Asset Bubbles: Re-thinking Policy for the Age of Asset Management

Brad Jones
Abstract

In distilling a vast literature spanning the rational—irrational divide, this paper offers reflections on why asset bubbles continue to threaten economic stability despite financial markets becoming more informationally-efficient, more complete, and more heavily influenced by sophisticated (i.e. presumably rational) institutional investors. Candidate explanations for bubble persistence—such as limits to learning, frictional limits to arbitrage, and behavioral errors—seem unsatisfactory as they are inconsistent with the aforementioned trends impacting global capital markets. In lieu of the short-term nature of the asset owner—manager relationship, and the momentum bias inherent in financial benchmarks, I argue that the business risk of asset managers acts as strong motivation for institutional herding and ‘rational bubble-riding.’ Two key policy implications follow. First, procyclicality could intensify as institutional assets under management continue to grow. Second, remedial policies should extend beyond the standard suite of macroprudential and monetary measures to include time-invariant policies targeted at the cause (not just symptom) of the problem. Prominent among these should be reforms addressing principal-agent contract design and the implementation of financial benchmarks.

JEL Classification Numbers: E44, E58, G01, G02, G12, G14, G15, G18

Keywords: Asset bubbles, Financial stability, Macroprudential policy, Monetary policy.

Author’s E-Mail Address: BJones@imf.org

1 The author is grateful for comments from Tamim Bayoumi, Olivier Blanchard, Pete Dattels, Luc Everaert, Gaston Gelos, Shuntaro Hara, Matthew Jones, Tommaso Mancini-Griffoli, and Miguel Savastano.
Contents

Abstract ........................................................................................................................................2

I. Introduction .................................................................................................................................4

II. The ‘Clean vs. Lean’ Debate: A Survey .....................................................................................8

III. Theories of (In)Efficient Markets and Speculative Bubbles..................................................15
   A. Bubbles and the (In)Efficiency of Markets – A Review ..........................................................15
   B. Competing Models of Bubble Formation and Persistence ....................................................19

IV. Policy Implications ..................................................................................................................31

V. Concluding Remarks and Future Research .............................................................................37

Tables
Table 1. Dimensions of the Traditional ‘Clean vs. Lean’ Debate ...................................................8
Table 2. Stylized Summary of Asset Pricing/Bubble Models ..........................................................20
Table 3. Mapping Policy Responses to Bubble Models .................................................................31

Figures
Figure 1. Worldwide Financial Assets and Institutional Assets ......................................................6
Figure 2. Bank Assets vs. Investment Firm Assets under Management .........................................6
Figure 3. Benchmark Decomposition of Hedge Fund Returns .......................................................25
Figure 4. Subjective vs. Objective Expected Returns .....................................................................30
Figure 5. Relative 10-year Annualized Outperformance of Fundamental-based Indices ..........34
I. INTRODUCTION

Though booms, bubbles and busts have littered the financial landscape for centuries, the devastating impact of the global financial crisis has ignited a vigorous policy debate over how best to address periods of excessive euphoria in asset markets. Notwithstanding Samuelson’s (1966) famous epigram that “the stock market has forecast nine of the last five recessions,” there is now broad agreement with the early analysis of Fisher (1933), Keynes (1936) and Kindleberger (1978) that large swings in asset markets are of great consequence to the real economy. There is, however, much less consensus over what to do about them.

The transmission mechanism between asset markets and the economy is generally well understood. Asset booms that compress risk premiums to abnormally low levels can lead to a misallocation of resources; induce excessive consumption, investment and debt accumulation; distort derivative markets commonly used to manage risk; and give an impression of false health for private and public sector balance sheets. As boom gives way to bust, the ensuing chain reaction can include: a negative wealth effect (stemming from a hit to the balance sheet and confidence of households and firms) which results in sharply reduced aggregate demand; a second round credit crunch, as banks respond to the decline in the value of collateral held on their balance sheet by tightening lending standards; and possibly sovereign-banking sector contagion, where the health of the sovereign balance sheet is impaired due to a collapse in tax revenues and bank recapitalization costs, with feedback effects for banks which are exposed to sovereigns in their asset holdings and as their funding costs derive from the sovereign borrowing rate. In extreme cases, severe asset price busts can also give rise to longer-term issues of intertemporal equity as future taxpayers become liable for ‘crimes they did not commit.’ More generally, the global financial crisis demonstrated that the adverse effects of financial disruptions on economic activity can be far worse than previously anticipated; with 93 countries (and most advanced economies) recording outright declines in real GDP in 2009, globally synchronous busts are especially virulent.

