The financial cycle and macroeconomics:
What have we learnt?

by Claudio Borio

Monetary and Economic Department

December 2012

JEL classification: E30, E44, E50, G10, G20, G28, H30, H50

Keywords: financial cycle, business cycle, medium term, financial crises, monetary economy, balance sheet recessions, balance sheet repair.
BIS Working Papers are written by members of the Monetary and Economic Department of the Bank for International Settlements, and from time to time by other economists, and are published by the Bank. The papers are on subjects of topical interest and are technical in character. The views expressed in them are those of their authors and not necessarily the views of the BIS.

This publication is available on the BIS website (www.bis.org).

© Bank for International Settlements 2012. All rights reserved. Brief excerpts may be reproduced or translated provided the source is stated.

ISSN 1020-0959 (print)
ISSN 1682-7678 (online)
The financial cycle and macroeconomics: What have we learnt?

Claudio Borio

Abstract

It is high time we rediscovered the role of the financial cycle in macroeconomics. In the environment that has prevailed for at least three decades now, it is not possible to understand business fluctuations and the corresponding analytical and policy challenges without understanding the financial cycle. This calls for a rethink of modelling strategies and for significant adjustments to macroeconomic policies. This essay highlights the stylised empirical features of the financial cycle, conjectures as to what it may take to model it satisfactorily, and considers its policy implications. In the discussion of policy, the essay pays special attention to the bust phase, which is less well explored and raises much more controversial issues.

JEL Classification: E30, E44, E50, G10, G20, G28, H30, H50

Keywords: financial cycle, business cycle, medium term, financial crises, monetary economy, balance sheet recessions, balance sheet repair.
Introduction

Understanding in economics does not proceed cumulatively. We do not necessarily know more today than we did yesterday, tempting as it may be to believe otherwise. So-called “lessons” are learnt, forgotten, re-learnt and forgotten again. Concepts rise to prominence and fall into oblivion before possibly resurrecting. They do so because the economic environment changes, sometimes slowly but profoundly, at other times suddenly and violently. But they do so also because the discipline is not immune to fashions and fads. After all, no walk of life is.

The notion of the financial cycle, and its role in macroeconomics, is no exception. The notion, or at least that of financial booms followed by busts, actually predates the much more common and influential one of the business cycle (eg, Zarnowitz (1992), Laidler (1999) and Besomi (2006)). But for most of the postwar period it fell out of favour. It featured, more or less prominently, only in the accounts of economists outside the mainstream (eg, Minsky (1982) and Kindleberger (2000)). Indeed, financial factors in general progressively disappeared from macroeconomists’ radar screen. Finance came to be seen effectively as a veil – a factor that, as a first approximation, could be ignored when seeking to understand business fluctuations (eg, Woodford (2003)). And when included at all, it would at most enhance the persistence of the impact of economic shocks that buffet the economy, delaying slightly its natural return to the steady state (eg, Bernanke et al (1999)).

What a difference a few years can make! The financial crisis that engulfed mature economies in the late 2000s has prompted much soul searching. Economists are now trying hard to incorporate financial factors into standard macroeconomic models. However, the prevailing, in fact almost exclusive, strategy is a conservative one. It is to graft additional so-called financial “frictions” on otherwise fully well behaved equilibrium macroeconomic models, built on real-business-cycle foundations and augmented with nominal rigidities. The approach is firmly anchored in the New Keynesian Dynamic Stochastic General Equilibrium (DSGE) paradigm.

The purpose of this essay is to summarise what we think we have learnt about the financial cycle over the last ten years or so in order to identify the most promising way forward. It draws extensively on work carried out at the BIS, because understanding the nexus between financial and business fluctuations has been a lodestar for the analytical and policy work of the institution. As a result, the essay provides a very specific and personal perspective on the issues, just one lens among many: it is not intended to survey the field.

The main thesis is that macroeconomics without the financial cycle is like Hamlet without the Prince. In the environment that has prevailed for at least three decades now, just as in the one that prevailed in the pre-WW2 years, it is simply not possible to understand business fluctuations and their policy challenges without understanding the financial cycle. This calls for a rethink of modelling strategies. And it calls for significant adjustments to macroeconomic policies. Some of these adjustments are well under way, others are at an early stage, yet others are hardly under consideration.

---

1 This essay is a slightly revised version of the one prepared for a keynote lecture at the Macroeconomic Modelling Workshop, “Monetary policy after the crisis”, National Bank of Poland, Warsaw, 13-14 September 2012. I would like to thank Carlo Cottarelli, Piti Disyatat, Leonardo Gambacorta, Craig Hakkio, Otmar Issing, Enisse Kharroubi, Anton Korinek, David Laidler, Robert Pringle, Vlad Sushko, Nikola Tarashev, Kostas Tsatsaronis, Michael Woodford and Mark Wynne for helpful comments and Magdalena Erdem for excellent statistical assistance. The views expressed are my own and do not necessarily reflect those of the Bank for International Settlements.
Three themes run through the essay. Think medium term! The financial cycle is much longer than the traditional business cycle. Think monetary! Modelling the financial cycle correctly, rather than simply mimicking some of its features superficially, requires recognising fully the fundamental monetary nature of our economies: the financial system does not just allocate, but also generates, purchasing power, and has very much a life of its own. Think global! The global economy, with its financial, product and input markets, is highly integrated. Understanding economic developments and the challenges they pose calls for a top-down and holistic perspective – one in which financial cycles interact, at times proceeding in sync, at others proceeding at different speeds and in different phases across the globe.

The first section defines the financial cycle and highlights its core empirical features. The second puts forward some conjectures about the elements necessary to model the financial cycle satisfactorily. The final one explores the policy implications, discussing in turn how to address the booms and the subsequent busts. The focus in the section is primarily on the bust, as this is by far the less well explored and still more controversial area.

1. The financial cycle: core stylised features

There is no consensus on the definition of the financial cycle. In what follows, the term will denote self-reinforcing interactions between perceptions of value and risk, attitudes towards risk and financing constraints, which translate into booms followed by busts. These interactions can amplify economic fluctuations and possibly lead to serious financial distress and economic dislocations. This analytical definition is closely tied to the increasingly popular concept of the “procyclicality” of the financial system (eg, Borio et al (2001), Danielsson et al (2004), Kashyap and Stein (2004), Brunnermeier et al (2009), Adrian and Shin (2010)). It is designed to be the most relevant one for macroeconomics and policymaking: hence the focus on business fluctuations and financial crises.

The next question is how best to approximate empirically the financial cycle, so defined. What follows considers, sequentially, the variables that can best capture it, its relationship with the business cycle, its link with financial crises, its real-time predictive content for financial distress, and its dependence on policy regimes.

1.1 Feature 1: it is most parsimoniously described in terms of credit and property prices

Arguably, the most parsimonious description of the financial cycle is in terms of credit and property prices (Drehmann et al (2012)). These variables tend to co-vary rather closely with each other, especially at low frequencies, confirming the importance of credit in the financing of construction and the purchase of property. In addition, the variability in the two series is dominated by the low-frequency components. By contrast, equity prices can be a distraction. They co-vary with the other two series far less. And much of their variability concentrates at comparatively higher frequencies.

It is important to understand what this finding does and does not say. It is no doubt possible to describe the financial cycle in other ways. At one end of the spectrum, like much of the extant work, one could exclusively focus on credit – the credit cycle (eg, Aikman et al (2010), Schularick and Taylor (2009), Jordá et al (2011), Dell’Arriccia et al (2012)). At the other end, one could combine statistically a variety of financial price and quantity variables, so as to extract their common components (eg, English et al (2005), Ng (2011), Hatzius et al (2011)). Examples of the genre are interest rates, volatilities, risk premia, default rates, non-performing loans, and so on. In between, studies have looked at the behaviour of credit and asset prices series taken individually, among other variables (eg, Claessens et al (2011a, 2011b)).
That said, combining credit and property prices appears to be the most parsimonious way to capture the core features of the link between the financial cycle, the business cycle and financial crises (see below). Analytically, this is the smallest set of variables needed to replicate adequately the mutually reinforcing interaction between financing constraints (credit) and perceptions of value and risks (property prices). Empirically, there is a growing literature documenting the information content of credit, as reviewed by Dell’Arricia et al (2012), and property prices (eg, IMF (2003)) taken individually for business fluctuations and systemic crises with serious macroeconomic dislocations. But it is the interaction between these two sets of variables that has the highest information content (see below).

