IMF Working Paper

Institute for Capacity Development and Research Department

Booms, Crises, and Recoveries: A New Paradigm of the Business Cycle and Its Policy Implications

Prepared by Valerie Cerra and Sweta C. Saxena

November 2017

Abstract

All types of recessions, on average, not just those associated with financial and political crises (as in Cerra and Saxena, AER 2008), lead to permanent output losses. These findings have far-reaching conceptual and policy implications. A new paradigm of the business cycle needs to account for shifts in trend output and the puzzling inconsistency of output dynamics with other cyclical components of production. The ‘output gap’ can be ill-conceived, poorly measured, and inconsistent over time. Persistent losses require more buffers and crisis-avoidance policies, affecting tradeoffs in prudential, macroeconomic, and reserve management policies. The frequency and depth of crises are key determinants of long-term growth and drive a new stylized model of economic development.

JEL Classification Numbers: E32, F41, O11, O47

Keywords: booms, crises, recoveries, business cycle, output gap, growth, development, fiscal policy, monetary policy, foreign reserves, macro-prudential policy, Global Financial Crisis

Authors’ E-Mail Addresses: vcerra@imf.org; ssaxena@imf.org

1 We are grateful to Charles Nelson for his continued encouragement of our work and Alejandro Werner for suggesting us to write this paper. We also appreciate comments from Martin Baily, Kimberly Beaton, Andy Berg, Helge Berger, Woon Gyu Choi, Rodrigo Cubero, Ted Gayer, Metodij Hadzi-Vaskov, Don Kohn, Ruy Lama, Daniel Leigh, David Lipton, Miko Mrkaic, Malhar Nabar, Maury Obstfeld, Jiri Podpiera, Tigran Poghasyan, Krishna Srinivasan, Thierry Tressel, Evan Tanner, Nour Tawk, Philip Turner, and seminar participants at the AEA Conference, Asia Pacific Economic Conference (Kolkata), Brookings Institute, International Monetary Fund, Research and Information Systems (Delhi), and the University of Washington.
Contents

Abstract ........................................................................................................................................... 1

I. Introduction .................................................................................................................................... 3

II. The New Business Cycle Paradigm .............................................................................................. 5

III. Unsustainable Booms and Financial Crises .............................................................................. 9

IV. Revisiting Business Cycle Methodology and Avoiding Spurious Cycles .................................. 11

V. Implications of New Paradigm for Policy Tradeoffs ................................................................. 16
   A. Financial regulation ......................................................................................................................... 16
   B. Monetary policy ............................................................................................................................. 17
   C. Foreign exchange reserve accumulation ......................................................................................... 18
   D. Fiscal policy .................................................................................................................................. 19

VI. Determinants of Economic Growth ........................................................................................... 20

VII. A Crisis-Driven Stylized Model of Economic Development ...................................................... 21

VIII. Conclusions ............................................................................................................................... 25

Tables
1: Strength of Recoveries .................................................................................................................. 7
2. Tests of Strong Boom Prior to Recession ....................................................................................... 10

Figures
1. U.S Real GDP Forecasts of Council of Economic Advisors ........................................................ 4
2. Impulse responses of output to crisis shock .................................................................................... 5
3. Concepts of the Business Cycle .................................................................................................... 6
4. Downward Revision in Potential GDP, U.S .................................................................................. 13
5. Filters Create Spurious Output Gaps .............................................................................................. 13
7. Estimates of Euro Area Output Gap and Growth by Vintages ................................................... 14
8. Revisions of Output Gap Estimates for Advanced Countries ....................................................... 14
9. Sweden versus OECD. GDP ........................................................................................................... 20
10: Absolute Convergence or Divergence .......................................................................................... 22
11. Frequency of Recessions ............................................................................................................... 23
12. Stylized Model of Economic Development .................................................................................. 24
13. Volatility versus Growth ............................................................................................................... 24

References ........................................................................................................................................... 26
I. INTRODUCTION

Weak economic activity and sluggish growth have persisted for nearly nine years after the Global Financial Crisis (GFC), puzzling most economic policymakers and academics. For example, in the Economic Report of the President for 2009, the Council of Economic Advisors (CEA) forecast a fast rebound of economic growth in the aftermath of the recession, with real GDP expected to rise by 5 percent in 2010 and 2011, nearly double the CEA’s estimate of long-run growth. In its Economic Projections and the Budget Outlook, the CEA reasoned: “…a key fact is that recessions are followed by rebounds. Indeed, if periods of lower-than-normal growth were not followed by periods of higher-than-normal growth, the unemployment rate would never return to normal.” Accordingly, the 2009 Report expected the level of real GDP to return to its pre-crisis trend level in 2011 (Figure 1).

Despite a deeper and more painful recession than expected, the CEA continued to forecast a strong rebound in economic growth, arguing: “The Administration forecast is based on the idea that real GDP fluctuates around a potential level that trends upward at a relatively steady rate.” … “The growth rate of the economy over the long run is determined by its supply side components.”

Scholars of the business cycle responded to these predictions in a vigorous blog debate on March 3, 2009. Greg Mankiw questioned the CEA’s prediction of a rapid recovery to trend real output based on the unit-root hypothesis (Campbell and Mankiw, QJE, 1987; Nelson and Plosser, JME, 1982). This work implies that an unexpected change in real GDP should change one’s forecast by at least the same amount over a long horizon. Brad Delong posted a rebuttal: “A fall in production that is accompanied by a big rise in the unemployment rate will in all likelihood be reversed.” … “And those post-recession periods of falling unemployment are also times of rapid output growth.” Paul Krugman also challenged Mankiw’s view: “For one thing is very clear: variables that measure the use of resources, like unemployment or capacity utilization, do NOT have unit roots: when unemployment is high, it tends to fall. And together with Okun’s law, this says that yes, it is right to expect high growth in future if the economy is depressed now. But to invoke the unit root thing to disparage growth forecasts now involves more than a bit of deliberate obtuseness. How can

---

2 February 28, 2009.


4 However, in his blog, Mankiw argued that if he knew the precise date the recession would end, his conditional forecast would be for higher-than-normal growth. His unconditional forecast, on the other hand, would be lower than the conditional one as it would take account of the possibility that the recession would linger.
you fail to acknowledge that there’s huge slack capacity in the economy right now? And yes, we can expect fast growth if and when that capacity comes back into use.”