---

2 Models addressing the misallocation of resources stemming from asset bubbles date back to Diamond (1965).

3 Baker (2009) and Chirinko and Schaller (2012) are recent studies showing how bubbles lower the cost of capital and discount rate used in the evaluation of new projects, thus fuelling excessive investment spending.

4 See most recently, Case and others (2005, 2013), and Carroll and others (2011).


6 Reinhart and Rogoff (2009a).

7 Mishkin (2011), Bluedorn and others (2013), and Reinhart and Rogoff (2013).
As global capital markets continue to expand in scale and scope, there are further reasons to expect asset price movements to attract the close attention of policy makers. The stock of globally traded financial assets (i.e. excluding real estate) has expanded sharply over recent decades, from US$7 trillion (71 percent of world GDP) to US$163 trillion (226 percent of world GDP) between 1980 and 2012. An extrapolation of per capita GDP trends suggests financial assets could reach $400 trillion by 2050 (Haldane, 2014). One implication of the rising importance of capital markets is that wealth and financial accelerator effects may become more important than in the past, fueled in part by institutional innovations that make it easier to monetize asset price gains through increased consumption and borrowing.

Another implication for financial stability central to the present analysis is that the behavior of agents (institutional investors) managing capital on behalf of principals (asset owners) could be a source of friction demanding much closer scrutiny in the years ahead—particularly if incentives in the asset management industry induce procyclical behavior. Indeed, not only is the stock of global financial assets rising rapidly, but so too is the share of those assets managed by institutional investors. Having increased from 30 to 39 percent between 2004 and 2012, some estimates suggests it could rise to 43 percent by 2020, totaling US$101 trillion (Figure 1, left panel). Around 80 percent of institutional assets are managed through mutual fund vehicles (Figure 1, right panel). Demographic trends (where the pool of global savers has become larger, older and richer), combined with increasingly liberalized financial systems, has proven a boon for the asset management industry. The trend is likely to continue as the process of capital market deepening, which is strongly linked to per capita income, is now firmly underway across the emerging markets. The fastest growing economies globally are those where holdings of financial assets and the penetration of the asset management industry is typically lowest. For these reasons we may well be entering the ‘Age of Asset Management’ (Haldane, 2014). It should be noted that the value of assets under management of the largest global investment firms is now approaching the value of assets held by the largest global banks (Figure 2, left hand panel). Since 2008, they have also

---

8 In a policy making context, increased recognition of the importance of asset price developments has taken a variety of forms since the financial crisis: it has motivated the use of unconventional monetary policies; sparked interest in macro-prudential policy measures to temper asset booms; central banks have implicitly or explicitly elevated financial stability to a policy objective; and new legislative powers have been bestowed upon official institutions to strengthen the oversight of financial markets.


10 The cross-country findings in Ludwig and Slok (2004) and Case and others (2013) point to this. Carvalho and others (2012) also provide evidence of a strengthening wealth effect over time.

11 The emerging market share of world financial assets is only half their share of world GDP (Haldane, 2014).
grown at a faster rate (Figure 2, right hand panel). Both the asset management and banking industries also have similar levels of asset concentration among the largest institutions.

![Figure 1. Institutional Assets under Management (left) and Investment Vehicles (right)](image)

Source: Pensions and Investments/Towers Watson (2014), Barclays Hedge, Preqin, ETFGI, Bloomberg, Author's estimates.

Notes: Figures in right hand panel are based on the relative share of assets in worldwide collective investment vehicles.

![Figure 2. Global Bank Assets vs. Investment Firm Assets under Management](image)

Source: Pensions and Investments/Towers Watson (2014), Bloomberg, Author's estimates.

Notes: Figures in right hand panel are based on the largest 500 banks and institutional investment firms respectively.

12 In the discussion of the relative size of banks and asset managers, it is worth pointing out however that the balance sheet of large institutional investment managers is typically only a fraction of their assets under management. This reflects their primary role as investment agent.

