudden stop. Quite the opposite occurred: During the crisis net capital inflows to the U.S. were a stabilizing rather than a destabilizing source. The U.S. as a whole never experienced, not even remotely, an external funding problem. This is an important observation to keep in mind as it hints that it is not the global imbalances per-se, or at least not through their conventional mechanism, that should be our primary concern.

I argue instead that the root imbalance was of a different kind: The entire world, including foreign central banks and investors, but also many U.S. financial institutions, had an insatiable demand for safe debt instruments which put enormous pressure on the U.S. financial system and its incentives (Caballero and Krishnamurthy 2008). The world economy experienced a massive safe-assets imbalance, which is not likely to go away any time soon.

In this view, the surge of safe-assets-demand is a key factor behind the rise in leverage and macroeconomic risk concentration in financial institutions in the U.S. (as well as the
U.K., Germany, and a few other developed economies), as these institutions sought the profits generated from bridging the gap between this rise in demand and the expansion of its *natural supply* (more on this later on). In all likelihood, the safe-asset shortage is also a central force behind the creation of highly complex financial instruments and linkages, which ultimately exposed the economy to panics triggered by Knightian uncertainty (Caballero and Krishnamurthy 2007, Caballero and Simsek 2009a,b).

This is not to say that the often emphasized regulatory and corporate governance weaknesses, misguided homeownership policies, and unscrupulous lenders, played no role in creating the conditions for the surge in real estate prices and its eventual crash. However, it is to say that these were mainly important in determining the minimum resistance path for the safe-assets imbalance to release its energy, rather than being the root cause of the dramatic recent macroeconomic boom-bust cycle.  

Similarly, it is not to say that global imbalances did not play a role. Indeed, there is a connection between the safe-assets imbalance and the more visible global imbalances: The latter were caused by the funding countries’ demand for financial assets in excess of their ability to produce them (Caballero et al 2008a,b), but this gap is particularly acute for safe assets since emerging markets have very limited institutional capability to produce these assets. Thus, the excess demand for safe-assets from the periphery greatly added to the U.S. economy’s own imbalance caused by a variety of collateral, regulatory, and mandated requirements for banks, mutual funds, insurance companies, and others. This safe-asset excess demand was exacerbated by the NASDAQ crash

---

2 Acharya and Schnabl (2009) document that it is precisely those developed economies where banks could exploit regulatory arbitrage through conduits that issued the ABCP sought by money market funds around the world, and that experienced the largest stock-market declines during the crisis. Moreover, they show that there is no relation between countries’ issuance of ABCP and their current account deficits.

3 See Krishnamurthy and Vissing-Jorgensen (2007) for persuasive evidence that safe-asset demand is quantitatively very significant and that it affects macro-prices and quantities. See Gorton and Metrick (2009) for a lucid description of some of the special features of safe debt (repos, in particular) that justify its high demand (beyond risk aversion).
which re-alerted the rest of the world of the risks inherent to the equity market even in developed economies.

The point, however, is that the gap to focus on is not along the external dimension we are so accustomed to, but along the safe-asset dimension. Shifting the focus to the latter provides a parsimonious account of many of the main events prior to, as well as during, the onset of the crisis — something the global (current account) imbalances view alone is unable to do.

Within this perspective the main pre-crisis mechanism worked as follows: As the demand for safe assets began to rise above what the U.S. corporate world and safe-mortgage-borrowers naturally could provide, financial institutions began to search for mechanisms to generate triple-A assets from previously untapped and riskier sources. Subprime borrowers were next in line, but in order to produce safe assets from their loans, “banks” had to create complex instruments and conduits that relied on the law of large numbers and tranching of their liabilities. Similar instruments were created from securitization of all sorts of payment streams, ranging from auto to student loans (see Gorton and Souleles 2006). Along the way, and reflecting the value associated to creating financial instruments from them, the price of real estate and other assets in short supply rose sharply. A positive feedback loop was created, as the rapid appreciation of the underlying assets seemed to justify a large triple-A tranche for derivative CDOs and related products. Credit rating agencies contributed to this loop, and so did greed and misguided homeownership policies, but they were not the root cause.

From a systemic point of view, this new found source of triple-A assets was much riskier than the traditional single-name highly rated bond. As Coval et al (2009) demonstrate, for a given unconditional probability of default, a highly rated tranche made of lower quality underlying assets will tend to default, in fact it can (nearly) only default, during a systemic event. This means that, even if correctly rated as triple-A, the correlation
between these complex assets distress and systemic distress is much higher than for simpler single-name bonds of equivalent rating.

The systemic fragility of these instruments became a source of systemic risk in itself once a significant share of them was kept within the financial system rather than sold to final investors. Banks and their SPVs, attracted by the low capital requirement provided by the senior and super-senior tranches of structured products, kept them in their books, sometimes passing their (perceived) infinitesimal risk onto the monolines and insurance companies (AIG, in particular). The recipe was copied in the main European financial centers (Acharya and Schnabl 2009). Through this process, the core of the financial system became interconnected in increasingly complex ways and as such, it developed vulnerability to a systemic event.

Much of the crisis is blamed on the crash of the real estate “bubble” and the rise in subprime mortgage defaults that followed it. But this cannot be all, or even much, of it. The global financial system went into cardiac arrest mode and was on the verge of imploding more than once, which seems hard to attribute to a relatively small shock such as the real-estate/subprime combo. Instead, the real damage came from the unexpected and sudden freezing of the entire securitization industry. Almost instantaneously, confidence vanished and the complexity which made possible the “multiplication of bread” during the boom, turned into a source of counterparty risk, real and imaginary. Senior and super-senior tranches were no longer perceived as invulnerable, and making matters worse, banks had to bring back into their balance sheets more of this new risk from the now struggling SIVs and conduits (see Gorton 2008). Knightian uncertainty took over, and pervasive flights to quality plagued the financial system. Fear fed into more fear, causing reluctance to engage in financial transactions, even among the prime financial institutions.