With the benefit of almost nine years of hindsight, we know now that economic growth did not rebound quickly. Instead, post-crisis growth continued to disappoint, significantly underperforming predictions even as medium-term forecasts were repeatedly marked down. In 2014, the level of real GDP was significantly below its pre-crisis trend and levels forecasted by the CEA in 2009 (Figure 1). The recurring forecast errors made by a variety of institutions, including the IMF, demonstrate that the weak recovery was not widely expected, reflecting the prevailing conception of the business cycle espoused by DeLong and Krugman.

Post-crisis average growth in the United States and the world declined to 2.1 and 3.8 percent, respectively, during 2010-17, significantly lower than their pre-crisis averages of 2.7 and 4.5 percent during 2000-07. This sharp deceleration in growth rates has attracted many interpretations, including a productivity slowdown, demographic and labor force participation challenges, and debt-crisis legacies (e.g., IMF 2014; Gordon 2015; Fernald et al. 2017; Rogoff, 2015; Reinhart, Reinhart, and Rogoff 2012; among others). Summers (2013, 2014) revives the concept of secular stagnation as a chronic deficiency in demand, an associated decline in the equilibrium real interest rate, and the inefficacy of monetary policy due to the zero lower bound on nominal interest rates.

While this large persistent output loss has been unprecedented for the United States in the post-World War II era, the phenomenon of a prolonged and sluggish recovery, especially after a crisis, is not new. Cerra and Saxena (2005a and 2005b) demonstrate the permanent effects of a financial crisis for Sweden and Asian economies, respectively, while Cerra and Saxena (2008) establishes that balance of payments (BOP) and banking crises, on average, generate a permanent loss in the level of real output in a panel of 190 countries. The analysis shows that output falls relative to a baseline and remains permanently lower following
financial and political crises. On average, the magnitude of the persistent loss in output is about 5 percent for BOP crises, 10 percent for banking crises, and 15 percent for twin crises (Figure 2).

**Figure 2. Impulse responses of output to crisis shock**

![Figure 2. Impulse responses of output to crisis shock](image)


In this paper, we argue that the traditional paradigm of the business cycle should be replaced with a new paradigm of output dynamics that incorporates hysteresis along the lines shown in the evidence above. The new paradigm requires a rethink of core business cycle methods, the source of shocks, and the comovement of components of demand and supply. Given the similar pattern of growth dynamics following crises and recessions around the world as found in Cerra and Saxena (2008), any U.S.-centric explanation for the sluggish post-GFC recovery (including secular stagnation) is not plausible. The explanation must encompass the cross-country experience, including for emerging markets that sustain similar post-crisis output losses but rarely share the same trends that have been attributed to U.S. secular stagnation. We argue that the so-called “output gap” is ill-measured and ill-conceived and we discuss the policy implications for assessing the state of the economy and responding to economic slack. The new paradigm signifies that recessions, particularly those stemming from crises, are very costly in their economic impact and in the changes in the political landscape that ensue. This has implications for a range of policy tradeoffs, including the optimal aggressiveness of financial regulation, the appropriate mandates for monetary and fiscal policies, and the optimal level of foreign exchange reserves as insurance against the impact of shocks. We end with a discussion of how the new paradigm changes interpretations of long-term growth and present a new stylized model showing how economic development is impeded by crises.

**II. The New Business Cycle Paradigm**

The traditional view of the business cycle is depicted in the left-side graph of Figure 3. As elaborated above by the CEA, this view assumes that supply-side components generate a steady upward trend in real GDP, typically called “potential output” and demand-side shocks
lead to temporary fluctuations around that trend. Consequently, a recession consists of a temporary decline in output below its trend line, but a fast rebound of output back to its initial upward trend line during the recovery phase of the business cycle.

**Figure 3. Concepts of the Business Cycle**

![Diagram of Business Cycle]

Historical data do not support the traditional view. Cerra and Saxena (2005c) find that all types of recessions lead to a persistent loss in output, on average, not just recessions associated with financial crises. This decade-old result holds in an updated sample of data through 2013. Table 1 shows that the average growth rate during expansion years of a recovery is significantly lower than the average growth rate of all expansion years, contrary to the assumption of high growth in recovery espoused by the CEA, DeLong, and Krugman. For example, when a recovery begins (i.e., the first year of positive growth following a recession), growth is 0.8 percentage points lower than the average in all expansion years if the recession is associated with a banking crisis and 0.5 percentage points lower than in all expansion years if the recession is not associated with a banking crisis. Growth is 0.6 percentage points and 0.3 percentage points lower during the first four years of recovery for banking crisis and non-banking crisis-related recoveries, respectively.

The lack of recovery—or the sluggish rebound—from recession is a widespread phenomenon across countries and time periods. The significantly lower growth immediately after a recession holds for all income groups and regions (Cerra and Saxena, 2005c, Tables 7-8), except low income and African countries in a sample from 1960-2001.

This evidence thus points to a revised concept of the business cycle, as shown in the right-side graph of Figure 3. At the end of a recession, the recovery consists “at best” of a return

---

5 In a limited sample of 23 advanced economies, Blanchard, Cerutti, and Summers (2015) find that two-thirds of recessions are followed by lower trend output and half of these have lower growth relative to pre-recession output trend.

6 Ball (2014) illustrates the long-term impact of the GFC of 2008-09 on a set of 23 OECD countries, with the pattern appearing very similar to the right-side illustration of Figure 3.
of growth to its long-term expansion rate without a high-growth rebound back to the initial trend. The model of the business cycle needs to reflect observed output dynamics. Output does not cycle around a long-term upward trend. Instead, shocks result in a complete shift in the trend line itself. In short, the “business cycle” is not a cycle.