13 For instance, the 25 largest banks in the world have a 51 percent share of assets of the largest 500 banks. The comparable figure for asset managers is 46 percent.
Despite these trends, the root causes and policy implications of asset bubbles continue to be a battleground for divergent views that often connect to fundamental and strongly held interpretations of the nature of markets, political economy and the role of the state. The key related question can be framed as follows: is the underlying tendency for asset price movements best represented by informationally-efficient and fully informed rational agents, where in the spirit of Friedman’s (1953a) ‘stabilizing speculators,’ asset valuation misalignments are quickly and smoothly corrected with few implications for economic activity—thus obviating the need for an interventionist policy? Or instead, can significant mispricings be sustained for long periods of time due to the combination of institutional limits to arbitrage and irrational and systematic behavioral errors that leave markets prone to self-reinforcing boom-bust cycles in the absence of policy activism? The implications derived from these alternative paradigms clearly diverge.

Against this backdrop, the contributions of this paper are threefold. First, the broad contours of the policy debate surrounding pre-emptive bubble intervention are distilled with an emphasis on the pendulum shift to a more activist approach since the global financial crisis. Second, as policy prescriptions vary widely across asset bubble models, I distinguish between competing explanations for why bubbles arise and might be impervious to self-correcting influences. I place emphasis on the institutional limits to arbitrage, notably incentives in the investment management industry. And third, I evaluate the merits of alternate policy options to deal with bubble persistence.

One of my central arguments is that the business risk of rapidly growing asset managers lends itself toward herding and ‘rational bubble-riding,’ and that even in the absence of leverage, this industry will tend to contribute to procyclicality and financial instability. It should perhaps serve as a shot across the bow that the rise of the institutional investment management industry—populated with what are presumably the most sophisticated, well resourced and rational speculators in the world—has coincided with three of history’s largest bubbles in the last twenty five years: the Japanese Heisei bubble of the late 1980s, the global equity bubble of the late 1990s, and the structured credit bubble of the mid 2000s. Because asset bubbles have continued to pose a threat to financial stability despite the increasing sophistication and institutionalization of financial markets, conventional explanations for bubbles—such as information shortages (‘limits to learning’) and incomplete markets (‘frictional limits to arbitrage’)—no longer seem a complete or useful guide. To mitigate the risk of ‘fighting the last war,’ I argue that remedial policies need to be multifaceted, extending beyond the current suite of counter-cyclical macroprudential and monetary policies to the realm of time-invariant measures such as principal-agent contract and financial benchmark reform. The motivation for this emphasis is that while countercyclical policies react to the symptoms of bubbles, they cannot adequately address the underlying institutional causes which are strongly related to incentives in the asset manager—owner relationship.
The structure of the paper is as follows. Section II outlines the competing paradigms in the debate of whether or not policy makers should lean against asset price booms (and if so, how). Section III reviews the evolution in asset pricing theories, before distilling explanations for why asset booms might be impervious to self-correcting forces (and thus prone to bubbles). The associated policy implications are addressed in Section IV. Concluding remarks and suggestions for future research are presented in Section V.