1.2 Feature 2: it has a much lower frequency than the traditional business cycle

*The financial cycle has a much lower frequency than the traditional business cycle* (Drehmann et al (2012)). As traditionally measured, the business cycle involves frequencies from 1 to 8 years: this is the range that statistical filters target when seeking to distinguish the cyclical from the trend components in GDP. By contrast, the average length of the financial cycle in a sample of seven industrialised countries since the 1960s has been around 16 years.

Graph 1, taken from Drehmann et al (2012), illustrates this point for the United States. The blue line traces the financial cycle obtained by combining credit and property prices and applying a statistical filter that targets frequencies between 8 and 30 years. The red line measures the business cycle in GDP obtained by applying the corresponding filter for frequencies up to 8 years, as normally done. Clearly, the financial cycle is much longer and has a much greater amplitude. The greater length of the financial cycle emerges also if one measures it based on Burns and Mitchell’s (1946) turning-point approach, as refined by Harding and Pagan (2006). As the orange (peaks) and green (troughs) bars indicate, the length is similar to that estimated through statistical filters, and the peaks and troughs are remarkably close to those obtained with it.

**Graph 1**

*The financial and business cycles in the United States*

Orange and green bars indicate peaks and troughs of the financial cycle measured by the combined behaviour of the component series (credit, the credit to GDP ratio and house prices) using the turning-point method. The blue line traces the financial cycle measured as the average of the medium-term cycle in the component series using frequency-based filters. The red line traces the GDP cycle identified by the traditional shorter-term frequency filter used to measure the business cycle.


It might be objected that this result partly follows by construction. The filters used target different frequencies. And Comin and Gertler (2006) have already shown, the importance of the medium-term component of fluctuations exceeds that of the short-term component also for GDP.
But interpreting the result in this way would be highly misleading (Drehmann et al (2012)). The business cycle is still identified in the macroeconomic literature with short-term fluctuations, up to 8 years. Moreover, the relative importance and amplitude of the medium-term component is considerably larger for the joint behaviour of credit and property prices than for GDP. And individual phases also differ between both cycles. The contraction phase of the financial cycle lasts several years, while business cycle recessions generally do not exceed one year. In fact, as discussed further below, failing to focus on the medium-term behaviour of the series can have important policy implications.

1.3 Feature 3: its peaks are closely associated with financial crises

Peaks in the financial cycle are closely associated with systemic banking crises (henceforth “financial crises” for short). In the sample of seven industrialised countries noted above, all the financial crises with domestic origin (i.e., those that do not stem from losses on cross-border exposures) occur at, or close to, the peak of the financial cycle. And the financial crises that occur away from peaks in domestic financial cycles reflect losses on exposures to foreign such cycles. Typical examples are the banking strains in Germany and Switzerland recently. Conversely, most financial cycle peaks coincide with financial crises. In fact, there are only three instances post-1985 for which the peak was not close to a crisis, and in all of them the financial system came under considerable stress (Germany in the early 2000s, Australia and Norway in 2008/2009).

Graph 2, again taken from Drehmann et al (2012), illustrates this point for the United States and United Kingdom. The black bars denote financial crises, as identified in well known data bases (Laeven and Valencia (2008 and 2010), Reinhart and Rogoff (2009)) and modified by the expert judgment of national authorities. One can see that the five crises occur quite close to the peaks in the financial cycles. In all the cases shown, the crises had a domestic origin.

The close association of the financial cycle with financial crises helps explain another empirical regularity: recessions that coincide with the contraction phase of the financial cycle are especially severe. On average, GDP drops by around 50% more than otherwise (Drehmann et al (2012)). This qualitative relationship exists even if financial crises do not
break out, as also confirmed by other work, which either considers credit and asset prices together (Borio and Lowe (2004) or focuses exclusively on credit (eg, Jordá et al (2011)).

1.4 Feature 4: it helps detect financial distress risks with a good lead in real time

The close link between the financial cycle and financial crises underlies the fourth empirical feature: it is possible to measure the build-up of risk of financial crises in real time with fairly good accuracy. Specifically, the most promising leading indicators of financial crises are based on simultaneous positive deviations (or “gaps”) of the ratio of (private sector) credit-to-GDP and asset prices, especially property prices, from historical norms (Borio and Drehmann (2009), Alessi and Detken (2009)). One can think of the credit gap as a rough measure of leverage in the economy, providing an indirect indication of the loss absorption capacity of the system; one can think of the property price gap as a rough measure of the likelihood and size of the subsequent price reversal, which tests that absorption capacity. The combination of the two variables provides a much cleaner signal – one with a lower noise – than either variable considered in isolation.

Graph 3, taken from Borio and Drehmann (2009), illustrates the out-of-sample performance of the corresponding leading indicator for the United States. Danger zones are shown as shaded areas. The graph indicates that by the mid-2000s concrete signs of the build-up of systemic risk were evident, as both the credit gap and property price gap were moving into the danger zone. And as discussed there, the out-of-sample performance is quite good across countries.

Graph 3
Estimated gaps for the United States

Credit-to-GDP gap (percentage points) Real property price gap (%)

The shaded areas refer to the threshold values for the indicators: 2–6 percentage points for credit-to-GDP gap; 15–25% for real property price gap. The estimates for 2008 are based on partial data (up to the third quarter).

1 Weighted average of residential and commercial property prices with weights corresponding to estimates of their share in overall property wealth. The legend refers to the residential property price component.

Source: Borio and Drehmann (2009).

---

2 See also Eichengreen and Mitchener (2004), who find that credit and asset price booms exacerbated the Great Depression, using similar indicators as Borio and Lowe (2002).

3 If a single variable has to be chosen, then the evidence indicates that credit is the most relevant one; see, eg, Borio and Lowe (2002), Drehmann et al (2011) and Schularick and Taylor (2009). Drehmann and Juselius (2012) find that over a one-year horizon, the debt-service ratio provides even more reliable signals. The credit gap, by contrast, is superior over longer horizons, providing warnings further ahead.
In addition, there is growing evidence that the cross-border component of credit tends to outgrow the purely domestic one during financial booms, especially those that precede serious financial strains (Borio et al (2011), Avdjiev et al (2012)). This typically holds for the direct component – in the form of lending granted directly to non-financial borrowers by banks located abroad⁴ – and for the indirect one – resulting from domestic banks’ borrowing abroad and in turn on-lending to non-financial borrowers.⁵ The reasons for this regularity are not yet fully clear. One may simply be the natural tendency for wholesale funding to gain ground as credit booms, which is then reflected in rising loan-to-deposit ratios.⁶ But, as discussed further below, no doubt more global forces influencing credit-supply conditions are also at work (e.g., Borio and Disyatat (2011), Shin (2011), Bruno and Shin (2011), CGFS (2011)).

Graph 4, from Borio et al (2011), illustrates this feature for Thailand, ahead of Asian financial crisis in the 1990s, and for the United States and United Kingdom, ahead of the recent financial crisis. It shows the tendency for the direct (continuous blue line) and indirect (included in the dashed blue lines) components of credit to grow faster than overall domestic credit (red line) during such episodes. This is true regardless of the overall size of the direct foreign component relative to the domestic one in the stock of credit (shaded areas).

1.5 Feature 5: its length and amplitude depend on policy regimes

The length and amplitude of the financial cycle are no constants of nature, of course; they depend on the policy regimes in place.⁷ Three factors seem to be especially important: the financial regime, the monetary regime and the real-economy regime (Borio and Lowe (2002), Borio (2007)). Financial liberalisation weakens financing constraints, supporting the full self-reinforcing interplay between perceptions of value and risk, risk attitudes and funding conditions. A monetary policy regime narrowly focused on controlling near-term inflation removes the need to tighten policy when financial booms take hold against the backdrop of low and stable inflation. And major positive supply side developments, such as those associated with the globalisation of the real side of the economy, provide plenty of fuel for financial booms: they raise growth potential and hence the scope for credit and asset price booms while at the same time putting downward pressure on inflation, thereby constraining the room for monetary policy tightening.