Table 1: Strength of Recoveries

<table>
<thead>
<tr>
<th>Troughs related to/Lags</th>
<th>-1</th>
<th>-1,-2</th>
<th>-1,-2,-3</th>
<th>-1,-2,-3,-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>All recessions</td>
<td>-0.53 ***</td>
<td>-0.40 ***</td>
<td>-0.49 ***</td>
<td>-0.40 ***</td>
</tr>
<tr>
<td>Banking recessions</td>
<td>-0.78 ***</td>
<td>-0.56 **</td>
<td>-0.67 **</td>
<td>-0.58 **</td>
</tr>
<tr>
<td>Non-banking recessions</td>
<td>-0.46 ***</td>
<td>-0.30 ***</td>
<td>-0.39 ***</td>
<td>-0.30 ***</td>
</tr>
<tr>
<td>Observations</td>
<td>4655</td>
<td>4523</td>
<td>4413</td>
<td>4280</td>
</tr>
<tr>
<td>Cross-sections</td>
<td>160</td>
<td>160</td>
<td>160</td>
<td>160</td>
</tr>
</tbody>
</table>


Note: We estimate the equation: $(g_i,t)/g_i,t > 0) = \alpha_i + \beta \text{Trough}_i,t-1 + \varepsilon_i,t$, where Trough is defined nonparametrically as a year of negative growth that is followed immediately by a year of positive growth. The table reports the value of $\beta$ which measures growth in the recovery year relative to other expansion years. Columns 2-4 include additional years after the trough. Recessions related to banking crises are those where the trough falls within a 5-year window of a banking crisis (t-1, t, t+1, t+2, t+3). Of the 584 troughs in the sample, 116 were related to banking crises. 16 percent of all observations are associated with negative growth rates. 

Data source: World Economic Outlook, International Monetary Fund, Laeven and Valencia (2012), and authors’ calculations.

How can the revised paradigm be explained in terms of supply versus demand shocks in driving output dynamics? The traditional view, exemplified by the statements from the CEA, would attribute a permanent fall in output to a decrease in aggregate supply and a temporary fall in output with a shock to aggregate demand. In a seminal paper, Blanchard and Quah (1989) impose a similar identification restriction in a vector autoregression of output and unemployment, where only supply disturbances have a long-run effect on output. Under this dichotomy, supply shocks would need to dominate cross-country experience to generate the output dynamics depicted in the right-side graph of Figure 3. The source of shocks is important to policymakers because conventional wisdom suggests that the temporary impact on output from a shortfall in demand could be ameliorated by expansionary macroeconomic policies, whereas structural reforms would be required to boost aggregate supply and the long-term level of output.
Real business cycle (RBC) theory assumes that economic fluctuations arise from productivity shocks, generating persistent changes to output that, in principle, could be consistent with the evolution of output in the right-side graph. However, this supply side explanation suffers from the criticism of implausibility that a country or the world suddenly loses productivity on a massive scale, with technology suddenly regressing and then remaining at a lower level to rationalize permanently lower output. In addition, the view of technology-driven fluctuations requires extreme assumptions for other economic behavior, such as the high intertemporal substitution of labor over the business cycle that is required to generate realistic employment variation (Stadler, 1994). Of course, supply shocks could encompass a variety of distortions that reduce productive efficiency, including firm-level misallocation of factors of production (Hsieh and Klenow, 2009).

Putting this debate in the perspective of the GFC, some scholars argue that productivity was slowing in the United States before the crisis (Fernald, 2014). But it would be unrealistic to argue that the recession in 2008-09 was due to a sudden loss of technological advancement. Instead, there is a broader consensus that the financial crisis reflected the U.S. housing bubble and associated leverage and led to a rise in uncertainty and various kinds of financial frictions (Krishnamurthy, 2010; Rogoff, 2015; Woodford, 2010). Bernanke (1983) argued along similar lines in explaining that the loss of banking infrastructure during the Great Depression made the financial system less effective in intermediating savings and investment. Should financial frictions and loss of intermediation be thought of as demand shocks or supply shocks? While a financial crisis may have some similarities to a technology shock, it is unlikely to be a pure supply shock. Many economic variables changed that reflected economic slack (unemployment rising, capacity utilization falling). Inflation also fell, although not much. Economic slack typically represents demand shocks.

More generally, do financial crises shift aggregate demand or aggregate supply? We find that, on average, banking crises lead to a fall in prices by about 4 percent in the long run in advanced countries relative to the baseline. A simultaneous fall in output and prices is consistent with the definition of a demand shock. However, the permanent loss in the level of output after a crisis raises the possibility that a shortfall in aggregate demand could eventually diminish aggregate supply (Yellen, 2016). Consider a low growth scenario, which discourages investment today (demand shock). The continuous shortfall in investment would reduce the capital stock in the long run (supply shock). In fact, IMF (2015) finds that business investment contracted sharply in advanced economies during the GFC, and the overriding factor holding it back has been the overall weakness of economic activity. Hence, reality is more complex than depicted by a traditional business cycle concept. Shocks having
initial attributes of shortfalls in aggregate demand, such as a rise in unemployment or a fall in inflation, can impact the supply side and permanently affect real GDP.\footnote{The idea that temporary shocks could produce a permanent macroeconomic impact was explored by Blanchard and Summers (1986) in the context of the rise in European unemployment in the 1970s. They attributed the hysteresis to labor market institutions such as labor union bargaining to support insiders. Ball (2009) argues that the natural rate of unemployment is affected by the actual rate of unemployment, which is in turn affected by shifts in aggregate demand.}

If an adverse demand shock can permanently depress the level of output, could expansionary macro policies during the recession help recover some of the lost output? Yellen (2016) notes that it may be worth running an aggressive “high-pressure economy” to reduce the depth and persistence of the downturn, thereby limiting the supply-side damage that might otherwise ensue. On the other hand, if there are deep hysteresis effects at work, operating through the real economy and financial sector, macroeconomic stimulus to support aggregate demand may be insufficient. Other policies such as structural reforms and measures to repair financial balance sheets may be required to recuperate any of the lost output.