II. The ‘Clean vs. Lean’ Debate: A Survey

The issue of what can and should be done about asset price booms has been, and remains, the topic of considerable debate. Prior to the global financial crisis it was one framed as follows: should policy makers confine their management of asset price cycles to cleaning-up the after-effects of bursting bubbles, only if and when they occur? Or should they instead be prepared to lean against asset price booms as they gather momentum, and thereby lower the probability of rare but disastrous busts, but at the potential cost of prematurely curtailting expansions? The pre-crisis dimensions of the competing ‘ex-post clean’ and ‘ex-ante lean’ paradigms are summarized in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>'Ex-post Clean' Paradigm (Jackson Hole Consensus)</th>
<th>'Ex-ante Lean' Paradigm (Basel Consensus)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability of policy makers to identify bubbles ex-ante</td>
<td>Virtually impossible—can’t know more than the market</td>
<td>Difficult but no harder than for other policy objectives</td>
</tr>
<tr>
<td>Does price stability guarantee financial stability?</td>
<td>Yes, but even if not, it’s the best contribution policy can make</td>
<td>No, it can even amplify financial instability</td>
</tr>
<tr>
<td>Welfare losses associated with pre-emptive tightening</td>
<td>Prohibitively high — cure is worse than the disease</td>
<td>Less than allowing bubbles to grow larger and then burst</td>
</tr>
<tr>
<td>Is monetary policy effective in dampening asset booms?</td>
<td>No, too blunt and low pass-through</td>
<td>Yes—even if just a verbal commitment</td>
</tr>
<tr>
<td>Are the clean-up costs of bursting bubbles manageable?</td>
<td>Yes with an aggressive and swift policy response</td>
<td>Can be too large to contain if the bubble expands unchecked</td>
</tr>
<tr>
<td>Welfare-enhancing objective function to include:</td>
<td>Only consumer prices</td>
<td>Consumer prices, credit growth, and risk premiums (or price indices with asset prices embedded)</td>
</tr>
<tr>
<td>Underlying tendency of asset price movements</td>
<td>Mostly efficient, rational, self-correcting and thus stabilizing</td>
<td>Often inefficient, irrational, self-reinforcing, destabilizing</td>
</tr>
<tr>
<td>Moral hazard risk</td>
<td>Unavoidable</td>
<td>Mitigated</td>
</tr>
</tbody>
</table>

14 Consistent with the general tone of the debate before the crisis, this section focuses largely on the role of monetary policy in curbing asset price booms, rather than other regulatory and macroprudential measures—interest in which has mushroomed mainly since the crisis.

15 In practice, there are of course many shades of gray between these two stark positions, but the demarcation helps to elucidate the areas of disagreement.
The ‘Ex-post Clean’ Paradigm

The ‘ex-post clean’ paradigm emphasizes the difficulties faced by authorities in reliably identifying a bubble in progress. Authorities have often set a high bar for establishing the existence of bubbles due to a perception they have less information about valuation misalignments than markets themselves (Cogley, 1999). If a central bank with no informational advantage over market participants believes a bubble has developed, investors—with more complete information and a vested interest in correcting irrational mispricings—will almost surely have arrived at the same realization and drive asset prices down to a level more consistent with fundamentals. To the extent that speculators stabilize markets by correcting valuation misalignments, they obviate the need for policy intervention in the first place (Friedman, 1953a). As pointed out in the late 1990s by U.S. Federal Reserve Chairman Alan Greenspan, “there is a fundamental problem with market intervention to prick a bubble: it presumes that you know more than the market … identifying a bubble in the process of inflating may be among the most formidable challenges confronting a central bank, pitting its own assessment of fundamentals against the combined judgment of millions of investors.” Bernanke (2002a) stressed that it was unlikely policy makers could accurately disentangle the rational and irrational component of asset prices, and expressed concern over “the effects on the long-run stability and efficiency of our financial system if the Fed attempted to substitute its judgments for those of the market.” Kohn (2006) similarly noted, “risk premiums often move in mysterious ways … we have a very poor understanding of the forces driving speculative bubbles.”

The ‘ex-post clean’ approach has also been advocated on the basis of counterfactuals and opportunity costs. For instance, policy makers cannot know ex-ante if a period of unusually elevated asset prices will plateau, benignly deflate, or be followed by a ‘catch up’ in fundamentals. As such, bubbles can only ever be identified (long) after the fact. Furthermore, a great deal of uncertainty surrounds the overall impact on the economy in the event of a sharp asset price reversal. The effects on inflation and output from booms and busts in asset prices sometimes reveal themselves with significant and variable lags (Yellen, 2009). In lieu of so much uncertainty, a more prudent course may therefore be to wait and observe what negative impact, if any, a downturn in asset prices exerts on the economy. Policies conditioned on noisy or unobserved variables can be destabilizing rather than stabilizing, rendering pre-emptive countervailing actions counterproductive (Friedman, 1953b).

Prior to the crisis, the ‘ex-post clean’ approach, characterized as the ‘Jackson Hole Consensus’ (Issing, 2009) or the ‘Greenspan Doctrine’ (Mishkin, 2011), prevailed most notably in U.S. policy circles; see for instance, Greenspan (1996, 2004); Bernanke and Gertler (1999, 2001); Bernanke (2002a); Bullard and Schaling (2002); Blinder and Reis (2005); Kohn (2008); and Mishkin (2008).