---

⁴ Importantly, but rarely appreciated, the commonly used monetary statistics do not capture this component (Borio et al (2011)).

⁵ For a detailed description of the stylised facts associated with credit booms combining both macro and micro data and comparing industrial and emerging market economies, see Mendoza and Terrones (2008); for a recent survey, Dell’ Arriccia et al (2012).

⁶ Why this tendency (Borio and Lowe (2004))? Recall that credit and asset price booms reinforce each other, as collateral values and leverage increase. As a result, credit tends to grow fast alongside asset prices. By contrast, opposing forces work on the relationship between money (deposits) and asset prices. Increases in wealth tend to raise the demand for money (wealth effect). However, higher expected returns on risky assets, such as equity and real estate, as well as a greater appetite for risk, induce a shift away from money towards riskier assets (substitution effect). This restrains the rise in the demand for money relative to the expansion in credit. See also Shin (2011) for an emphasis on the role of non-core (wholesale) deposits.

⁷ This underlines a critical point: the financial cycle as defined in this essay should not be considered a recurrent, regular feature of the economy, which inevitably unfolds in a specific way (i.e., a regular and stationary process). Rather, it is a tendency for a set of variables to evolve in a specific way responding to the economic environment and policies within it. The key to this cycle is that the boom sets the basis for, or causes, the subsequent bust.
Graph 4

Credit booms and external credit: selected countries

<table>
<thead>
<tr>
<th>Country</th>
<th>In billions of US dollars</th>
<th>Year-on-year growth, in per cent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thailand in the 1990s</td>
<td><img src="image" alt="Graph" /></td>
<td><img src="image" alt="Graph" /></td>
</tr>
<tr>
<td>United Kingdom</td>
<td><img src="image" alt="Graph" /></td>
<td><img src="image" alt="Graph" /></td>
</tr>
<tr>
<td>United States</td>
<td><img src="image" alt="Graph" /></td>
<td><img src="image" alt="Graph" /></td>
</tr>
</tbody>
</table>

The vertical lines indicate crisis episodes end-July 1997 for Thailand and end-Q2 2007 and end-Q3 2008 for the United States and the United Kingdom. For details on the construction of the various credit components, see Borio et al (2011).  

- **Credit to the private non-financial sector granted by banks from offices located outside the country.**
- **Credit as in footnote (1) plus cross-border borrowing by banks located in the country.**
- **Credit as in footnote (2) minus credit to non-residents granted by banks located in the country.**


The empirical evidence is consistent with this analysis. As Graph 1 indicates, the length and amplitude of the financial cycle has increased markedly since the mid-1980s, a good approximation for the start of the financial liberalisation phase in mature economies (Borio and White (2003)). This date is also an approximate proxy for the establishment of monetary regimes more successful in controlling inflation. And the cycle appears to have become especially large and prolonged since the 1990s, following the entry of China and other former communist countries into the global trading system. By contrast, prior to the mid-1980s in, say, the United States the financial and traditional business cycles are quite similar in length.

---

1. Estimate of credit to the private non-financial sector granted by banks from offices located outside the country.
2. Estimate of credit as in footnote (1) plus cross-border borrowing by banks located in the country.
3. Estimate as in footnote (2) minus credit to non-residents granted by banks located in the country.

---

8. Indeed, the link between financial liberalisation and credit booms is one of the best established regularities in the literature, drawing in particular on the experience of emerging market economies. It was already evident following the experience of liberalisation in the Southern Cone countries of Latin America in the 1970s (eg, Diaz-Alejandro (1985), Baliño (1987)).
and amplitude (left-hand panel). In fact, across the seven economies covered in Drehmann et al (2012), the average length of the financial cycle is 16 years over the whole sample; but for cycles that peaked after 1998, the average duration is nearly 20 years, compared with 11 for previous ones.

Moreover, it is no coincidence that the only significant financial cycle ending in a financial crisis pre-1985 took place in the United Kingdom, following a phase of financial liberalisation in the early 1970s (Competition and Credit Control). That this was also a period of high inflation indicates that financial liberalisation, by itself, is quite capable of generating sizable financial cycles. That said, in those days the rise in inflation and/or the deterioration of the balance of payments that tended to accompany economic expansions would inevitably quickly call for a policy tightening, constraining the cycle compared with the policy regimes that followed.\(^9\)

2. The financial cycle: analytical challenges

A systematic modelling of the financial cycle should be capable of accommodating the stylised facts just described. This raises first-order analytical challenges. What follows considers three basic features that satisfactory models should be able to replicate and then makes some conjectures about what strategies could be followed to do this.

2.1 Essential features that require modelling

The first feature is that the financial boom should not just precede the bust but cause it. The boom sows the seeds of the subsequent bust, as a result of the vulnerabilities that build up during this phase. This perspective is closer to the prewar prevailing view of business fluctuations, seen as the result of endogenous forces that perpetuate (irregular) cycles. It is harder to reconcile with today’s dominant view of business fluctuations, harking back to Frisch (1933), which sees them as the result of random exogenous shocks transmitted to the economy by propagation mechanisms inherent in the economic structure (Borio et al (2001)).\(^{10}\) And it is especially hard to reconcile with the approaches grafted on the real-business-cycle tradition, in which in the absence of persistent shocks the economy rapidly returns to steady state. In this case, much of the persistence in the behaviour of the economy is driven by the persistence of the shocks themselves (eg, Christiano et al (2005), Smets and Wouters (2003)). Arguably, since shocks can be regarded as a measure of our ignorance, rather than of our understanding, this approach leaves much of the behaviour of the economy unexplained.

The second feature is the presence of debt and capital stock overhangs (disequilibrium excess stocks). During the financial boom, credit plays a facilitating role, as the weakening of financing constraints allows expenditures to take place and assets to be purchased. This in turn leads to misallocation of resources, notably capital but also labour, typically masked by the veneer of a seemingly robust economy. However, as the boom turns to bust, and asset prices and cash flows fall, debt becomes a forcing variable, as economic agents cut their

---

\(^9\) In addition, it is no coincidence that financial booms and busts of this kind were quite common during the gold standard, all the way up to the 1930s. This was the previous time in history in which a liberalised financial system coincided with a monetary regime that yielded a reasonable degree of price stability over longer horizon. See Goodhart and De Largy (1999) and Borio and Lowe (2002) for a discussion of these issues and for evidence.

\(^{10}\) See, in particular, Zarnowitz (1992) for a historical review of the business cycle literature. Of course, unless one is prepared to endogenise everything and shift to a deterministic world, shocks will inevitably play a role.
expenditures in order to repair their balance sheets (e.g., Fisher (1932)). Similarly, too much capital in overgrown sectors holds back the recovery. And a heterogeneous labour pool adds to the adjustment costs. Financial crises are largely a symptom of the underlying stock problems and, in turn, tend to exacerbate them. Current models generally rule out the presence of such disequilibrium stocks, and when they incorporate them, they assume them exogenously, do not see them as the legacy of the preceding boom and treat them as exogenous cuts in borrowing limits (Eggertsson and Krugman (2012)).

The third feature is a distinction between potential output as non-inflationary output and as sustainable output (Borio et al (2012)). Current thinking implicitly or explicitly identifies potential output with what can be produced without leading to inflationary pressures, other things equal (Okun (1962), Woodford (2003), Congdon (2008), Svensson (2011a)). In turn, it regards sustainability as a core feature of potential output: if the economy reaches it, and in the absence of exogenous shocks, the economy would be able to stay there indefinitely. To be sure, the specific definition of potential output is model-dependent. DSGE models rely on notions that are much more volatile than those envisaged by traditional macroeconomic approaches (Mishkin (2007), Basu and Fernald (2009)). That said, inflation is generally seen as the variable that conveys information about the difference between actual and potential output (the “output gap”), drawing on various versions of the Phillips curve. This is reflected in how potential output and the corresponding output gap are measured in practice. Except in purely statistical models based exclusively on the behaviour of the output series itself, the vast majority of approaches rely on the information conveyed by inflation (Boone (2000), Kiley (2010)). And yet, as the previous analysis indicates, it is quite possible for inflation to remain stable while output is on an unsustainable path, owing to the build-up of financial imbalances and the distortions they mask in the real economy. Ostensibly, sustainable output and non-inflationary output need not coincide.