Along the way, the underlying structural deficit of safe assets that was behind the whole cycle worsened as the newly found source of triple-A assets from the securitization
industry dried up, and the spike in perceived uncertainty further increased demand for these assets. Safe interest rates plummeted to record low levels. Initially, the flight to quality was a boon for money market funds, which suddenly found themselves facing a herd of new clients. In order to capture a large share of this expansion in demand from clients with a higher risk-tolerance than their usual clients, some money market funds began to invest in short-term commercial paper issued by the investment banks in distress. This strategy backfired after Lehman’s collapse, when the Reserve Primary Fund “broke-the-buck” as a result of its losses associated with Lehman’s bankruptcy. Perceived complexity reached a new level as even the supposedly safest private funds were no longer immune to contagion. Widespread panic took over and were it not for the massive and concerted intervention taken by governments around the world, the financial system would have imploded.

Global imbalances and their feared sudden reversal never played a significant role during this deep crisis. In fact, the worse things became, the more domestic and foreign investors ran to U.S. Treasuries for cover and the dollar appreciated. Instead, the largest reallocation of funds matched the downgrade in perception of the safety of the newly created triple-A securitization based assets.4

From this perspective, the core policy problem to deal with is how to bridge the safe-asset gap without over-exposing the financial sector to systemic risk. Raising capital requirements is a kneejerk policy reaction to reduce vulnerability but it does not help to deal with the structural problem of excess safe-asset demand. Quite the opposite, by reducing the financial sector’s ability to grow its balance sheet, it may worsen the safe-asset gap. The cost of this policy distortion is stronger headwinds for the recovery and the risk that the same pattern of systemically-vulnerable safe-asset creation may

---

4 Adelino (2009) documents that while the issuance prices of lower rated RMBS reflected their relative exposure to a crisis even controlling for rating, this was not the case for triple-A securities. Investors in these securities seemed to have been less informed about the quality of the securitized assets than investors in riskier, more information sensitive, securities.
migrate to somewhere else in the world that is even less prepared to absorb the systemic risk.

There are essentially two options to deal with the core problem of triple-A shortage: Either the government takes care of supplying much of the triple-A assets or it lets the private sector take the lead role with government support only during extreme systemic events. Here I develop the latter option. The reason this is an option is that the main failure during the crisis was not in the private sector’s ability to create triple-A assets through complex financial engineering, but in the systemic vulnerability created by this process. We should preserve the good aspects of this process while finding a mechanism to relocate the systemic risk component generated by this asset creation activity away from the banks and into private investors (for small and medium size shocks) and the government (for tail events). This transfer should be done on an ex-ante basis and for a fair fee, which can incorporate any concerns with the size, complexity, and systemic exposure of specific financial institutions. There are many options to do so, all of which amount to some form of partially mandated insurance provision from the government to the financial sector against a systemic event.

The rest of the paper goes into a more detailed analysis of the three components highlighted in this introduction: The prelude of the crisis, the crisis, and the appropriate policy framework in a safe-asset shortage environment.

II. The Prelude

The most visible anomaly in international financial markets prior to the crisis was the large and sustained current account deficit of the U.S. Paradoxically, perhaps its biggest danger was that it distracted attention from a more serious and critical imbalance, that between the global demand for safe assets and the ability of the U.S. private sector to
generate these assets without over-stretching its financial system. In this section I develop this argument in more detail.

II.A. Global Imbalances

Prior to the crisis, there was a widespread concern with “global imbalances,” which essentially refer to the massive and persistent current account deficits experienced by the U.S. and financed by the periphery (see Figure 1).

Figure 1.

![Current Account Graph](image)


In Caballero et al (2008a) we argued that the emerging market crises at the end of the 1990s, the subsequent rapid growth of China and other East Asian economies, and the associated rise in commodity prices in recent years reoriented capital flows from
emerging markets toward the United States. In effect, emerging markets and commodity producers in need of sound and liquid financial instruments to store their newfound wealth turned to the U.S. financial markets and institutions, which were perceived as uniquely positioned as providers of these instruments.

Concern about these “imbalances” was intellectually grounded on the devastating crises often experienced by emerging market economies that run chronic current account deficits. The main trigger of these crises is the abrupt macroeconomic adjustment needed to deal with a sudden reversal in the net capital inflows that supported the previous expansion and current account deficits (the so called “sudden stops”).

Figure 2 from Calvo and Talvi (2005) and Calvo, Izquierdo and Talvi (2006) captures the dramatic events during a sudden stop such as the one that affected South East Asia at the end of the 1990s. In a matter of months, net capital flows as a share of GDP declined by double-digits, with a corresponding adjustment in the current account and aggregate demand. The exchange rate collapsed, and along with it so did national income.

The global concern by the mid-2000s was that the U.S. would experience a similar fate, which unavoidably would drag the world economy into a deep recession. However, when the crisis finally did come, the mechanism did not at all resemble the feared sudden stop. Quite the opposite occurred, during the crisis net capital inflows to the U.S. were a stabilizing rather than a destabilizing source. The U.S. as a whole never experienced, not even remotely, an external funding problem. Moreover, during the

---

5 In Caballero and Krishnamurthy (2007) we described the emerging markets side of the story. We argued there that the financial underdevelopment of these economies naturally led to the formation of domestic asset bubbles during periods of large net capital inflows. The crash of these bubbles lead to large capital flow reversals. The events of the late 1990s corresponded to a coordinated crash in emerging markets and hence a reallocation of funds toward the U.S. large enough to lead to bubbles in the U.S. and other developed economies (Caballero et al 2008b).
early phase of the crisis, Sovereign Wealth Funds were the primary providers of fresh capital to distressed U.S. banks.\(^6\)

Figure 2.