A decision tree response matrix incorporating elements of uncertainty can be found in Rudebusch (2005).

If authorities believe they have the tools available to keep the harmful effects of a bursting bubble at a manageable level, then a policy of leaning against the wind may also be unnecessary. In addition to countercyclical fiscal policy, authorities have various means by which to arrest the fallout of bursting bubbles: policy rates can be taken down to the zero bound; a commitment can be made to maintain policy rates at exceptionally low levels for an extended period; risk premia can be lowered through central bank asset purchases; inflation expectations can be bolstered (lowering real interest rates) with a credible commitment to increase the supply of domestic currency; and authorities can intervene in the foreign exchange market to weaken the domestic currency and stimulate exports (Svensson, 2001; Bernanke 2002c). Recapitalization measures are also available in the event a bust threatens the broader financial sector. Prior to the global financial crisis, a number of researchers pointed out that lessons had been learned from the timidity of the Japanese policy response to the bursting Heisei bubble in the 1990s (Bernanke, 2000, 2003; Ahearne and others, 2002; Posen, 2003; Kohn, 2006). The relatively modest economic fallout associated with the stock market crashes of 1987 and 2000-2002 helped solidify the view that asset price busts, if and when they occurred, could be comfortably contained by ‘cleaning’ rather than ‘leaning.’

Even if bubbles could be identified with confidence in real time, advocates of the ‘ex-post clean’ approach have also argued that monetary policy is not the right tool to safely diffuse them. Implicit in this view has been the belief that the best contribution central banks could make to financial stability was to focus solely on price stability. If the sensitivity of asset markets to interest rate movements is low (and/or uncertain), curbing asset price increases

19 Representative of this view, Blinder and Reis (2005, pp. 67-68) concluded at the 2005 Jackson Hole Conference: “The ‘mop up after’ strategy received a severe real world stress test in 2000–2002, when the biggest bubble in history imploded, vaporizing some $8 trillion in wealth in the process. It is noteworthy, but insufficiently noted, that the ensuing recession was tiny and that not a single sizable bank failed. In fact, and even more amazingly, not a single sizable brokerage or investment bank failed either. Thus the fears that the ‘mop up after’ strategy might be overwhelmed by the speed and magnitude of the bursting of a giant bubble proved to be unfounded. Regarding Greenspan’s legacy, then, we pose a simple rhetorical question. If the mopping up strategy worked this well after the mega-bubble burst in 2000, shouldn’t we assume that it will also work well after other, presumably smaller, bubbles burst in the future? Our suggested answer is apparent.”

20 The limits of monetary policy in curbing asset booms in an open economy have also been emphasized in a different context, where large cross-border capital flows (rather than domestic central bank policy rates) exert the primary influence on domestic asset prices (Reinhart and Reinhart, 2011). This argument was not intended to absolve central banks from the responsibility of leaning against the wind per-se, but rather, simply highlights there may be limitations as to what is achievable from monetary policy when vast amounts of international capital can flow freely across borders. Though this is a more pressing issue for countries that lack depth in domestic capital markets (such as emerging market economies), the divergent behavior of long-term bond yields from movements in central bank policy rates was particularly noticeable in the U.S. in the years preceding the global financial crisis. This interpretation is consistent with the bond “conundrum” cited by Greenspan (2005) and the “savings glut” hypothesis of Bernanke (2005), where changes in long-term forward interest rates in the U.S. (and other developed bond markets) were negatively correlated with short term interest rates during the Federal Reserve’s two year tightening cycle in 2004-06.
may require a policy response so vigorous it risks unnecessary collateral damage to the broader economy. The crudeness of monetary policy might induce the very outcome it sought to avoid, namely a sharp recession. In this sense, the cure would be as bad, or possibly worse, than the disease (Bernanke, and others 1999; Greenspan, 2002; Bernanke, 2002a; Gruen and others, 2005; Bernanke and Kuttner, 2005; Kohn, 2006; Kuttner, 2012). For instance, invoking the Tinbergen separation principle (where to achieve a certain number of objectives, at least an equal number of instruments are required), Bernanke (2002a) argued,

“We cannot practice safe popping, at least not with the blunt tool of monetary policy. Certainly there is no way to direct the effects of monetary policy at a single class of assets while leaving other financial markets and the broader economy untouched. One might as well try to perform brain surgery with a sledgehammer. The problem of safe popping applies with double force to the aggressive bubble-popping strategy. A truly vigorous attempt by a central bank to rein in a supposed speculative bubble may well succeed but only at the risk of throttling a legitimate economic boom or, worse, throwing the whole economy into depression.”