Graph 5
Output gap estimates: comparing methodologies (full-sample estimates)
In percentage points

United States

Spain


This can make a considerable difference in practice. Graph 5, from Borio et al (2012), plots three different measures of the output gap for the United States and Spain: a traditional one based on a Hodrick-Prescott filter (green line), one derived from a full production function approach by the OECD (blue line) and one that adjusts the Hodrick-Prescott filter drawing on information about the behaviour of credit and property prices (red line). The production function approach relies on information about inflation: estimates of the non-accelerating inflation rate of unemployment (NAIRU) help pin down potential output. The estimates based on information about the financial cycle combine the growth rates of credit and property prices so as to identify financial booms and busts and to capture more precisely the cyclical
component of output. All of the estimates use observations for the full sample (they rely on two-sided filters). The graph clearly shows that, especially in the 2000s, the credit-adjusted output gaps pointed to output being considerably higher than potential than the other two indicators. By contrast, before the mid-1980s, the various estimates tracked each other quite closely for the United States, which is consistent with much more subdued financial cycles at the time.

Moreover, the differences are much larger if the estimates are based on real-time information, i.e., if the filters are only one-sided. For instance, Graph 6 shows that while the estimates based on the financial cycle indicate that output was well above potential during the financial boom of the 2000s, their real-time Hodrick-Prescott filter counterparts miss this completely. Moreover, in particular for the United States, for the financial cycle based estimates there is hardly any difference between the real-time and full-sample results, again in sharp contrast to those purely based on the Hodrick-Prescott filter or production function approach. This is critical for policy: the passage of time, by itself, does not rewrite history—a well known major drawback of traditional output gap measures.

Graph 6
Output gaps: real-time (one-sided) vs full-sample (two-sided) estimates
In percentage points

United States
Spain


2.2 How could this be done?

How best to incorporate the three key features just described into models is far from obvious. Even so, it is possible to make some preliminary suggestions. To varying degrees, they could help capture the intra-temporal and inter-temporal coordination failures that no doubt lie at the heart of financial and business cycles.

---

11 Allowing also for data revisions makes little difference to the results; see Borio et al (2012).

12 For a more comprehensive discussion of, and references to, the relevant literature, see Borio (2011a) and, for a technical treatment, Gertler and Kiyotaki (2010). For a discussion of the range of limitations in risk measurement and incentives that can explain the corresponding procyclicality, see, eg, Borio et al (2001). For a recent influential treatment of incentive problems, see Rajan (2005).

13 For those who prefer to derive macroeconomic models from micro foundations, this inevitably requires moving away from the representative agent paradigm. Assuming heterogeneous agents has already become quite common in order to incorporate features such as credit frictions; see Gertler and Kiyotaki (2010).
One step would be to move away from model-consistent (“rational”) expectations. Modelling the build-up and unwinding of financial imbalances while retaining the assumption that economic agents have a full understanding of the economy is possible, but artificial.\textsuperscript{14} Heterogeneous and fundamentally incomplete knowledge is a core characteristic of economic processes. As we all see in our daily lives, empirical evidence is simply too fuzzy to allow agents to resolve differences of views.\textsuperscript{15} And this fundamental uncertainty is a key driver of economic behaviour.

A second step is to allow for state-varying risk tolerance, ie for attitudes towards risk that vary with the state of the economy, wealth and balance sheets (eg, Borio and Zhu (2011)). There are many ways that this can be done. And even without assuming that preferences towards risk vary with the state of the economy, behaviour can replicate state-varying degrees of prudence and risk-taking.\textsuperscript{16} While strictly speaking not necessary, this assumption would naturally amplify financial booms and busts, by strengthening the effect of state-varying financing constraints. In addition, if one wished to model the serious dislocations of financial crises, allowing more meaningfully for actual defaults would be an important modification (eg, Goodhart and Tsomocos (2011)).

A third, arguably more fundamental, step would be to capture more deeply the monetary nature of our economies. As discussed in more detail in the next sub-section, models should deal with true monetary economies, not with real economies treated as monetary ones, as is sometimes the case (eg, Borio and Disyatat (2011)).\textsuperscript{17} Financial contracts are set in nominal, not in real, terms. More importantly, the banking system does not simply transfer real resources, more or less efficiently, from one sector to another; it generates (nominal) purchasing power. Deposits are not endowments that precede loan formation; it is loans that create deposits. Money is not a “friction” but a necessary ingredient that improves over barter. And while the generation of purchasing power acts as oil for the economic machine, it can, in the process, open the door to instability, when combined with some of the previous elements. Working with better representations of monetary economies should help cast further light on the aggregate and sectoral distortions that arise in the real economy when credit creation becomes unanchored, poorly pinned down by loose perceptions of value and risks.\textsuperscript{18} Only then will it be possible to fully understand the role that monetary policy plays in the macroeconomy. And in all probability, this will require us to move away from the heavy focus on equilibrium concepts and methods to analyse business fluctuations and to rediscover the merits of disequilibrium analysis, such as that stressed by Wicksell (1898) (Borio and Disyatat (2011)).\textsuperscript{19}

\textsuperscript{14} See, for example, the interesting approaches followed by Christiano et al (2008) and, more recently, Boissay et al (2012), He and Krishnamurthy (2012) and Brunnermeier and Yannikov (2012). More generally, it is easy to model coordination failures without assuming model inconsistent expectations.

\textsuperscript{15} For an interesting approach along these lines, see Kurz’s (1994) notion of “rational beliefs”. Also, Frydman and Goldberg (2011) allow for imperfect information in a way that is consistent with the predictive information content of “gap” measures such as those discussed above. For a recent survey, see Woodford (2012).

\textsuperscript{16} See, for instance, Borio and Zhu (2011) for a non-technical review of ways to introduce state-varying effective risk tolerance in the context of the “risk-taking channel” of monetary policy, ie the impact of monetary policy on risk perceptions and risk tolerance. For a recent formalisation of one of the possible mechanisms at work, via leverage and binding value-at-risk constraints, see Bruno and Shin (2011). For a short review of the empirical evidence, see Gambacorta (2009).

\textsuperscript{17} On the difference between “real” and “nominal” analysis, see in particular Schumpeter (1954) and Kohn (1986).

\textsuperscript{18} Building on the work of Wicksell (1898), such distortions played a key role in the work of economists such as von Mises (1912) and Hayek (1933).

\textsuperscript{19} Some analyses do consider contracts set in nominal terms (eg, Diamond and Rajan (2006)). Moreover, there is a growing literature that treats money as essential, improving over barter; see Williamson and Wright (2010)
2.3 The importance of a monetary economy: an example

It is worth illustrating the importance of working with better representations of monetary economies by considering an example: this is the popular view that global current account imbalances were at the origin of the financial crisis – what might be called the “excess saving” view. Arguably, this represents a questionable application of paradigms appropriate for “real” economies to what are in fact monetary ones (Borio and Disyatat (2011)).

According to the “excess saving” view, global current account surpluses, especially in Asia, led to the financial crisis in two ways. First, current account surpluses in those economies, and the corresponding in net capital outflows, financed the credit boom in the deficit countries at the epicentre of the crisis, above all the United States. Second, the ex ante excess of saving over investment reflected in those current account surpluses put downward pressure on world interest rates, especially on US dollar assets, in which much of the surpluses were invested. This, in turn, fuelled the credit boom and risk-taking, thereby sowing the seeds of the global financial crisis.