![Sudden Stop in South East Asia*](image)

Source: Calvo and Talvi (2005) and Calvo, Izquierdo and Talvi (2006).

Figure 3 shows that the U.S. current account deficit was already experiencing a turn-around since 2006, and that it indeed dropped sharply after the Lehman episode. But the key contrast with emerging markets experiencing large external adjustments is that the U.S. dollar appreciated sharply during this episode. That is, the adjustment was the result of a contraction in aggregate demand due to domestic financial problems rather than due to a sudden shortage of net external funding. This is an important observation

\(^6\) They did so until Secretary Paulson decided to exemplary punish shareholders during the Bear Stearns intervention.
to keep in mind as it hints that it is not the global imbalances per-se, or at least not through their conventional mechanism, that should be our primary concern.

Figure 3.

Source: IMF International Financial Statistics

II.B. **The Critical Safe Assets Imbalance**

Since the 1980s, there has been a surge in financial assets to store value (in addition to those for risk “diversification”). Figure 4 shows that U.S. financial assets grew from less than 160 percent of GDP in 1980 to almost 480 percent in the third quarter of 2007. Debt plays an important role in this surge, especially in the 1980s and in the post September 11, 2001 and Nasdaq crash period.
The growth in the ratio of debt to GDP was due to an increase in debt generated by the private sector, especially the financial sector. The domestic financial sectors’ share of U.S. debt outstanding rose from 12 percent in 1980Q1 to 34 percent in 2007Q3 (Figure 5). At the same time, the government share of outstanding U.S. debt fell from 24 percent to 15 percent. (Unsurprisingly, this trend has been reversed of late.) Households and non-financial firms, which are the residual, were each responsible for a relatively flat share of the growing debt-outstanding pie.

A fundamental driver of this increase in debt liabilities was the insatiable demand for safe debt instruments. This demand came from foreign central banks and investors, but also from many U.S. financial institutions. The demand for safe debt instruments could not be readily met by existing sources of triple-A debt. Only a sliver of corporate debt, for example, carries a triple-A rating. Thus, the world economy experienced a massive

---

7 Debt is the sum of debt outstanding for households, businesses (non-financial), state and local governments, the federal government and domestic financial sectors. Equity is the sum of market capitalizations of the NYSE, Nasdaq and Amex exchanges.
safe-assets imbalance, which is not likely to go away any time soon (Caballero and Krishnamurthy 2009).

Figure 5.

![Government and financial sector debt](image)

Source: Federal Reserve.  

This imbalance presented a large profit opportunity for the U.S. financial system. However, creating safe assets to meet this demand put enormous pressure on the U.S. financial system and its incentives. The U.S. financial system created safe assets from unsafe ones by pooling assets and issuing senior claims on the payoffs of the pools. The senior claims are protected against losses because more junior claims absorb losses before the senior claims. The claims on these pools of assets are called collateralized debt obligations, or CDOs.

---

8 Government includes federal, state, and local government. The denominator is total debt outstanding by households, businesses (non-financial), federal, state and local government, and the domestic financial sector.
Global issuance of CDOs grew from $185 billion in 2000 to $1.3 trillion in 2007.\(^9\) A large share of these assets carried triple-A ratings, at least before the crisis. As of June 2007, Fitch, a ratings agency, gave a triple-A rating to almost 60 percent of all global structured products, according to a Fitch Ratings (2007) document cited in Coval, Jurek and Stafford (2009).\(^10\) Benmelech and Dlugosz (2009) find that, in a large sample of collateralized loan obligations (CLOs) issued between 2000 and 2007, 71 percent of issuance is rated triple-A, a percentage that is fairly stable throughout the period.

In contrast, Fitch (in June 2007) and S&P (in early 2008) gave a triple-A rating to less than one percent of corporate issuers. Moreover, the increase in issuance of corporate debt was slower than the increase in the issuance of CDOs.

Figure 6 compares the volume of U.S. issuance of corporate debt and CDOs. Corporate debt issuance grew between 2000 and 2006, but not nearly as fast as did CDO issuance. In 2000, the volume of CDO issuance was only 18 percent of the volume of corporate debt issuance; by 2006, this ratio had risen almost four-fold.\(^11\)

\(^9\) IMF Global Financial Stability report (2009), Figure 2.2. These numbers include both CDOs and CDO\(^2\).


\(^11\) Due to the crisis, issuance of CDOs collapsed to less than 24 billion dollars during the first five months of 2009, while corporations rushed to secure financing in the newly thawed debt markets, issuing 465 billion dollars in the corporate debt market in the same period. Thus, for the first five months of 2009, CDO issuance volume as a percentage of corporate debt issuance volume is only five percent.
Figure 6.

A shift toward CDOs


Corroborating the demand-pull for these instruments, as the quantity of non-government triple-A assets expanded, the yields required to hold them tightened. In January 2002, the spread between Moody’s triple-A seasoned corporate bond rate and a 30-year U.S. Treasury was 1.19 percentage points.¹³ In early 2007, before the crisis began, the spread reached 0.55 percentage points. Figure 7 shows the spread for Moody’s triple-A seasoned corporate bond rate relative to a 20-year Treasury bond.

¹² U.S. corporate bond issuance includes “all non-convertible debt, MTNs and Yankee bonds, but excludes CDs and federal agency debt.” Both investment-grade and high-yield issues are included. Corporate debt issuance volumes are from the Securities Industry and Financial Markets Association (SIFMA) and based on Thomson Reuters data. CDO issuance includes the issuance of CDOs and CDO^2, from the IMF Global Financial Stability report (2009). The 2009 data covers through the end of June.