In fact, Cerra, Panizza, and Saxena (2013) show that macroeconomic policies can influence the speed of economic recovery, helping to recover some of the lost ground from recessions and financial crises. Monetary and fiscal stimulus, real depreciation, foreign aid and more flexible exchange rate regimes can spur a rebound. In advanced countries suffering from recessions associated with banking crises, fiscal policy is particularly effective in boosting growth during the recovery.

The welfare gains from macroeconomic stabilization policies can be potentially very large, in contrast to conventional theory. Lucas (2003) argues that the benefit of using macroeconomic stabilization policies to eliminate “business cycle fluctuations” is extremely small. However, his calculations assume that output (and consumption) fluctuate around a trend and that macroeconomic stabilization policies do not affect the trend. The evidence presented above contradicts both assumptions. Crises and other negative shocks generate permanently lower output and consumption, generating substantially higher welfare costs. Stabilization policies can limit these welfare losses by supporting a rebound.

III. UNSUSTAINABLE BOOMS AND FINANCIAL CRISES

One potential explanation for the empirical regularity of sluggish recoveries relates to unsustainable booms leading up to recessions and crises, including the GFC. If excessively high growth pushes output above its sustainable trend in productive capacity, then the permanent loss in output associated with the recession may simply reflect a reversion to the sustainable trend.

\footnote{The idea that temporary shocks could produce a permanent macroeconomic impact was explored by Blanchard and Summers (1986) in the context of the rise in European unemployment in the 1970s. They attributed the hysteresis to labor market institutions such as labor union bargaining to support insiders. Ball (2009) argues that the natural rate of unemployment is affected by the actual rate of unemployment, which is in turn affected by shifts in aggregate demand.}
Using several decades of historical data for a large panel of at least 160 countries, we tested whether GDP growth tends to be abnormally high before recessions, banking crises, and currency crises. We compared the rate of growth during the 3 years up to and including the peak year, with each country’s historical average rate of growth during expansion years (Table 2). We find that GDP growth is lower just before the peak of a business cycle and just before a banking or currency crisis than in an average expansion year. In fact, growth is significantly lower before a crisis. This finding refutes the argument that unsustainable booms are responsible for the permanent output loss in the subsequent recessions and crises.

Table 2. Tests of Strong Boom Prior to Recession

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>3 years prior to recession</th>
<th>3 years prior to banking crises</th>
<th>3 years prior to currency crises</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.36***</td>
<td>-0.84***</td>
<td>-0.28**</td>
</tr>
<tr>
<td></td>
<td>-0.41***</td>
<td>-0.42**</td>
<td>0.10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No of Observations</th>
<th>6854</th>
<th>5208</th>
<th>5474</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Effects</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

Note: The numbers reported in the table compare the growth in expansion years with the growth in the 3 years prior to the recession, banking crises, and currency crises. These episodes eliminate any recessions occurring within the 3-year window.

Data source: World Economic Outlook, International Monetary Fund, Laeven and Valencia (2012), and authors’ calculations.

The evidence also does not support the hypothesis that crises reflect realizations of a declining trend: Cerra and Saxena (2008) demonstrate that growth forecasts prior to crises have been optimistic and expectations have been sluggish to adjust downward.

Financial and economic imbalances may be building before a financial crisis even while GDP grows at normal rates. Some financial variables, including credit growth and house prices, tend to rise strongly before financial crises (Kaminsky and Reinhart, 1999; Cerutti, Dagher, and Dell’Ariccia, 2015). Borio and others (2015) find that credit booms create resource misallocation, which leaves scarring effects on the economy after the bust. These studies imply that there is information in economic and financial variables that could shed light on the sources and transmission of shocks. However, while unsustainable booms may occur in some sectors of the economy, our results show that these financial imbalances do not manifest themselves in above-normal overall GDP growth. In short, we find no evidence of

---

8 Dell’Ariccia and others (2012) find that growth is higher during credit booms relative to non-boom periods, but their non-boom periods include years of recession. They also find that only one-third of credit booms end in financial crises; others are followed by extended periods of below-trend economic growth or, conversely, by financial deepening that supports long-term growth. The lack of strong growth before a crisis may be less

(continued…)
stronger output growth leading up to the crisis or recession. Thus, the puzzle of lack of reversion of output to its pre-crisis trend is complemented by an additional puzzle of the lack of strong output growth leading into the crisis.

IV. REVISITING BUSINESS CYCLE METHODOLOGY AND AVOIDING SPURIOUS CYCLES

RBC practitioners often detrend output series, even though the theory sees output as driven by technology shocks that would typically be related to the supply side. RBC economists compare the deviations from trend in the actual historical data and the simulated data from their model to argue the validity of their model. They typically match second moments—the variability and co-movements between output and other variables.

The revised conception of the business cycle implies that the RBC methodology needs to be improved or abandoned. First, it melds together fluctuations caused by different disturbances and obscures the propagation dynamics. A more appropriate methodology would be to examine the time path of output and other pertinent variables in response to a specific shock. Comin and Gertler (2006) emphasize that conventional detrending methods used in business cycle analysis sweep longer-term oscillations into the trend, thereby removing them from analysis. Incorporating longer frequency oscillations, they show that business cycles are more persistent. Second, given that output dynamics can be characterized by persistent shifts in the trend, the RBC practice of filtering output generates spurious cycles (Cogley and Nason, 1993; Canova, 1998). Detrending the actual and simulated data in the same spurious manner may also create relationships that do not exist in the raw data.

The practice of decomposing output into trends and cycles pervades the economic literature, irrespective of the school of thought. The ‘output gap’ is a standard representation of the cycle, defined as the difference between the actual output and the potential output of a country. ‘Potential output’ signifies the economy’s capacity to produce (i.e., the supply side) and is typically viewed as evolving smoothly based on technology, labor, and capital, as exemplified by the quotes from the CEA. Consequently, under the traditional paradigm of the business cycle, the output gap can be thought of as the temporary deviation of actual output around potential output corresponding to demand shocks. However, as explained above, the data do not support the traditional paradigm of temporary fluctuations. Instead, even demand shocks can have a permanent impact on output. If output movements are permanent and cycles do not exist, then we need to reconsider the meaning and measurement of the output gap.