The core objection to this view is that it arguably conflates “financing” with “saving” – two notions that coincide only in non-monetary economies. Financing is a gross cash-flow concept, and denotes access to purchasing power in the form of an accepted settlement medium (money), including through borrowing. Saving, as defined in the national accounts, is simply income (output) not consumed. Expenditures require financing, not saving. The expression “wall of saving” is, in fact, misleading: saving is more like a “hole” in aggregate expenditures – the hole that makes room for investment to take place. For example, in an economy without any investment, saving, by definition, is also zero. And yet that economy may require a lot of financing, such as that needed to fund any gap between income from sales and payments for factor inputs. In fact, the link between saving and credit is very loose. For instance, we saw earlier that during financial booms the credit-to-GDP gap tends to rise substantially. This means that the net change in the credit stock exceeds income by a considerable margin, and hence saving by an even larger one, as saving is only a small portion of that income.

As specifically applied to the “excess saving” view, this translates into two criticisms: one concerns identities, the other behavioural relationships.

The criticism concerning identities is that it is gross, not net, capital flows that finance credit booms. In fact, the US credit boom was largely financed domestically (Graph 4). But to the extent that it was financed externally, the funding came largely from countries with a current account deficit (the United Kingdom) or in balance (the euro area). This explains why it was largely banks located there that faced serious financial strains. Moreover, a considerable portion of the funding was round-tripping from the United States (He and McCauley (2012)). More generally, the financial crisis reflected disruptions in financing channels, in borrowing and lending patterns, about which saving and investment flows are largely silent.20

The criticism concerning behavioural relationships is that the balance between ex ante saving and investment is best thought of as affecting the natural, not the market, interest

---

20 Consistent with this view, Jordá et al (2011) find that current account deficits do not add to the predictive content of credit booms for banking crises. Indeed, as noted in Borio and Disyatat (2011), some of the most disruptive credit booms in history that ushered in banking crises took place in countries that were experiencing surpluses, including the United States ahead of the Great Depression and Japan in the 1980s. By contrast, credit and asset price booms no doubt tend to go hand-in-hand with a deterioration of the current account, as domestic expenditures run ahead of output.
A long tradition in economics sees market interest rates as fundamentally monetary phenomena, reflecting the interplay between the policy rate set by central banks, market expectations about future policy rates and risk premia, as affected by the relative supply of financial assets and risk perceptions and preferences. By contrast, natural rates are unobservable, equilibrium concepts determined by real factors. And, just like with any other asset price, there is no reason to believe that market rates may not deviate from their natural counterparts for prolonged periods. This is true for both policy rates, set by central banks, and long term rates, critically influenced by market expectations and risk preferences. In fact, it is hard to see how natural rates, being an equilibrium phenomenon, could have been at the origin of the huge macroeconomic dislocations associated with the financial crisis. Moreover, empirically, the link between global saving and current accounts, on the one hand, and both short and long real interest rates, on the other, is quite tenuous (Graph 7).

Graph 7
Global current account imbalances, saving and interest rates

---

3. The financial cycle: policy challenges

What are the policy implications of the previous analysis? What follows considers, in turn, policies to address the boom and the bust. However, since policies that target the boom have

---

21 In the international context, the famous Metzler (1960) diagram, postulating that a real world interest rate equates the global supply of saving and the global demand for investment, or the more modern rendering by Caballero, Farhi and Gourinchas (2008), are clear examples of purely real analysis as applied to what are in fact monetary economies. In such models, by construction, there is no difference between saving and financing.

22 For variations on this theme, see Tobin (1961), Cochrane (2001) and Hördahl et al (2006).

23 This is why Borio and Disyatat (2011) argue that the real cause of the financial crisis was not "excess saving" but the "excess elasticity" of the international monetary and financial system: the monetary and financial regimes in place failed to restrain the build-up of unsustainable credit and asset price booms ("financial imbalances") (see below). Credit creation, a defining feature of a monetary economy, plays a key role in this story. For a similar overall perspective, and an attempt to formalise some of the mechanisms at work, see Shin (2011). For a recent discussion of the relevance of current accounts, see in particular Obstfeld (2011).
been extensively considered in the past and now command a growing consensus, the discussion focuses mainly on the bust. This is less well explored and more controversial.

3.1 Dealing with the boom

Addressing financial booms calls for stronger anchors in the financial, monetary and fiscal regimes. These can help constrain the boom and, failing that, improve the defences and room for policy manoeuvre to deal with the subsequent bust. Either way, they would help to address what might be called the “excess elasticity” of the system (Borio and Disyatat (2011)), ie, its failure to restrain the build-up of unsustainable financial booms (“financial imbalances”) owing to the powerful procyclical forces at play.

In the case of prudential policy, the main adjustment is to strengthen the macroprudential, or systemic, orientation of the arrangements in place (eg, Borio (2011a), Caruana (2010), CGFS (2010, 2012)). A key element is to address the procyclicality of the financial system head-on. The idea is to build up buffers in good times, as financial vulnerabilities grow, so as to be able to draw them down in bad times, as financial stress materialises. There are many ways of doing so, through the appropriate design of tools such as capital and liquidity standards, provisioning, collateral and margining practices, and so on. Basel III, for instance, has put in place a countercyclical capital buffer (BCBS (2010a), Drehmann et al (2010, 2011)). And the G20 have endorsed the need to set up fully fledged macroprudential frameworks in national jurisdictions; considerable progress has already been made (FSB-IMF-BIS (2011)).

In the case of monetary policy, it is necessary to adopt strategies that allow central banks to tighten so as to lean against the build-up of financial imbalances even if near-term inflation remains subdued (eg, BIS (2010), Caruana (2011) and Borio (2011b), Eichengreen et al (2011)) – what might be called the “lean option”. Operationally, this calls for extending policy horizons beyond the roughly 2-year ones typical of inflation targeting regimes and for giving greater prominence to the balance of risks in the outlook (Borio and Lowe (2002), Bean (2003)), fully taking into account the slow build-up of vulnerabilities associated with the financial cycle. As the timing of the unwinding of financial imbalances is highly uncertain, extending the horizon should not be interpreted as extending point forecasts mechanically. Rather, it is a device to help assess the balance of risks facing the economy and the costs of policy action and inaction in a more meaningful and structured way. Increasingly, central banks have been shifting in this direction, albeit quite cautiously (Borio (2011b)).

In the case of fiscal policy, there is a need for extra prudence during economic expansions associated with financial booms. The reason is simple: financial booms do not just flatter the balance sheets and income statements of financial institutions and those to whom they lend (Borio and Drehmann (2010)), they also flatter the fiscal accounts (Eschenbach and Schuknecht (2004), BIS (2010, 2012), Borio (2011a), Benetrix and Lane (2011)). Potential output and growth tend to be overestimated. Financial booms are especially generous for the public coffers, because of the structure of revenues (Suárez (2010), Price and Dang (2011)). And the sovereign inadvertently accumulates contingent liabilities, which crystallise as the boom turns to bust and balance sheet problems emerge, especially in the financial sector.

24 This implies raising interest rates more than conventional Taylor rules would call for, at least if output gaps are estimated in the traditional way (Taylor (2010)). Whether the new measures proposed in Borio et al (2012) and discussed above would make a significantly large difference is a question that deserves further research.

25 The latest addition has been the Bank of Canada (2011): in its recent review of its inflation targeting framework, it has adjusted it so as to allow for the lean option. The issue, however, remains controversial, see, eg, Shirakawa (2010), Issing (2012) and, for a critical view, Svensson (2011b). For a recent formalisation of the policy prescription, see also Woodford (2012b).
The recent experiences of Spain and Ireland are telling. The fiscal accounts looked strong during the financial boom: the debt-to-GDP ratios were low and falling and fiscal surpluses prevailed. And yet, following the bust and the banking crises, sovereign crises broke out.

Estimating cyclically-adjusted balances using financial cycle information can help to address these biases (Borio et al (2012)). Graph 8 illustrates this, by applying the measures of the output gap that incorporate financial cycle information to the fiscal balances of the United States and Spain. As can be seen, the difference between the corresponding estimates and those derived from simple Hodrick-Prescott filters or the production function approach can be very large and persistent, especially for real-time (one-sided) estimates. Moreover, these corrections relate only to the different measures of potential output: they do not allow for the structure of revenues or growing contingent liabilities.