¹³ The spread between Moody’s triple-A rate and a 20-year Treasury was 0.9 percentage points in January 2002. In early 2007, it reached 0.45 percentage points.
Coval, Jurek and Stafford (2008) find that CDX tranches – essentially, tranches of a CDO with credit default swaps (CDS) as assets – carried similar yields to single-name CDS with similar loss rates (“despite their highly dissimilar economic risks”). Their period of analysis is September 2004 to September 2007. This suggests that the tight pricing of triple-A corporate bonds carried over into the structured-finance market.

The creation of CDOs and the leveraging in financial institutions helped meet the global demand for safe assets. However, these securities and the linkages among firms became very complex, leaving the system vulnerable to panic (Caballero and Simsek 2009a,b). Also, senior CDO tranches concentrate macroeconomic risk, and banks

---

14 This figure plots the spread between Moody’s triple-A rate and a 20-year constant maturity Treasury. There is a small maturity mismatch because “Moody's tries to include bonds with remaining maturities as close as possible to 30 years.” However, there is no 30-year constant maturity Treasury bond for a good part of the 2000s, since the Treasury stopped issuing 30-year debt for a period. Thus, the spread shown in the figure is a noisy measure of the maturity-matched spread.
retained a large number of these assets. Thus, the response of the financial system to the safe-assets imbalance produced conditions in which an episode of Knightian uncertainty could do significant damage (Caballero and Krishnamurthy 2007).

Having downplayed the conventional concern with global imbalances, it is important to note that there is a connection between the safe-assets imbalance and the more visible global imbalances: The latter were caused by the funding countries’ demand for financial assets in excess of these countries’ ability to produce them (Caballero et al 2008), but this gap is particularly acute for safe assets since emerging markets have very limited institutional capability to produce them. Thus, the excess demand for safe-assets from the periphery greatly added to the U.S. economy’s own imbalance caused by a variety of collateral requirements and mandates for mutual funds, insurance companies, and others. The point, however, is that the gap to focus on is not along the external dimension we are so accustomed to, but along the safe-asset dimension.

Relative to the liabilities of the rest of the world owned by the U.S., the liabilities of the U.S. owned by the rest of the world are skewed toward debt, rather than equity and FDI. This is shown in Figure 8, which gives the ratio of debt to equity and FDI for U.S. assets and liabilities, according to careful estimates by Gourinchas and Rey (2007). The figure shows that after the emerging markets crisis of the late 1990s, there was a sharp decline in the ratio of U.S. debt liabilities to U.S. equity and FDI liabilities as foreign investors demanded U.S. assets across the board and particularly high return risky assets. However, this trend turned around sharply after the dot-com crash and the September 11, 2001 attacks on the U.S. This is partially a price effect; but it also reflects a shift in allocation toward debt and away from equity. We see this in Figure 9, which shows the cumulated flows of foreign assets into U.S. debt liabilities and equity and FDI liabilities since 1990. Cumulated flows into debt were growing exponentially throughout this period, whereas flows into equity and FDI reached an inflection point.

---

15 This phase resembled the Japanese (real estate) asset shopping spree during the 1980s, which was not targeted to safe assets or their safe tranche but to whole assets.
after the dot-com crash and 9/11 attacks, tapering off thereafter. After the crash the world came to realize that there was substantial risk in U.S. assets as well and decided to refocus on the safe tranches of the asset distribution.

Figure 8.

Source: Gourinchas and Rey (2007).
II.C. Complexity and Systemic Risk Build-up

“Banks” generated safe assets from risky ones by creating complex instruments that pooled assets, such as subprime mortgages, and creatively divvied up the flows from those assets.

Figure 10 shows a caricature of the balance sheet of a CDO. A CDO can be defined by two properties: (1) pooling of assets; and (2) tranching of liabilities. The left (asset) side of the figure shows individual assets owned by the CDO, represented by blue boxes; if the CDO is a RMBS or CMBS CDO, the blue boxes are bundles of residential or commercial mortgages. These blue boxes are the CDO’s assets. It is here on the asset side that pooling takes place. The right (liability) side are the liabilities of the CDO; it is
here is where the tranching takes place. The bottom tranche is the equity tranche, sometimes referred to as “toxic waste.” Although the implementation details vary from CDO to CDO, a given non-equity liability is protected from loss because equity and all the more junior liabilities have to be wiped out before the given liability suffers any loss. These more junior liabilities serve as a buffer. The arrow shows the direction of the cash flows. Similar instruments were created from securitization of all sorts of payment streams, ranging from auto to student loans.

Figure 10. The Balance Sheet of a CDO

Consider a CDO with two $1 mortgages as assets. Suppose the probability of default for each mortgage is 10 percent and that if in default, the mortgage is worthless. Suppose further the defaults are uncorrelated. There are (essentially) three states of the world: no defaults; one default; and two defaults. What if the CDO had only one class of liabilities consisting of two $1 liabilities that each received equal payoffs from the assets? Then the $1 liability would pay in full 81 percent of the time; it would pay 50 cents 18 percent of the time; and it would be worthless one percent of the time. An individual mortgage is worthless ten percent of the time, so this asset is “safer” than the individual mortgages, because of diversification. But CDOs create even safer assets by
dividing their liabilities into tranches with different seniorities. Suppose there are a $1 junior tranche and a $1 senior tranche. The senior tranche pays in full 99 percent of the time. Thus, a “safe” asset is created. However, it is key to note that the senior tranche is very exposed to “systematic” risk; it only defaults when both of the mortgages default. (Consider an additional asset in the economy, which also pays in full 99 percent of the time; suppose the asset’s payoffs are uncorrelated with the mortgage payoffs. During the global downturn, the new asset defaults only 1 percent of the time, whereas the senior tranche defaults 100 percent of the time.) This point has been rigorously explored by Coval, Jurek, and Stafford (2009a,b), who call structured finance products like CDOs “economic catastrophe bonds.” By design, it is precisely during economic catastrophes when defaults on the asset side of the balance sheet are large enough to eradicate the buffer protecting the senior tranches.