In a related context, Greenspan (2004) suggested that in instances where a moderate first round of monetary tightening failed to elicit the desired effect, markets could become emboldened with a sense of immunity and rise to even loftier heights (thus posing an even larger threat to longer-term financial stability). A rise in interest rates could lead to rational agents boosting their assessment of the equilibrium growth rate during a speculative episode, thus negating the negative impact on fundamental asset values from a higher discount rate.

Much of the skepticism over the role that monetary policy could play in safely diffusing bubbles owed to the criticism leveled at the Federal Reserve for its role in tightening monetary conditions in response to frothy stock market conditions ahead of the Great Depression (Friedman and Schwartz, 1963; Schwartz, 1981; White, 1990; and Bernanke 2002a, 2002b). As Bernanke (2002a) argued, “Early in 1928, the Fed passed into the control of a coterie of aggressive bubble-poppers … The correct interpretation of the 1920s is not the popular one – that the stock market got overvalued, crashed, and caused a Great Depression. The true story is that monetary policy tried overzealously to stop the rise in stock prices. But the main effect of tight monetary policy … was to slow the economy … The slowing economy, together with rising interest rates, was in turn a major factor in precipitating the stock market crash.”

21 This is by no means a unanimous interpretation of events preceding the Depression. As Eichengreen (1992, p. 213) points out, Adolph Miller, the one professional economist on the Federal Reserve Board in the late 1920s, argued the easy money proclivities of the Federal Reserve System in the years immediately prior to the crash were “one of the most costly errors committed by it or another other banking system in the last 75 years.” To this day there remains much debate over the degree to which monetary settings contributed to the Depression.
The ‘Ex-ante Lean’ Paradigm

The ‘ex-ante lean’ paradigm trades off the possibility of subpar macroeconomic performance in the near term, for the prospect of significantly better performance over a longer horizon, on the understanding that the repercussions of a major asset price collapse could be unmanageable. According to this view, a policy of ‘benign neglect’ toward asset price booms can store up trouble, making it difficult and costly to contain at a later date. Put differently, the long-term welfare costs associated with allowing asset bubbles to grow unchecked may be prohibitively high—although remedial policy tools may not be costless, they can be deployed effectively enough to curb the accumulation of dangerous excesses and thus aid in maximizing welfare over the longer-term. As such, authorities may need to take out insurance against the possibility of a large crash in the future by tightening financial conditions earlier than would otherwise be the case.

By cooling off booms before they spiral out of hand, these efforts can contribute to lengthening the economic cycle (Roubini, 2006). Barlevy (2012) suggests theoretical economic models of asset bubbles, which have tended to downplay the case for pre-emptive intervention, have yet to adequately address the long-term welfare implications associated with their bursting that the recent crisis brought to light.

Proponents of pre-emptive intervention argue that the difficulties associated with identifying bubbles in real time should not absolve policy makers from this responsibility. Although the precise magnitude of a bubble cannot be known with certainty, authorities are frequently tasked with making decisions in real time with less than perfect information. Constructing estimates of market risk premia may be no more difficult than constructing estimates of the natural rate of interest or output (Blanchard 2000; Cecchetti, 2003). Incorporating uncertainty into the decision making process, Dupor (2005) suggested policy makers should lean against the wind of asset bubbles with an intensity linked to the confidence in which they consider asset prices to be misaligned. Others have suggested policy makers have a stronger case in leaning against credit-fuelled asset bubbles, in part because authorities have informational advantages over the market in terms of their access to timely qualitative and quantitative credit-related data (White, 2009; Mishkin, 2010).

---

22 The ‘ex-ante lean’ paradigm, or ‘Basel Consensus,’ tended to attract relative more support in Europe than the U.S.; see for instance, Cecchetti and others (2000); Bordo and Jeanne (2002); Borio and Lowe (2002, 2004); Issing (2003, 2009); Roubini (2006); White (2006, 2009); Cooper (2008); and Taylor (2009).