Graph 8
Budget balances and cyclical adjustments
In percentage points

![Graph 8](image)


More generally, there is a risk that failing to recognise that the financial cycle has a longer duration than the business cycle could lead policymakers astray. This occurs in the context of what might be called the “unfinished recession” phenomenon (Drehmann et al (2012)). Specifically, policy responses that fail to take medium-term financial cycles into account can contain recessions in the short run but at the cost of larger recessions down the road. In these cases, policymakers may focus too much on equity prices and standard business cycle measures and lose sight of the continued build-up of the financial cycle. The bust that then follows an unchecked financial boom brings about much larger economic dislocations. In other words, dealing with the immediate recession while not addressing the build-up of financial imbalances simply postpones the day of reckoning.

Graph 9 (overleaf), from Drehmann et al (2012), illustrates this for the United States, although the phenomenon is more general. The graph focuses on two similar episodes: the mid-1980s-early 1990s and the period 2001-2007. In both cases, monetary policy eased strongly in the wake of the stock market crashes of 1987 and 2001 and the associated weakening in economic activity. At the same time, the credit-to-GDP ratio and property prices continued their ascent, soon followed by GDP, only to collapse a few years later and cause much bigger financial and economic dislocations. Partly because inflation remained rather subdued in the second episode, the authorities raised policy rates much more gradually. And the interval between the peak in equity prices and property prices (vertical lines) was considerably longer, roughly 5 rather than 2 years. From the perspective of the medium-term financial and business cycles, the slowdowns or contractions in 1987 and 2001 can thus be regarded as “unfinished recessions”.

15
3.2 Dealing with the bust

Not all recessions are born equal. The typical recession in the postwar period, at least until the mid-1980s, was triggered by a tightening of monetary policy to constrain inflation. The upswing was relatively short. Heavily regulated financial systems resulted in little debt or capital stock overhangs. And high inflation boosted nominal asset prices and, with financial restrictions in place, eroded the real value of debt.

The most recent recession in mature economies is the quintessential “balance sheet recession,” which follows a financial boom gone wrong against the backdrop of low and stable inflation. The preceding boom is much longer. The debt, capital stock and asset price overhangs are much larger. And the financial sector is much more damaged. The closest equivalent in mature economies is Japan in the 1990s.

---

26 Koo (2003) seems to have been the first to use such a term. He employs it to describe a recession driven by non-financial firms’ seeking to repay their excessive debt burdens, such as those left by the bursting of the bubble in Japan in the early 1990s. Specifically, he argues that the objective of financial firms shifts from maximising profits to minimising debt. The term is used here more generally to denote a recession associated with the financial bust that follows an unsustainable financial boom. But the general characteristics are similar, in particular the debt overhang. That said, we draw different conclusions about the appropriate policy responses, especially with respect to prudential and fiscal policy.
As noted earlier, and as confirmed by broader empirical evidence, these recessions are particularly costly (eg, BCBS (2010b), Reinhart and Rogoff (2009), Reinhart and Reinhart (2010), Dell’ Arriccia (2012)). They tend to be deeper, to be followed by weaker recoveries, and to result in permanent output losses: output may regain its previous long-term growth but fails to return to its previous trajectory. Arguably, these features reflect a mixture of factors: the overestimation of both potential output and growth during the boom; the misallocation of resources, notably the capital stock but also labour, during that phase; the oppressive effect of the debt and capital overhangs during the bust; and the disruptions to financial intermediation once financial strains emerge.

This suggests that the key policy challenge is to prevent a stock problem from leading to a long-lasting flow problem, weighing down on income, output and expenditures. And the goal has to be achieved in the context of limited room for policy manoeuvre: unless policy has actively leaned against the financial boom to start with, policy buffers will be very low. The capital and liquidity cushions of financial institutions will be strained; the fiscal accounts will show gaping holes; and policy interest rates will not be that far away from the zero lower bound.

Against this backdrop, it is critical to distinguish two different phases, which differ in terms of priorities: crisis management and crisis resolution (Borio (2011b), Caruana (2012a)). In crisis management the priority is to prevent the implosion of the financial system, warding off the threat of a self-reinforcing downward spiral with economic activity. If room is available, policies should be deployed aggressively to that end. This is the phase historically linked to central banks’ lender-of-last resort function, which, unless constrained by exchange rate pegs, can be accompanied by sharp cuts in policy rates, especially helpful in boosting confidence. In crisis resolution, by contrast, the priority is balance sheet repair, so as to lay the basis for a self-sustained economic recovery. Here addressing the debt overhang is essential. And policies need to be adjusted accordingly.

A problem is that traditional rules of thumb for policy may be less effective in addressing balance sheet recessions, because of the legacy of the financial booms and the headwinds of the bust. There is a risk that, to different degrees, these policies may buy time but also make it easier to waste it. This may store up bigger problems further down the road. What follows explores this issue considering, sequentially, prudential, fiscal and monetary policy. But before doing so, and in order to fix ideas, it is worth discussing more concretely two historical polar cases.

The first case, universally recognised as a good example, is how the Nordic countries addressed the balance sheet recessions they confronted in the early 1990s (Borio et al (2010)). The crisis management phase was prompt and short. The authorities stabilised the financial system through public guarantees and, where necessary, central bank liquidity support. Then, almost without any discontinuity, they tackled the crisis resolution phase. With an external crisis constraining the room for manoeuvre for monetary and fiscal policy, they addressed balance sheet repair head-on. They enforced comprehensive loss recognition (writedowns); they recapitalised institutions subject to tough tests, including though temporary public ownership; they sorted institutions based on viability; they dealt with bad assets, including though disposal; they reduced the excess capacity27 in the financial system and promoted operational efficiencies, so as to lay the basis for sustainable profitability. The recovery was comparatively quick and self-sustained.

The second case, generally regarded as an example not to follow, is what happened in Japan in the wake of its financial bust, which took place roughly at the same time in the early

---

27 This excess capacity is a natural legacy of the previous boom; see, eg, Philippon (2008) for the extraordinary expansion of the US financial system ahead of the recent crisis and, more generally, BIS (2012).
The authorities were slow to recognise the balance sheet problems and, without an equivalent external crisis, had much more room to use expansionary monetary and fiscal policies. Faced with political resistance to the use of public money, balance sheet repair was delayed for several years. The economy took much longer to recover.

Consider now the role of prudential policy more specifically. Ideally, if effective macroprudential frameworks were in place, capital and liquidity buffers could be drawn down to soften the blow to the financial system and the economy. But if the authorities have failed to build up buffers in good times and financial strains emerge, the challenge is to repair financial institutions’ balance sheets.

A possible pitfall here is to focus exclusively on recapitalising banks with private sector money without enforcing full loss recognition. While such a policy is intended to prevent a credit crunch and disorderly deleveraging, it is arguably suboptimal. In the presence of investors’ doubts about the quality of banks’ balance sheets, it fails to reduce the cost of equity and funding more generally. Just like Caesar’s wife, not only do banks’ balance sheets have to be impeccable, they also have to be seen to be impeccable. In addition, it can generate wrong incentives: to avoid the recognition of losses; to misallocate credit, by keeping bad borrowers afloat (ever-greening) while charging higher rates to healthy borrowers; and possibly to bet for resurrection. The key point is that in the crisis resolution phase, when reduction in overall debt and asset prices is inevitable, the issue is not so much the overall amount of credit but its quality (allocation). Over time, any misallocation can reduce potential output and growth – a form of “hysteresis” that can help explain the persistent output losses.

Consider fiscal policy next. The challenge here is to use the typically scarce fiscal space effectively, so as to avoid the risk of a sovereign crisis. A widespread view among macroeconomists is that expansionary fiscal policy through pump-priming (increases in expenditures and reductions in taxes) is comparatively more effective when the economy is weak. Economic agents are “finance-constrained”, unable to borrow as much as they would like in order to spend: their propensity to spend any additional income they receive is high (eg, Gali et al (2007), Roeger and in ’t Veld (2009), Eggertsson and Krugman (2012)). In addition, with slack in the economy and interest rates possibly already constrained by the zero lower bound, there is no incentive to tighten monetary policy in response (eg, Eggertsson and Woodford (2003), Christiano et al (2011)).