As private-label securitizations increased, the price of real estate and other assets in short supply rose sharply, in part reflecting the value associated with creating financial instruments from them. A positive feedback loop was created, as the rapid appreciation of the underlying assets seemed to justify a large triple-A tranche for derivative CDOs and related products. Credit rating agencies contributed to this loop, and so did greed and misguided homeownership policies, but as stated earlier: They were not the root cause.

Figure 11 shows the fast growth of global private-label securitization issuance during this period. The systemic fragility of these instruments became a source of systemic risk in itself once a significant share of them was kept within the financial system rather than sold to final unleveraged investors. Banks and their SIVs, attracted by the high return and low capital requirement combination provided by the senior and super-senior tranches of structured products, kept them on their books, sometimes passing

---

16 The light-blue bars represent CDO^2, CDOs that have other CDOs as their assets. Global issuance of CDO^2 increased from $25 billion in 2000 to $338 billion in 2006. As Coval, Jurek, and Stafford (2009a,b) point out, the use of CDO^2 only increases the correlation between default and systemic events.
their (then perceived as) infinitesimal risk onto the monolines and insurance companies (AIG, in particular).\footnote{17 An important reason for the creation of SIVs stemmed from regulatory arbitrage. From Brunnermeier (2009): “In hindsight, it is clear that one distorting force leading to the popularity of structured investment vehicles was regulatory and ratings arbitrage. The Basel I Accord (an international agreement that sets guidelines for bank regulation) required that banks hold capital of at least eight percent of the loans on their balance sheets; this capital requirement (called a ‘capital charge’) was much lower for contractual credit lines. Moreover, there was no capital charge at all for ‘reputational’ credit lines – noncontractual liquidity backstops that sponsoring banks provided to structured investment vehicles to maintain their reputation. Thus, moving a pool of loans into off-balance-sheet vehicles, and then granting a credit line to that pool to ensure a AAA-rating, allowed banks to reduce the amount of capital they needed to hold to conform with Basel I regulations while the risk for the bank remained essentially unchanged. The subsequent Basel II Accord, which went into effect on January 1, 2007 in Europe but is yet to be fully implemented in the United States, took some steps to correct this preferential treatment of noncontractual credit lines, but with little effect. While Basel II implemented capital charges based on asset ratings, banks were able to reduce their capital charges by pooling loans in off-balance-sheet vehicles. Because of the reduction of idiosyncratic risk through diversification, assets issued by these vehicles received a better rating than did the individual securities in the pool. In addition, issuing short-term assets improved the overall rating even further, since banks sponsoring these structured investment vehicles were not sufficiently downgraded for granting liquidity backstops.” According to Tett (2009), one reason the large banks kept exposure to the super senior tranches on their books was that AIG, which in the early stages of the boom provided insurance on these tranches, eventually decided it had too much exposure and stopped providing this insurance.}}
Through this process, the core of the financial system became interconnected in increasingly complex ways and vulnerable to a systemic event. According to Acharya and Schnabl (2009), “about 30% of all triple-A asset-backed securities remained within the banking system, and if one includes ABCP conduits and SIVs that had recourse, this fraction rises to 50%.” A November 2007 announcement by Citigroup shows its investment bank had $43 billion in exposure in super senior tranches of ABS CDOs “primarily” backed by subprime residential mortgages. Notably, Citi’s investment bank’s total direct exposure to U.S. subprime mortgages was $55 billion, so almost 80 percent of its exposure was through super senior tranches. By 2007, the Federal Reserve Bank of New York calculated that SIVs and similar vehicles had combined assets of $2.2 trillion, more than the assets of hedge funds ($1.8 trillion), and more than half the total assets of the five largest broker dealers ($4 trillion) (Tett 2009).

**III. The Crisis**

The conditions were thus ripe for a severe systemic event, which eventually came with far more force than anyone had anticipated. There was the shock from declining real estate prices and the corresponding rise in subprime defaults, but even under the most pessimistic scenarios, these shocks pale in comparison with the magnitude of the crisis that eventually followed once the panic set in. In this section I attempt to disentangle these two components, shock and panic, as a clarifying distinction is important for the policy prescription that follows in the next section.

**III.A. The Shock**

Much of the crisis is blamed on the crash of the real estate “bubble” and the rise in subprime mortgage defaults that ensued. But this cannot be all, or even much, of the story. The global financial system went into cardiac arrest mode and was on the verge of imploding more than once, which seems hard to attribute to a relatively small shock such as the real-estate/subprime combo (see Caballero 2009).
In Caballero and Kurlat (2009), we constructed an estimate of how much banks’ initial mortgage-related losses were amplified by the crisis that these losses sparked. For this estimate, we computed the evolution of the market value (equity plus long term debt) of the major U.S. banks since January 2007, which yielded an estimate of total losses on the right side of these banks’ balance sheets. Absent any feedback effects, these losses should be equal to the losses suffered by the assets on the left side of the balance sheets. However, as illustrated in Figure 12, we find that losses on the right side are on the order of three times the IMF’s (evolving) estimates of losses related to mortgage assets accruing to U.S. banks.

Beginning in 2008, and increasingly after the fall of Bear Stearns, the overall loss in market value became larger than the losses from subprime assets alone. The market began to price its losses from the overall disruption of financial markets, the severe recession, and losses on other types of assets which far exceeded the estimated losses from the mortgage market itself.