*surprising than at first glance given that many papers use the strong growth of credit relative to GDP growth as a predictor of banking crises (Demirgüç-Kunt, Asli, and Enrica Detragiache, 1998; Gourinchas, Valdes, and Landerretche, 2001; and Honohan, 1997).*
The output gap is also a common feature of analysis in policy institutions around the world. Many central banks base their decisions for setting the policy interest rate on the deviation of expected inflation from its target and a measure of the output gap as an indicator of economic slack. Similarly, fiscal policy analysis includes measuring cyclically-adjusted fiscal positions based on the output gap to ascertain the level of sustainable debt. Often, the design of structural policies also incorporates estimates of the output gap (IMF 2016b).9

An implication of the new stylized fact on the permanent impact of crises and recessions on output is that the output gap may be extremely difficult to measure and more difficult to interpret, especially around severe recessions when a proper assessment is crucial. Indeed, the output gap may not be defined well if trend and actual output move together. For example, Figure 4, copied from the Congressional Budget Office, shows that the estimated path of potential GDP has been revised down over time after U.S. output fell sharply during the crisis. Moreover, the output gap has been closing in recent years, but this has happened through potential GDP coming down to actual GDP, not the other way around as normally assumed.

In fact, this recent pattern of continued downward revisions of potential GDP and potential growth has contributed to the debate on secular stagnation. But these revisions may be partly just an artifact of measurement issues. Figure 5 shows the estimate of potential GDP when we apply a statistical smoothing method to output data a few years after a recession. The filter artificially constructs a positive output gap before the downturn and a negative one afterward, even if such gaps may not exist. This faulty measurement of the output gap could mislead policymakers, and it occurs precisely at the time when a good calibration of the output gap is most needed.

---

9 Fiscal structural reforms in the labor market area, such as reduced labor tax wedges and increases public spending on ALMPs, have larger effect during periods of economic slack, in part because they usually entail some degree of fiscal stimulus. In contrast, reforms to employment protection arrangements and unemployment benefit systems have positive effects in good times, but can become contractionary in periods of slack because they weaken aggregate demand.
The problem with measuring and defining the output gap is not just hypothetical. In the April 2008 World Economic Outlook (WEO), output gaps in the United States and the euro area were measured to be broadly closed, as shown in the blue dotted lines (Figure 6). But the view of history has since changed. When output data after the crisis is added to the sample, the estimated path of potential GDP is pulled down and the output gap at the peak becomes positive, as shown by the red lines for the U.S. and the euro area using the October 2015 WEO database. Ex-post, the pre-peak period appears as if it were overheating. In fact, the estimate of the 2008 output gap in the euro area continued to increase, even though the 2008 growth rate turned out much lower than originally projected (Figure 7).
Stark revisions of the output gap over time are pervasive. In a sample of 18 advanced countries, current WEO estimates of the 2007 output gaps are all higher than the estimates made at the time (Figure 8). This artificial construction of pre-crisis output gaps is also prevalent in literature studying the economic conditions and predictors of crises. For example, Gourinchas and Obstfeld (2012) show positive output gaps before some types of crises, including the GFC. However, their application of an HP filter to ex-post data leads to the same bias in inference as described above.

The inability to estimate a stable output gap is not limited to the use of an HP filter. OECD estimates based on the production function approach display similar revisions over time (Borio, Disyatat, and Juselius, 2013). More sophisticated multivariate models also lack robustness in real time versus ex-post estimates. For example, Benes et al. (2010) find substantial uncertainty in their estimates of potential output for 10 countries using a multivariate filter (inflation, unemployment, and capacity utilization). If additional information is used (e.g., growth expectations), the end of the period problem is reduced slightly, although significant uncertainty still surrounds the estimates, particularly around turning points in the business cycle (Blagrave et al. 2015). Blanchard, Dell’Ariccia, and Mauro (2010) discuss the breakdown of the relationship between inflation and potential output during the Great Moderation. Borio, Disyatat, and Juselius (2014) show that structural models that impose a Phillips curve relationship lead to specification errors and biases, especially due to the low-frequency trend in inflation. Borio, Disyatat, and Juselius (2013) augment the HP filter to include financial variables in the estimation of the output gap and

---

10 Cerra and Saxena (2000) estimate potential output for Sweden using a range of different methods. Berger et al. (2015) discuss various approaches and apply the multivariate filter to some Euro area countries.

11 Many filters are “one-sided” at the beginning and end of a time series. Estimates of trend output are therefore overly sensitive to the actual data at the end of the sample.
show that this method provides more stable estimates over time. However, their model does not allow financial factors to influence potential output and their large negative output gap estimates post-GFC suffer from the same failure to recognize the permanent output loss.

Although the output gap is used extensively in policy analysis, its concept and measurement should depend on the purpose for which it is intended. Thus, the output gap employed for monetary policy should be different from the one employed for fiscal policy. Monetary policymakers need to assess the slack in the economy and the signs of pressure for inflation or deflation. Given that the standard definition of potential output and the output gap focuses on inflation stability and the full utilization of resources, the standard measure of the output gap may still be useful for a monetary policy focused on inflation stabilization. During a recession, resource utilization and inflation typically fall. The Okun’s law relationship prevails, with unemployment rising as GDP growth declines. Even so, there is a puzzle as to why Okun’s law is not as robust in the recovery. GDP growth does not accelerate beyond a normal expansion rate while the post-recession/post-crisis unemployment rate falls. The negative output gap in the recession—with falling inflation and rising unemployment—do not translate into any useful prediction for actual output dynamics in the recovery.