This view, however, does not seem to take into account the specific features of a balance sheet recession. If agents are overindebted, they may naturally give priority to the repayment of debt and not spend the additional income: in the extreme, the marginal propensity to consume would be zero. Moreover, if the banking system is not working smoothly in the background, it can actually dampen the second-round effects of the fiscal multiplier: the funds need to go to those more willing to spend, but may not get there. Importantly, the available empirical evidence that finds higher fiscal multipliers when the economy is weak does not condition on the type of recession (eg, IMF (2010)). And some preliminary new

---

28 The fact that, even prior to the recent financial crisis, fiscal positions were already on an unsustainable path, has further narrowed the room for manoeuvre; see Cecchetti et al (2010) and BIS (2012)).

29 Perotti (1999) is one exception, as he argues that this need not be the case if debt is high and increases in taxes distortionary. For a review of the literature, see Corsetti et al (2012).

30 Koo (2003) argues that absent such a fiscal expansion, the economy would go into a tail spin: nothing would offset firms’ attempts to cut debt; the economy is in a fundamental disequilibrium.

31 This applies symmetrically to a tightening of fiscal policy, if agents have already cut spending as far as possible to reduce the debt burden.
research that controls for such differences actually finds that fiscal policy is less effective than in normal recessions (see below). This is clearly an area that deserves further study.

More generally, if the problem is a stock problem, it stands to reason that fiscal policy could be more effective if it targeted this problem directly. The objective would be to use the public sector balance sheet to support repair and strengthen the private sector’s balance sheet. This applies to the balance sheets of financial institutions, through injections of public sector money (capital) subject to strict conditionality on loss recognition and possibly temporary public ownership. And it applies also to the balance sheets of the non-financial sectors, such as households, including possibly through various forms of debt relief. If the diagnosis is correct, this use of public money could establish the basis of a self-sustaining recovery by removing a key impediment to private sector expenditures. Moreover, as an owner or co-owner, the sovereign could actually make capital gains in the longer term, as was the case in some Nordic countries.

Importantly, this is not a passive strategy, but a very active one. It inevitably substitutes public sector debt for private sector debt. And it requires a forceful approach, in order to address the conflicts of interests between borrowers and lenders, between managers, shareholders and debt holders, and so on. It is not pure fiscal policy in the traditional macroeconomic sense, as it generally calls for a broader set of measures supported by the public purse. But it arguably holds a better prospect of truly jump-starting the economy, rather than risking being a bridge to nowhere. And a bridge too far can mean a sovereign crisis.

What about monetary policy? The key pitfall here is that extraordinarily aggressive and prolonged monetary policy easing can buy time but may actually delay, rather than promote, adjustment. This is true for both interest policy (changes in the short-term policy rate) and central bank balance sheet policy (changes in those balance sheets aimed at influencing financial conditions beyond the short-term interest rates, such as through large-scale asset purchases and liquidity support) (Borio and Disyatat (2010)).

Context matters. Monetary policy is likely to be less effective in stimulating aggregate demand in a balance sheet recession. Overly indebted economic agents do not wish to borrow in order to spend. A damaged financial system is less effective in transmitting the policy stance to the rest of the economy. All this means that, in order to have the same short-term effect on aggregate demand, policy will naturally be pushed further. But this also increases its side-effects.

There are at least four possible side-effects of extraordinarily accommodative and prolonged monetary easing. First, it can mask underlying balance sheet weakness. It is all too easy to underestimate what the ability to repay of both private and public sector borrowers would be under more normal conditions. It is easier to delay the recognition of losses (eg, ever-greening). And except when refinancing options can be exercised for free, as in the case of the US

---

32 For a discussion of this issue, see BIS (2012) and IMF (2012). Non-recourse mortgages, allowing households to “walk away” when the value of the property falls below that of debt, can be helpful from this perspective: they simultaneously facilitate the reduction of the debt burden and speed up the lenders’ recognition of losses. This is probably one reason why debt reduction in the United States has proceeded faster and gone further than in Spain.

33 This relies on the public sector’s superior ability to bring forward (real) resources from the future, underpinned by its power to tax (eg, Holmström and Tirole (2011)). This is why the creditworthiness of the sovereign is so critical.

34 See Borio (2011b), Hannoun (2012), Caruana (2012a) and BIS (2012) for details. See also White (2012).
mortgage market, the policy does not reduce the (present value of the) debt burden; in fact, it increases it.

Second, it can numb incentives to reduce excess capacity in the financial sector and even encourage betting-for-resurrection behaviour. Even if the economy is not buoyant, institutions may take on risks not commensurate with rewards, such as in trading activities, or in asset classes that may even inhibit growth, such as in commodities like oil. One may wonder whether JP Morgan’s highly publicised losses in the second quarter of 2012 might partly reflect such incentives.

Third, over time, it can undermine the earnings capacity of financial intermediaries. Extraordinarily low short-term interest rates and a flat term structure, associated with commitments to keep policy rates low and with bond purchases, compress banks’ interest margins. And low long-term rates sap the strength of insurance companies and pension funds, in turn possibly weakening the balance sheets of non-financial corporations, households and the sovereign. It is no coincidence that in Japan insurance companies came under serious strains a few years after banks did (eg, Fukao (2002)).

Finally, it can atrophy markets and mask market signals, as central banks take over larger portions of financial intermediation. Interbank markets tend to shrink (Baba et al (2005), BIS (2010)), and risk premia and activity tend to become unusually compressed as policymakers risk becoming the marginal buyers. For example, it is hard to believe that the highly negative risk premia that prevailed in the government bond markets of even highly indebted sovereigns in the first half of 2012 were not to a considerable extent the result of central bank purchases. In fact, lowering those yields was a policy objective.

Over time, political economy considerations may add to the side-effects (Borio and Disyatat (2010)). The central bank’s autonomy and, eventually, credibility may come under threat. Technically, central banks have a monopoly over interest rate policy, but not over balance sheet policy. Balance sheet policies can be properly assessed only in the context of the consolidated public sector balance sheet, which is much larger. They make central banks vulnerable to losses, which can undermine their financial independence. And they make them open to the criticism of overstepping their role, such as if they are perceived to finance public sector deficits directly (purchases of sovereign assets) or to favour one sector over another (purchases of private sector assets).

The key risk is that central banks become overburdened (Borio (2011b), BIS (2012)) and a vicious circle develops. As policy fails to produce the desired effects and if adjustment is delayed, central banks come under growing pressure to do more. A widening “expectations gap” then emerges, between what central banks are expected to deliver and what they can deliver. All this makes eventual exit harder and may ultimately threaten the central bank’s credibility. One may wonder whether some of these forces have not been at play in Japan, a country in which the central bank has not yet been able to exit.

On reflection, the basic reason for the limitations of monetary policy in a financial bust is not hard to find (Caruana (2012a)). Monetary policy typically operates by encouraging borrowing, boosting asset prices and risk-taking. But initial conditions already include too much debt, too-high asset prices (property) and too much risk-taking. There is an inevitable tension between how policy works and the direction the economy needs to take.

---

35 See BIS (2012) for a discussion of these issues and supporting empirical evidence. For the impact on pension funds, see Ramaswamy (2012).

36 For overviews of the empirical evidence on the impact of central bank balance sheet policies on financial markets and the macroeconomics in the major advanced economies, see, eg, Williams (2011), Cecioni et al (2011) and Meaning and Zhu (2012).
Recent empirical evidence is consistent with the relative ineffectiveness of monetary policy in a balance sheet recession (Bech et al (2012)). The authors examine seventy-three recessions in twenty-four advanced economies since the 1960s, distinguishing the twenty-nine that coincided with financial crises from the rest. They find that, considering the recession and subsequent recovery, monetary policy has less of an impact on output when financial crises occur (Graph 10, top panels). Moreover, in normal recessions, the more accommodative monetary policy is in the downturn, the stronger the subsequent recovery, but this relationship is no longer apparent if a financial crisis erupts (Graph 10, bottom panels). In addition, the same study finds that the faster the deleveraging in such recessions, the stronger the subsequent recovery. And the link between fiscal policy and the recovery is similar to that for monetary policy, also pointing to its relative ineffectiveness in balance sheet recessions.