18 The procedure for estimating this was as follows: For equity, we simply tracked the evolution of each bank’s market capitalization, excluding increases in the market cap due to issues of new shares. For debt, we estimated the duration of each bank’s long term debt (including any preferred shares) from the maturity profiles described in the 10-K statements as of December 2007, assuming the interest rate was equal to the rate on 10-year Treasuries plus the spread on 5-year CDS for each bank, obtained from JP Morgan. Assuming an unchanged maturity profile, we then tracked the changes in the implied market value of each bank’s long term debt on the basis of the evolution of the CDS spread. The banks included in the calculation are the 19 banks that underwent the “stress tests” plus Lehman, Bear Stearns, Merrill Lynch, Wachovia, and Washington Mutual.

19 The IMF uses a projection of macroeconomic variables and default rates to estimate losses on loans, and market values to estimate losses on subprime-related securities. To the extent that market prices of securities overreacted due to fire sales, our procedure understates the multiplier.
Figure 12. Losses from mortgage assets, total loss of market value and multiplier.

![Graph showing losses from mortgage assets, total loss of market value and multiplier.](image)


### III.B. The Panic

The real damage came from the unexpected and sudden freezing of the entire securitization industry. The blue line in Figure 13 is “New Issuance of Asset-Backed Securities in Previous 3 Months,” from Adrian and Shin (2009); the data originally come from JP Morgan Chase. The crisis in this market is apparent from the disappearance of new issuances. The red line is the implied spread on the 2006-1 AAA ABX, which measures the cost of insuring against default by triple-A tranches of subprime mortgage-backed securities of the first-half-of-2006 vintage. The spread data are from JP Morgan Chase and not only corroborates the crisis impression from the quantity side but also makes it clear that the collapse in quantity is demand rather than supply driven.

---

The new issuance series is the sum over the following categories of ABS: “home equity (subprime)”; commercial real estate; autos; credit cards; student loans; non-US residential mortgages; and other. The data were provided by Tobias Adrian.
Confidence vanished and the complexity which made possible the “multiplication of bread” during the boom, turned into a source of counterparty risk, real and imaginary (see Gorton 2008). Senior and super-senior tranches were no longer perceived as invulnerable, and making matters worse, banks had to bring back into their balance sheets more of this new risk from the now struggling SIVs and conduits.

In December 2007, Citigroup provided a guarantee facility to its SIVs, essentially bringing the SIVs’ $49 billion in assets onto its balance sheet.21 About 60 percent of the SIVs’ assets were financial-institutions debt. Thirty-nine percent were structured-finance assets, including: U.S. (7 percent of total assets) and non-U.S. (12 percent) residential

---

MBS; CBOs, CDOs, and CLOs (6 percent); student loans (5 percent), and credit card (5 percent) assets. One hundred percent of the structured-finance assets carried Moody’s triple-A ratings.

Knightian uncertainty took over, and pervasive flight to qualities plagued the financial system. Fear fed into more fear, and caused reluctance to engage in financial transactions, even among the prime financial institutions (see Figure 14).

**Figure 14.**

![TED Spread](chart)

Source: Global Financial Data.

As the crises spread there was a sharp shift of the maturity structure of asset-backed commercial paper (ABCP) toward very short-term maturities. At the height of the crisis, nearly 80 percent of the asset-backed paper issued had a maturity of only one to four days, an increase of more than 40 percentage points since January 2004. This
shortening of the maturity structure was fairly gradual, but there was a pronounced decline in September 2008, when Lehman collapsed. As shown in Figure 15, the average maturity of newly issued ABCP declined from 17 days in August 2008, to 9 days in September 2008. It has long been understood that funding through short-maturity liabilities generates rollover risk, but as it has been documented in the context of emerging market crises by Broner et al (2008), that borrowers do not have many other options during severe financial crises.

Figure 15.

![Average maturity of newly issued ABCP (estimate)](image)

Sources: Federal Reserve, author’s estimate.\(^{22}\)

In the months following Lehman’s collapse, there was also a sudden move toward very short-term maturities in the CP market for financial companies, as shown in Figure 16.\(^{23}\)

\(^{22}\) The weighted average maturity is estimated as follows: [1-4] day CP is treated as \((4+1)/2 = 2.5\) days; [41-80] day CP is treated as \((80+41)/2 = 60.5\) days; etc. For [80+] day CP, a value of 90 days was used.
The decline was precipitous because market participants were surprised by the depth of financial firms’ difficulties and the chaotic aftermath of Lehman’s collapse, whereas participants knew (since at least summer 2007) about problems in the ABS market. In fact, as problems festered in the ABS market during the early phases of the crisis, financial firms conservatively lengthened their CP maturities. In August 2008, 50 percent of the commercial paper issued by financial companies matured in one to four days. By November, 87 percent of the paper issued matured in one to four days.

Figure 16.

![Average maturity of newly issued financial CP (estimate)](chart)

Sources: Federal Reserve, author’s estimate.

Along the way, the underlying structural deficit of safe assets that was behind the whole cycle, worsened as the newly found source of triple-A assets from the securitization industry dried up, and the spike in uncertainty further increased demand for these assets. Safe interest rates plummeted to record low levels (see Figure 17).

23 There was also substantial volatility of in the maturity structure during this period, likely the result of a desire of financial firms to push out the maturity structure and a willingness of investors to absorb longer maturities that varied with volatile market conditions and expectations of government involvement.
Initially, the flight to quality was a boon for the money market funds, which suddenly found themselves with a herd of new clients. In order to capture this expansion in demand from clients having a preference for riskier assets than their usual clients, some money market funds began to invest in short-term commercial paper from the investment banks in distress. This strategy backfired after Lehman’s collapse, when the Reserve Primary Fund “broke-the-buck” as a result of its losses associated with Lehman’s bankruptcy. Perceived complexity reached a new level as even the supposedly safest private funds were no longer immune to contagion. Widespread panic
took over and had it not been for the massive and concerted intervention taken by governments around the world, the financial system would have imploded.\textsuperscript{24}

While the end of the panic has removed some of the pressure on the safe-assets world, the crisis destroyed a significant share of the financial industry created by the private sector to satisfy the large demand for safe assets around the world. The shortage of triple-A assets is now worse than it was before the crisis, and unless a solution to this gap is found soon, many of the weaknesses that were created before the crisis will reemerge in the same or a mutated form.