23 The cost of this insurance policy might be loosely measured as foregone output in the intervening period resulting from tighter-than-otherwise monetary conditions.

24 Consider the uncertainty associated with estimates of the output gap introduced by the impact of data revisions: the average absolute difference between the first reported quarterly estimate of U.S. GDP (upon which policy decisions might be based) and the final estimate was 1.30 percentage points between 1983 and 2010. Given the average GDP growth rate over this period was 3.06 percent, this suggests authorities are making policy decisions based on real time data with a significant degree of error.
The near-exclusive focus of monetary policy on goods and services inflation (relative to asset price inflation) might also contribute to the formation of bubbles and financial instability, particularly in instances where positive supply shocks occur in the market for goods but not assets. If positive productivity or supply shocks lower the rate of goods and services inflation below the target of the central bank, policy rates may be lowered to levels that ferment speculative risk taking. Notably, all the major U.S. stock market booms and busts over the past two centuries, including the 1920s and 1990s, have occurred against the backdrop of relatively easy monetary conditions and strong productivity growth which kept consumer price inflation low and stable (Bordo and Wheelock, 2004; Christiano and others, 2010). Bordo and Wheelock (2006, 2007) document the tendency for asset booms to occur during periods of below target inflation (or disinflation) across a host of industrialized countries over the course of the twentieth century.25

From an analytical standpoint, the observation that asset valuations are mean reverting and that asset booms frequently accompany periods of benign inflation and/or robust credit growth might suggest a role for an augmented Taylor-rule to guide monetary policy settings.26 This would dictate leaning against below-average risk premia and/or above-average credit growth, in addition to above-target consumer price inflation (Blanchard, 2000; Borio and Lowe, 2002, 2004; Cecchetti, 2003; Christiano and others, 2010; Acharya and Naqvi, 2012). Alternatively, because they impact the economy through the wealth effect and embed expectations about future economic conditions that can be relevant for inflation-targeting central banks, asset prices could be incorporated directly into the inflation metric that monetary authorities target (Trichet, 2003; Goodhart and Hofmann, 2007).

Finally, moral hazard considerations could also support a policy of leaning against the wind. By pre-committing to an asymmetric response to asset price booms and busts such that investors believe policy makers will ease aggressively in an asset price bust, but refrain from tightening in a boom, asset prices might take on a more dangerous speculative dynamic than would otherwise be the case (Caballero and Krishnamurthy, 2003; Roubini, 2006; Cooper, 2013). 

---

25 Accommodative monetary conditions (as measured by deviations of monetary policy from a Taylor rule) and strong credit growth have also been associated with stock and real estate booms across a host of OECD countries since the end of the Bretton Woods system (Detken and Smets, 2004).

26 It should be noted that this argument is distinct from one (for which there is little support) which suggests policy makers should target the level of asset prices on an on-going basis. Moreover the case for an augmented Taylor rule along these lines has by no means gone unchallenged. Bernanke and Gertler (1999) argued including asset prices in a Taylor-type interest rate rule was likely to be destabilizing. Bullard and Schaling (2002) argue that including equity prices in a Taylor-type policy rule will likely degrade economic performance by creating indeterminacy of rational expectations equilibrium where such indeterminacy did not otherwise exist. Trichet (2003) cautioned that such an approach could open a “Pandora’s Box.”
At the very least, as evidence of a bubble begins to mount, a verbal policy strategy of moral suasion could have few costs relative to potential benefits. Sober assessments of ‘frothy’ market conditions by authorities in public forums, or the more aggressive announcement that policy makers stand ready to act should asset prices continue to soar, might be sufficient to prevent the emergence of destabilizing bubbles in the first place (Blanchard, 2000).