**Graph 10**

**Monetary policy in the business cycle with and without financial crises**

The dotted line means that the regression coefficient is not statistically significant.


What about the exchange rate? No doubt, to the extent that monetary policy induces a depreciation of the exchange rate it can support balance sheet repair, except in the shortrun when that debt is denominated in a foreign currency (eg, Krugman (1999), Tovar (2005)). The depreciation boosts output and cash flows, helping repair. In fact, export-driven creditless recoveries have typically been the way out of financial crises (eg, Calvo et al (2006), Tang and Upper (2010)). The Nordic countries were no exception. That said, this mechanism does have limitations. The link between monetary policy easing and exchange rate depreciation is not always water-tight, as global factors may overwhelm domestic ones. In Japan, for instance, the yen proved very resilient despite the balance sheet strains. The option is less effective for large, relatively closed economies. It may be perceived to have
beggar-thy-neighbour connotations. And it may result in unwelcome exchange rate pressures and capital flows elsewhere, especially if economic and financial cycles are not synchronised.

This brings us to the global implications of national monetary policies. There is a risk that monetary policies that appear reasonable from the perspective of individual economies result in policies that are not appropriate in the aggregate (Borio (2011b)). Unusually accommodating policies in the core countries in the world can easily get transmitted to the rest of the world partly through resistance to exchange rate appreciation, as other countries find it hard to insulate themselves from the associated capital flows and pressures on their currency.

**Graph 11**  
**Very accommodating global monetary conditions**

In fact, for the world as a whole, aggregate monetary conditions appear to have been unusually accommodative for quite some time. There is a clear tension between the downward trend in the average inflation-adjusted policy rate and the upward trend in estimates of world trend growth (Graph 11, left-hand panel). Likewise, the policy rate has also been below the range of estimates of Taylor rules, even when these are based on traditional output gap measures that do not take into account the impact of the financial cycle.

37 For a similar view, see Eichengreen et al (2011) and, for a recent elaboration, Caruana (2012b). Borio (2011b) develops this argument also in the context of commodity prices. More generally, Borio and Filardo (2007) put forward, and find supporting evidence for, the hypothesis that in a world with highly integrated factor input and product markets, global measures of economic slack have gained in significance relative to purely domestic measures in the determination of inflation. For a supporting view, based on a theoretical macroeconomic model, see Martínez-García and Wynne (2010) and Engle (2012); for a sceptical one, see, eg, Ball (2006) and Woodford (2010).
(same graph, right-hand panel, from Hofmann and Bogdanova (2012)). And, of late partly as a result of central bank balance sheet policies, long term interest rates have followed a similar path and are now at historically low levels (same graph, centre panel). Worryingly, several emerging market economies have been seeing symptoms of the build-up of financial imbalances that are eerily reminiscent of those seen in mature economies ahead of their financial crisis. As some economies are struggling with financial busts, others are struggling with equally serious financial booms (BIS (2012)).

More generally, we may be witnessing a new form of time inconsistency, both within individual economies and at the global level (Borio (2011b)).38 We are all familiar with time inconsistency in the context of inflation. In this case, taking wages and prices as given, policymakers may be tempted to produce inflation in an ultimately unsuccessful effort to raise output and employment, as prices and wages catch up (Kydland and Prescott (1977)). Over time, inflation trends higher without lasting gains in output or employment. In the case of financial cycles in individual economies, policies fail to constrain the boom but respond aggressively to the bust. The end-result can be a downward trend in policy rates across cycles and increasing resort to balance sheet policies without lasting gains in terms of financial and macroeconomic stability. Globally, the monetary stance in the core economies is then transmitted to the rest of the world, reinforcing that downward trend. Such a trend is all too obvious across financial cycles in the data, as confirmed by the previous evidence.

4. Conclusion

It is high time we rediscovered the role of the financial cycle in macroeconomics: Hamlet has to get its Prince back. This is essential for a better understanding of the economy and for the design of policy. This essay has taken a small step in that direction, by highlighting some of the core empirical features of the financial cycle, suggesting what might be necessary to model it, and proposing adjustments in policy to address it more effectively.

At least five stylised empirical features of the financial cycle stand out. The financial cycle is best captured by the joint behaviour of credit and property prices. It is much longer, and has a much larger amplitude, than the traditional business cycle. It is closely associated with systemic banking crises, which tend to occur close to its peak. It permits the identification of the risks of future financial crises in real time and with a good lead. And it is highly dependent of the financial, monetary and real-economy policy regimes in place.

Modelling the financial cycle raises major analytical challenges for prevailing paradigms. It calls for booms that do not just precede but generate subsequent busts, for the explicit treatment of disequilibrium debt and capital stock overhangs during the busts, and for a clear distinction between non-inflationary and sustainable output, ie, a richer notion of potential output – all features outside the mainstream. Moving in this direction requires capturing better the coordination failures that drive financial and business fluctuations. This suggests moving away from model-consistent expectations, thereby allowing for endemic uncertainty and disagreement over the workings of the economy. It suggests incorporating perceptions of risk and attitudes towards risk that vary systematically over the cycle, interacting tightly with the waxing and waning of financing constraints. Above all, it suggests capturing more deeply the monetary nature of our economies, ie, working with economies in which financial intermediaries do not just allocate real resources but generate purchasing power ex nihilo and in which these processes interact with loosely anchored perceptions of value, thereby

38 Some aspects of this time inconsistency have been formalised recently by Fahri and Tirole (2009) and Diamond and Rajan (2009).
generating instability. In turn, this in all probability means moving away from equilibrium settings and tackling disequilibrium explicitly. In many respects, all this takes us back to previous economic intellectual traditions that have been progressively abandoned in recent decades.

The financial cycle also raises first-order policy challenges. To take it fully into account, prudential, monetary and fiscal policies need to keep a firm focus on the medium term. The basic principle is to build up buffers during the financial boom so as to be able to draw them down during the bust, thereby stabilising the system. And if policy is unable to constrain the boom sufficiently and the financial bust generates a serious balance sheet recession, policies need to address balance sheet repair head-on. The overarching priority is to structure them so as to encourage and support the underlying balance sheet adjustment, rather than unwittingly delaying it. The priority is to prevent a stock problem from resulting in a persistent and serious flow problem, weighing down on expenditures and output. In turn, this means recognising the limitations of traditional fiscal expansion and of protracted and aggressive monetary easing. And, from a global perspective, it means recognising and internalising the unwelcome spillovers that policies can have, especially if financial cycles are not synchronised across countries. These challenges are especially tough for monetary policy, which risks being seriously overburdened and, over time, losing its effectiveness and credibility.

More generally, the emergence of the financial cycle as a major force highlights a growing tension between “economic” and “calendar” time. The financial cycle slows down economic time, as it stretches out the period over which the relevant economic phenomena play themselves out. Financial vulnerabilities take a long time to grow and the wounds they generate in the economic tissue a long time to heal. But the horizon of policymakers does not seem to have adjusted accordingly. If anything, it has shrunk in an attempt to respond to the high-frequency vagaries of the markets. This tension can be a major source of economic damage. The short horizons of market participants and policymakers contributed in no small measure to the financial crisis. They should not be allowed to generate the next one.

---

39 For an elaboration of this point, see Borio (2012). The notion of economic time dates back to at least Burns and Mitchell (1946), who take averages within business cycle phases. For a more formal recent treatment, see Stock (1987).
References


Bank of Canada (2011): Renewal of the inflation-control target – Background information, November.


——— (2010b): An assessment of the long-term economic impact of stronger capital and liquidity requirements, August.


FSB-BIS-IMF (2011): *Progress report from the FSB, IMF and BIS on macroprudential policy tools and frameworks*, report to the G20, October.


Okun, A (1962): “Potential GNP, its measurement and significance”, *Cowles Foundation*, Yale University.