The concept of the output gap used for monetary policy purposes would be inappropriate for determining public debt sustainability. The reason is that tax revenues tend to move with actual GDP. So, a persistent decline in the level of output also implies a persistent decline in the tax base. With permanently lower tax revenue, fiscal retrenchment through lower spending would be required to preserve debt sustainability. If instead, the fiscal authorities use the standard output gap (based on inflation pressure and other measures of slack) to calculate structural revenues after a crisis, they might incorrectly and dangerously assume no need for fiscal retrenchment, leading to a debt bias. Why? Because the output gap calculated by the standard method is likely to be negative after the crisis, due to falling inflation and rising unemployment. The fiscal authorities would be misled into believing that structural revenues are higher than current actual revenues and that revenues would rebound during the recovery as the slack disappears. But as discussed above, the historical evidence shows that actual GDP (i.e., the tax base) does not revert up to the prior trend, even when the unemployment rate declines and resource utilization rises. Given this, the output gap for fiscal purposes should be calculated as zero after a crisis and structural fiscal revenue would be the same as actual fiscal revenue.\(^\text{12}\)

\(^\text{12}\) Borio, Disyatat, and Juselius (2013) argue that their “finance neutral” output gap is better for cyclically adjusting budget balances because the estimates were positive before the GFC, which warned of unsustainably high output. However, their post-crisis output gap estimates are large negative values that suggest actual output below trend output, and imply a rebound in output that never happened. As with an output gap from a naïve HP filter or one based on declining inflation, the negative output gap estimates inappropriately suggest more fiscal space than existed.
V. Implications of New Paradigm for Policy Tradeoffs

The large and persistent output losses from financial crises and deep recessions depress employment and household income, which often leads to populist pressure for policy changes that inadvertently further reduce growth and social welfare. For example, Mian et al. (2014) document an increase in political polarization after systemic financial crises, including an increase in voting shares earned by far-right political parties (Funke, et al., 2016).

Another risk from sluggish recoveries is the backlash against globalization. The longer-term outlook remains disappointing as growth going forward is a full percentage point lower for advanced economies compared to their historical average (1990-2007). The resulting weakness in the global outlook can further lower incentives to invest and slow trade, which will, in turn, weigh on underlying potential growth. Moreover, in a negative feedback loop, trade protectionism in many countries has risen since the GFC, with potentially adverse consequences for growth and productivity.

The high cost of financial crises and recessions raises the need for prudent macroeconomic and financial policies to reduce the risk of such volatility. We focus on three crucial policy trade-offs. In particular, avoiding crises and large recessions would tilt the balance in favor of more financial regulation, financial stability to be included as a consideration for monetary policy, building a larger war chest of foreign reserves, and maintaining a conservative fiscal stance during booms.

A. Financial regulation

Financial policy involves a tradeoff. A vibrant financial sector generates intermediation between savers and investors and assures that funds are available for risk-taking in projects that have a probable but uncertain return. So, regulation shouldn’t be so excessive that it wipes out all innovation and risk-taking. Indeed, Ranciere, Tornell, and Westermann (2006) find that financial liberalization improves long-term growth on balance, even when taking into account the indirect negative impact of higher probability of financial crises.13 On the other hand, if regulation is too limited, it can misalign incentives toward too much financial risk-taking, especially if adverse outcomes accrue and spread throughout the financial sector and the economy to bystanders.

13 However, they only consider twin crises (banking and BOP). In addition, their equity market liberalization variable may have limited impact on generating a twin crisis and their other liberalization proxy, actual capital inflows, may be endogenous with higher growth. Moreover, crises have other determinants than just liberalization, and financial regulation has many facets.
Prior to the GFC, the view of many, including Greenspan (1997), was to limit regulation to foster financial innovation. However, the limited perimeter of regulation gave incentives for banks to create off-balance-sheet entities to avoid some prudential rules and increase leverage (Blanchard, Dell’Ariccia, and Mauro 2010). Adrian and Obstfeld (2017) note that financial regulations were put in place in the aftermath of the great recession, but there have been subsequent calls to relax them on the grounds that they go too far and hurt the economy by reducing financial institutions’ profits and their ability to provide essential services. While regulation can raise the cost of credit for households and businesses or reduce market liquidity, Adrian and Obstfeld (2017) argue that these unintended consequences are relatively small compared to the large benefits of a safer financial system. There is also a somber realization that regulation must adapt to the constantly innovating financial markets (Caruana, 2014; Turner, 2017).\footnote{Caruana (2014) notes that at a meeting of bankers, the suggestions to deal with financial market risks included establishing thicker buffers, setting suitable incentives and penalties for banks, and adopting more proactive, intrusive supervision.} In short, regulation needs to weigh the benefits of innovation and prudent risk-taking against the externalities of excessive risk-taking with the latter including the high costs borne by many bystanders.

Another key lesson is to avoid politically-driven cycles in regulation and deregulation that exacerbate financial cycles. As demonstrated by the experience of the past two decades, it is politically difficult to impose regulation in a fast-growing economy. But regulation would be most useful during credit and asset price booms to curb financial excesses and the build-up of vulnerabilities in the financial sector. In contrast, public outcry in the aftermath of crises may generate heavy-handed regulation that undermines any fragile recovery by further reducing access to credit.

### B. Monetary policy

Another policy tradeoff relates to the mandate of monetary policy. In recent decades, price stability has become the principle objective of monetary policy, especially for central banks with explicit inflation targeting frameworks. In addition, some central banks, such as the U.S. Federal Reserve, treat economic stability as an explicit or implicit objective. Given the substantial real economic costs associated with financial crises, central banks would need to incorporate financial stability in their policy analysis and decisions, even if only to safeguard economic stability. The persistent economic costs following the GFC demonstrate that financial stability needs to be a component of the central bank mandate even in advanced countries with modern financial and economic institutions.

Central banks also need to consider the tradeoffs and effectiveness of using different tools to ensure financial stability and address asset price booms and busts. Ideally, central banks
could reduce the probability of bubbles and crises through regulation and prudential supervision. However, to the extent that regulation and supervision are insufficient, central banks may need to use monetary policy tools to reduce the incidence of financial crises as well as to stabilize inflation and output directly. Indeed, if overly accommodative monetary policy—responding to low inflation risk—contributes to the boom and bust cycles that end in financial crises and output losses, then central banks would certainly need to incorporate financial cycles in their policy modeling and analysis.