\textbf{IV. The Policy Problem}

The core policy problem to deal with is how to bridge the safe-asset gap without over-exposing the financial sector to systemic risk. There are essentially two options: Either the government takes care of supplying much of the triple-A assets or it lets the private sector do much of it but it backstops extreme systemic events, a sort of public-private venture. Here I develop the latter option.

The reason I view the public-private venture as an option despite the recent crisis, is that that the main failure was not in the private sector’s ability to create triple-A assets through complex financial engineering (although rating agencies may have excessively facilitated the process), but in the systemic vulnerability created by this process. The idea of the public-private venture is to preserve the successful parts of this asset

\textsuperscript{24} Reserve Primary Fund had invested $785 million in Lehman debt, which constituted about 1.2\% of its assets. Immediately after Lehman filed for bankruptcy, the fund suffered a massive run, with over $30 billion in redemption requests (about half of its total assets) before it stopped accepting redemption requests at $1 at 11 a.m. the following day. Money market funds had been considered extremely safe, and had indeed benefited from the flight to quality during the previous year, growing by about $850 billion (34\%) since mid 2007. The drop in the Reserve Primary Fund’s NAV caused investors to question the safety of the entire industry. There were net redemptions for about $170 billion during that week, as well as a large shift from prime funds towards funds investing exclusively in government debt. In order to stem the panic, on September 19 the U.S. Treasury announced a guarantee program that would compensate investors if the NAV of participating funds fell below $1.
creation activity while finding a mechanism to reallocate the systemic risk component it creates from the banks’ balance sheet to private unlevered investors (for small and medium size shocks) and the government (for tail events).

Raising capital requirements is a kneejerk policy reaction to reduce vulnerability but it does not help to deal with the structural problem of excess safe-asset demand and it may well worsen it. We need a more balanced response, trading off vulnerability reduction and the safe-asset gap, in determining the socially optimal level of capital requirements (which may well be higher than the pre-crisis levels, especially for illiquid assets) and complementary measures.

There are two broad categories of recent proposals to reduce crisis risk without excessively limiting the financial sector’s ability to bridge the safe-assets gap:

- Pre-paid/arranged contingent capital injections, and
- Pre-paid/arranged contingent asset and capital insurance injections.

The basic purpose of the former, contingent capital injections, is to reduce the costs associated with the holding of capital when it is not needed. However, and centrally, these approaches recognize that access to capital during crises needs to be arranged in advance, since it is often hard to raise capital during a severe crisis. Proposals of this kind differ in their sources of this contingent capital, in particular, between the private sector and the government. Within the former, in some proposals the contingent funds come primarily from existing stakeholders (e.g., through contingent debt/equity swaps) while in others the funds come from outsiders. However, outsiders’ commitments problems limit the extent to which the private sector can serve as the source of this capital during extreme events, a point highlighted by Holmström and Tirole (1998) in theory and AIG (and the monolines) in practice.
Flannery’s (2002) proposal made one of the first significant steps in this direction with his proposal for “reverse convertible debentures.”25 Such debentures would convert to equity whenever the market value of a firm’s equity falls below a certain threshold.

One problem of this early proposal is that it made no distinction between aggregate and idiosyncratic shocks. The Kashyap et al (2008) proposal deals with this distinction and calls for banks to buy capital insurance policies that pay off when the banking sector experiences a negative systemic shock.26 Private investors would underwrite the policies and place the amount insured into a “lock box” invested in US Treasuries. Investors who are themselves subject to capital requirements would not be allowed to supply this insurance. The insurance would be triggered when aggregate bank losses over a certain number of quarters exceed some significant amount; losses at the covered bank would not be included in determining whether the insurance is triggered.

Combining both contributions, the Squam Lake Working Group on Financial Regulation has a proposal (2009) similar to Flannery’s except that conversion from debt to equity is triggered only during systemic events and only for banks that violate certain capital-adequacy covenants.27

Yet another variant on capital insurance is for the insurance policy to pay out to the regulator, instead of the firm. Under this proposal, by Acharya and others, the amount of insurance required would be proportional to an estimate of the systemic risk posed by the bank, in order to discourage firms ex-ante from taking on excessive systemic risk.

Hart and Zingales advocate an alternative approach; when spreads on a bank’s CDS rise above a certain threshold, a regulator allows the bank a window of time to issue equity

in order to bring the CDS spread back below the threshold. If the bank is unable to reduce its CDS spread, the regulator reviews the bank’s books and determines whether the bank’s debt is at risk. If the regulator determines the bank’s debt is not at risk, the regulator invests in the bank by lending to the bank; otherwise, the regulator replaces the CEO with a trustee, who will liquidate the bank and pass the proceeds to the bondholders. Although this approach does have a contingent capital-injection component, it also relies heavily on the resolution of financial firms, which can be a useful disciplinary device during normal times but can be highly counterproductive during a systemic episode.

More generally, the contingent capital approach is the right one when the crisis is mostly one of fundamentals. However, if the panic component is significant, a central feature of most financial crises, then it is not the most cost-effective, and it may well trigger further panic as fear of dilution and forced conversion increases.