Economists have also debated the timing of a monetary policy intervention. Blanchard, Dell’Ariccia, and Mauro (2010) argue that the Federal Reserve successfully responded to the 1987 stock market crash, the Long-Term Capital Management (LTCM) collapse, and the bursting of the tech bubble. These successes with ex-post intervention reinforced the view prevailing before the GFC that monetary policy would be adequately equipped to handle the financial consequences of asset prices busts. However, the inability of the Fed and other central banks to adequately address the aftermath of the Great Recession questions the philosophy of “benign neglect” of financial bubbles. Svensson (2016) argues that leaning against the wind (keeping interest rates higher than required for price stability to avoid a financial bubble) may increase the cost of a crisis by weakening the economy and increasing unemployment. IMF (2015) maintains that monetary policy should not be altered to contain financial stability risks and that the case for leaning against the wind is limited. However, Gourio, Kashyap, and Sim (2016) embed the risk and cost of a financial crisis underpinned by “excessive credit” into a New Keynesian DSGE model and show that leaning against the wind is welfare improving. This also means accepting higher volatility of inflation and output in exchange for reducing the risk of a financial crisis. Likewise, Christiano et al (2010) argue that stock market booms have historically coincided with low inflation and strong credit growth. They show that monetary policy should include an independent role for credit growth to reduce output and asset price volatility. All in all, the new paradigm and the corroborating post-GFC experience tip the balance of the debate in favor of complementing prudential tools with monetary policy tools and to respond preemptively to avoid financial crises and severe recessions.

C. Foreign exchange reserve accumulation

The third macroeconomic policy that needs to be reconsidered in light of the new paradigm is the appropriate amount of foreign exchange reserve accumulation. A large literature discusses the optimal level of reserves. One of the key objectives of accumulating foreign exchange is to serve as a buffer against shocks that would require foreign payments. So, the foreign reserves can be thought of as insurance, with the costs of accumulation offset by their benefits in the event of a sudden stop or a balance of payments crisis. The central bank can draw upon its reserves: i) to increase liquidity when relying on sound macroeconomic policies isn’t enough in the face of contagion (Feldstein, 1999); ii) to avoid costly liquidation
of long-term projects when the economy is susceptible to sudden stops (Aizenman and Lee, 2005); and iii) to manage financial instability as well as exchange rates in a world of increasing financial globalization (Obstfeld, Shambaugh, and Taylor, 2008).\textsuperscript{15}

Policy work on the appropriate level of foreign exchange reserves for crisis prevention is based on a model developed by Jeanne and Ranciere (2011). In their model, foreign reserves serve as insurance against consumption losses during the period of a \textit{temporary} \textit{BOP} crisis. A key assumption of the model is that the temporary BOP crisis generates a temporary loss in output and consumption. However, optimal reserve coverage metrics developed from this model are misspecified given that the empirical evidence contradicts the assumption of \textit{temporary} output and consumption losses. Instead, the permanent output and consumption losses that characterize the new paradigm require a significantly larger level of foreign exchange reserve buffers as insurance. Given the high output and social costs of crises, the insurance premium can also pay for itself by reducing the probability of a financial crisis (Rodrik, 2006).\textsuperscript{16} Conversely, permanent losses argue for immediate adjustment of consumption. In the same vein, Cerra, Panizza, and Saxena (2013) suggest that allowing exchange rate depreciation could help support a rebound in economic activity, limiting the permanent losses.

D. Fiscal policy

The pattern of output dynamics through booms, crises, and recoveries needs to be incorporated into fiscal policy decisions and the design of fiscal rules. During economic expansions associated with strong credit growth and asset price booms, financial vulnerabilities may be building even if they do not generate high output growth or inflation. These booms in particular raise the possibility of a subsequent financial crisis that will have scarring effects on output and the tax base, which will deteriorate the post-crisis fiscal balance, holding constant the policy-determined expenditure plans. Thus, policy makers should maintain conservative fiscal positions during the boom phase, both to avoid intensifying any financial bubble and to preserve fiscal space that might be required in the aftermath of a potential crisis.

The fiscal policy response to a crisis or large recession, in contrast, would need to weigh several considerations. The permanent decline in fiscal revenue would need to be matched by expenditure restraint to avoid a deterioration of debt dynamics. Policymakers would need to

\textsuperscript{15} This insurance is for central banks to deal with both sudden stops as well as sudden flights in face of open capital accounts.

\textsuperscript{16} Bilateral swaps and some financing arrangements were introduced in the aftermath of the GFC, but the terms and access have been uncertain and the arrangements have been intended only as temporary support.
assess whether to adjust immediately to the permanent shock or to smooth the transition to it. On the other hand, if fiscal space is available, a temporary fiscal expansion could help support a recovery (Cerra, Panizza, and Saxena, 2013) and at least partly reduce the scarring effect of the crisis/recession.

VI. DETERMINANTS OF ECONOMIC GROWTH

The new paradigm implies that long-term economic growth is affected by financial crises and other adverse shocks. Thus, it is critical to account for the impact of crises on output to avoid misinterpreting the role of policies and other determinants of growth. Cerra and Saxena (2005a) describe a cautionary example. In particular, Sweden was ranked 3rd of OECD countries in per capita GDP in 1970, but fell to 14th place in the 1991 rankings. Some Swedish economists argued that this relative decline was caused by the fiscal and social policies of the Swedish welfare state model. However, the time series in Figure 9 shows that the decline in Sweden’s ranking occurred precisely during its financial crisis in 1991-92. In contrast, an abrupt decline in Sweden’s relative per capita GDP is not consistent with a story of distortions and disincentives related to welfare state policies, as any such distortions would have been ongoing and would have been associated with consistently lower growth year after year. Thus, a lesson from this example is to examine the time path of growth rather than merely the long-run average growth rate to better attribute economic policies and conditions as determinants of growth.

Figure 9. Sweden versus OECD. GDP per capita
Economic development is also impeded by crises and adverse shocks, which conventional development theory does not suitably incorporate. The conventional neoclassical model of economic development would predict that poor countries, which are those with lower capital stocks, would grow faster than rich countries because of diminishing returns to capital. In addition, according to this theory, an adverse shock that lowers a country’s capital stock would lead to even faster growth because the increase in the return on capital would spur investment. The neoclassical theory thus implies convergence in per capita incomes.

However, the development literature has found that empirical evidence contradicts the prediction of such convergence (Barro, 1991). The top left graph in Figure 10 shows that countries starting out rich in 1960 grew faster between 1960 and 2000 than those starting out poor. This contradiction from neoclassical theory has been termed ‘absolute divergence.’ The theory, nonetheless, allows for rich countries to grow faster than poor countries if they are converging to a higher steady state (Barro and Sala-i-Martin, 2004, Figure 1.10), a notion termed ‘conditional convergence.’ In this case, the rich country would be farther from its (higher) steady state, where steady state positions in turn depend on economic characteristics such as population growth, saving rates, and government policies. Under conditional convergence, growth in per capita GDP depends on the gap between the initial and the steady state capital stock. A less noticed implication of this theory is that the growth of the rich country should be consistently higher than that of the poor country as both converge to their respective steady states.

The pattern of growth over time is not consistent with the conditional convergence explanation. In the second and third row of Figure 10, we separate each country’s average long-run growth rate into its average growth during expansions (i.e., positive growth) and its average during recessions (i.e., negative annual growth). We find that poor countries grow faster than rich countries during their expansions. Thus, there is absolute convergence during expansions, which would be consistent with poor countries being farther from their steady states.
Absolute divergence during 1960-2000 derives instead from the fact that poor countries’ recessions had been much deeper than those of rich countries (Figure 10; bottom left chart) and poor countries had more frequent recessions (Figure 1, left chart). Thus, adverse shocks and crises leading to recessions obstructed poor countries’ growth, rather than differences in their gaps to steady state.
These findings combined with our previous evidence of permanent output loss during recessions lead us to propose a new stylized model of economic development (Figure 12). In line with neoclassical growth theory, economic expansions in poor countries are stronger than in rich countries, which in the absence of negative shocks could over time lead to absolute convergence in the levels of output per capita. Both sets of countries have negative shocks with persistent impacts. However, recessions in rich countries are shallow and infrequent compared to those in poor countries. Volatility leads to lower growth (Ramey and Ramey, 1995; Cerra and Saxena, 2005). Indeed, there is a strong negative relationship between the proportion of recession years and the average growth rate (Figure 13). Poor countries’ policies and institutions likely still contribute to underdevelopment, but the mechanism occurs mainly through generating frequent and large crises and recessions which commonly set poor countries back from their development path.

However, the past decade has been different (right hand side in Figure 10). From 2001-2011, poor countries grew faster than rich countries. As before, they grew faster during expansions. But this time, poor and rich countries had roughly similar losses during recessions. In addition, although poor countries had more frequent recessions during the period 1960-2000, rich countries had a slightly higher frequency of recession years during 2001-2011 (Figure 11). The main reason for the difference during the last decade is that the rich countries experienced financial and debt crises leading to severe recessions that set them back from their previous trends. In contrast, lower income and emerging market countries had developed stronger policy frameworks and built policy space to respond to shocks (IMF 2012) although their policy buffers have since partly eroded (IMF 2016a).

---

17 Cerra and Saxena (2008) show that negative political shocks, such as weaker constraints on executive power, also lead to permanent output loss in low and lower-middle income countries.
Figure 12. Stylized Model of Economic Development

Figure 13. Volatility versus Growth
VIII. Conclusions

Our work has shown that crises and recessions tend to permanently push down the level of a country’s output. As argued throughout the paper, this fact has powerful implications for our concept of the business cycle and for macroeconomic and prudential policies to avoid and react to crises and recessions. The cross-country evidence and the U.S. experience following the GFC demonstrate that the business cycle is not a cycle. This stylized fact presents a new puzzle: mean reversion in the unemployment rate and capacity utilization does not translate into trend reversion in output. Understanding the mechanism that underlies this apparent inconsistency sets an important agenda for future research. We need to eschew reasons that focus on the specific characteristics of the U.S., and find explanations that are robust to the similar experiences of many countries. In addition, the traditional distinction between supply and demand shocks and the assumption that demand shocks have only a transitory economic impact need to be revisited. If output does not exhibit cycles, the rationale for constructing an ‘output gap’ as a measure of the cyclical position is not well justified and could generate misleading signals. Instead, policy analysis should focus on indicators that better measure the purpose for which they are intended—such as direct measure of price pressure for monetary policy and measures of the permanent level of revenue for fiscal policy.

Economic policies should be geared toward avoiding crises and severe recessions and responding with appropriate stimulus and safety nets. Financial regulation that contains excessive risk-taking and sustainable macroeconomic policies constitute the first best options. If these policies are insufficient, monetary policy may need to play a role in addressing financial stability risks. Where policy space permits, economic stimulus may also be helpful in the aftermath of a crisis or severe recession. Foreign exchange reserves can help insure against losses due to external payment crises. However, previous estimates of the optimal level of reserves assumed temporary losses. New estimates are required in light of the evidence that the losses associated with crises are persistent.

Beyond short-run output dynamics and stabilization policies, our analysis also points to a new model of long-term economic development. Poor countries have had more frequent crises and deeper recessions than rich countries over most of the last several decades. This volatility and its persistent adverse impact on output has held back poor countries’ development irrespective of any factors that support development as in the canonical model of gradual capital accumulation.
References:


Blanchard, Olivier, Giovanni Dell’Ariccia and Paolo Mauro, 2010. “Rethinking Macroeconomic Policy,” IMF Staff Discussion Note SDN/10/03.


Dell’Ariccia, Giovanni, Deniz Igan, Luc Laeven, and Hui Tong, 2012. “Policies for Macrofinancial Stability: How to Deal with Credit Boom?” *IMF Staff Discussion Note SDN/12/06*.