This takes us to the second set of proposals, contingent insurance injections. The basic idea of this approach is that the pure panic component of a crisis does not require a costly capital injection to subside. All that is needed is a broad guarantee that resources will be available should conditions worsen. Despite its high notional value, the expected cost of such a policy is low because it derives its power from the very same feature that underlies the panic. That is, the enormous distortion in perceived probabilities of a catastrophe also means that economic agents greatly overvalue public insurance and guarantees. Providing these can be as effective as capital injections in dealing with the panic at a fraction of the expected cost (when assessed at reasonable rather than panic-driven probabilities of a catastrophe).

In Caballero and Krishnamurthy (2007) we showed that during an episode of Knightian uncertainty, a government or central bank concerned with the aggregate will want to provide insurance against extreme events even if it has no informational advantage over the private sector. The reason is that during a panic of this kind, each individual bank
and investor fears itself to be in a situation worse than the average, an event that cannot be true for the collective. By providing a broad guarantee, the government gets the private sector to react more than one-for-one with this guarantee since it also closes the gap between the true average and the average of panic-driven expectations.

During the current crises, there were many asset-insurance injection proposals.28 The argument for why it may be optimal to support assets rather than inject capital during a panic is developed in Caballero and Krishnamurthy (2008b). In practice, financial institutions face a constraint such that value-at-risk must be less than some multiple of equity. In normal times, this structure speaks to the power of equity injections, since these are “multiplied” many times in relaxing the value-at-risk constraint. In contrast, insuring assets reduces value-at-risk by reducing risk directly, which typically does not involve a multiplier. However, when uncertainty is rampant, some illiquid and complex assets, such as CDOs and CDO-squared, can reverse this calculation. In such cases, insuring the uncertainty-creating assets reduces risk by multiples, and frees capital, more effectively than directly injecting equity capital.

Moreover, it turns out that the same principle of insurance-injection can be used to recapitalize banks when this the chosen solution. Rather than directly injecting capital, the government can pledge a minimum future price guarantee for newly privately raised capital (Caballero 2009a). This mechanism is very powerful because private investors overvalue the guarantee, and because the recapitalization itself makes a catastrophic event less likely. Caballero and Kurlat (2009a) quantified this mechanism and showed that once the equilibrium response of equity prices is taken into account, this mechanism significantly reduces the effective exposure of government resources relative to a public equity injection.

Many of the actual programs implemented during the crisis had elements of guarantees rather than being pure capital injections. Perhaps the clearest case of this approach is

28 See, e.g., Caballero (2009a,b); Mehrling and Milne (2008) and Milne (2009) for proposals.
that followed by the UK. Their asset protection scheme, announced in January 2009, provided insurance against 90 percent of losses above a “first-loss” threshold on portfolios of corporate and leveraged loans, commercial and residential property loans, and structured credit assets such as RMBS, CMBS, CLO, and CDO obligations. The insurance is provided in exchange for a fee. The APS covered £552 billion portfolios of RBS and Lloyds Banking Group, with a first-loss amount of £19.5 billion and £25 billion, respectively. The main criticism to the U.K.’s approach is that they charged such a high fee for the insurance that most banks chose not to engage, leaving the overall economy more exposed to their failure than socially optimal.

In Caballero and Kurlat (2009b) we proposed a policy framework which would not only guarantee access to insurance in the event of an SFA episode, but it would do so in a flexible manner that integrates the role of the government as an insurer of last resort with private sector information on the optimal allocation of contingent insurance.

Under our proposal, the government would issue tradable insurance credits (TICs) which would be purchased by financial institutions, some of which would have minimum holding requirements. During a systemic crisis, each TIC would entitle its holder to attach a central bank guarantee to newly-issued and legacy securities. All regulated financial institutions would be allowed to hold and use TICs, as well as private equity funds, corporations, and possibly hedge funds. In principle, TICs could be used as a flexible and readily available substitute for many of the facilities that were created during the crisis. The basic mechanism would consist of attaching them to assets, but variants could include attaching them to liabilities and even equity, depending on the particular needs of the distressed institutions and markets, and they could also operate as collateral-enhancers for discount window borrowing.

---

29 Recent press reports suggest that Lloyds may seek to reduce dilution related to its participation in the APS by asking the government to insure fewer assets than it agreed in March. “Lloyds May Still Join Government Asset-Protection Plan,” Wall Street Journal, September 18, 2009.
TICs are equivalent to CDS during systemic crises but not during normal times. That is, TICs are contingent-CDS. They become activated only when a systemic crisis arises. By targeting the event that needs protection, this contingent feature significantly lowers the cost of insurance for financial institutions.

Note also that TICs’ tradability would allow private agents to use markets to reallocate the access to insurance toward financial institutions in most dire need. And if distressed institutions chose to not seek to stock up on TICs and risk their survival for a higher return (as probably Lehman did and failed), at the very least the rest of the financial system would be better protected against the turmoil that could arise if the misbehaved institutions fails as they would be holding the TICs.

To conclude, it is important to highlight that the point of these insurance arrangements is to remove (for a fee) the systemic risk from leveraged and interconnected financial institutions, while they continue to produce the triple-A assets whose shortage is behind many of the main global macroeconomic phenomena of the last two decades.

V. Final Remarks

The world entered the crisis with a shortage of safe financial assets and emerged with an even more acute deficit. Moreover, the crisis itself was exacerbated by a widespread panic that rendered the private sector unable to supply these assets precisely when demand was at its peak.

It turns out that a single policy instrument, a contingent insurance to the key financial players, can help with both problems: It can preserve the private sectors ability to engineer safe assets from relatively risky assets through securitization and tranching; and it can reduce the severity of panics during financial crises.
Bibliography


Lehman Brothers. 2008. “Residential Credit Losses - Going into Extra Innings?”

Mehrling, Perry and Alistair Milne, 2008. “Government’s role as credit insurer of last resort and how it can be fulfilled,” mimeo October.