The New Consensus and the Road Ahead

The aftermath of the global financial crisis has resolved many important elements of the ‘clean vs. lean’ debate in favor of the latter. Prior to the crisis, asset bubbles and financial stability were in some quarters viewed as a “potentially dangerous distraction, risking the central bank’s attention to, and credibility in, achieving its price-stability mandate … even at the Federal Reserve, financial stability was rarely discussed” (Williams, 2014, p. 2). Yellen (2009, p. 8) noted immediately following the crisis that while most central bankers (particularly in the United States) have historically resisted the idea of leaning against asset price increases, “now that we face the tangible and tragic consequences of the bursting of the bubble, I think it is time to take another look.” And as Stein (2013a, p. 1) summarized, “Post-crisis, I think it’s safe to assume that nobody would now argue that we should be putting all our eggs in the ‘clean’ basket.” Broadly speaking, a new consensus has emerged in policy making circles to recognize that:

- In both developed and emerging economies, the long-term welfare costs associated with allowing bubbles to grow unchecked can be more damaging than previously thought (in this sense, asset bubbles pose an ‘equal opportunity menace’);
- Long periods of stability can be destabilizing to the extent they promote ever more aggressive forms of risk-taking in asset markets (Minsky, 1986, 1992);
- Frictions in asset markets can have a pronounced, non-linear impact on the economy, especially if the imbalances to which they give rise are allowed to accumulate unchecked;

27 Ravn (2011) and Hall (2011) both find strong empirical evidence of U.S. monetary policy pursuing an asymmetric monetary policy with respect to the stock market.

28 More forcefully, Malliaris (2012, p. 427) has suggested the ex-post clean approach has lost is “commanding power among central bankers,” and that, “instead of central bankers asking how to respond to asset price bubbles, they should be asking ‘How may central banks cause asset bubbles?’”

29 While policy makers prior to the crisis understood that financial frictions could be economically damaging, Mishkin (2011, p. 12) noted, “the macroeconomic models used for forecasting and policy analysis, whether they were dynamic stochastic general equilibrium models or more traditional macro-econometric models like FRBUS in use at the Federal Reserve, did not allow for the impact of financial frictions and disruptions on economic activity.” Financial distortions, though theoretically relevant, were considered quantitatively too small to sway the conduct of monetary policy, especially in advanced economies (Bayoumi and others, 2014). Brunnermeier and others (2012) survey the macroeconomic implications of financial frictions.
While long-term price stability should remain a primary objective of monetary policy and can make a contribution to financial stability, it is not a sufficient condition to insure against asset bubbles or financial instability more generally;\(^{30}\)

- Highly targeted prudential measures should form the first line of defense against nascent signs of asset bubbles and concomitant threats to financial stability;\(^{31}\)
- More generally, a new macro-financial stability architecture is required at the national and global level to address the threat that asset bubbles and financial imbalances pose to economic stability.

This said, old uncertainties may have been resolved, but new ones have emerged in their wake. With this in mind, the remainder of the paper is motivated by the concept that the optimal mix of policy responses—regulatory, monetary, and macro-prudential in nature—can be refined with a clearer distinction among the underlying forces that give rise to asset bubbles in capital markets.

### III. Theories of (In)Efficient Markets and Speculative Bubbles

#### A. Bubbles and the (In)Efficiency of Markets – A Review

Explanations over how and why asset bubbles arise have fascinated and perplexed financial economists for decades. Broad-based interest in the topic intensified half a century ago, following Samuelson's (1965) proof that stock prices should follow a random walk if rational competitive investors require a fixed rate of return, and Fama’s (1965a, 1965b) demonstration that stock prices appeared virtually indistinguishable from a random walk.\(^ {32,33}\)

\(^{30}\) In Norway, policymakers increased the policy interest rate in mid-2010 when they were facing escalating household debt despite inflation below target and output below capacity, in part as a way of “guarding against the risk of future imbalances.” Similarly, Sweden's Riksbank held its policy rate “slightly higher than we would have done otherwise” because of financial stability concerns. These effectiveness of such measures has attracted considerable debate (for instance, see Svensson, 2014).

\(^{31}\) See most recently IMF (2011, 2013), Viñals (2013), Bayoumi and others (2014), and Yellen (2014).

\(^{32}\) Though the ideas of the Efficient Market Hypothesis (EMH) and the Random Walk Hypothesis (RWH) are often closely linked, they are not directly substitutable. The EMH evolved to accommodate rationally time-varying risk premia (expected returns), while the RWH is a special case of constant expected returns: evidence of time-varying return predictability does not invalidate the EMH as it does the RWH. More generally, to make any statement about market efficiency first requires an assertion of how prices should be set using a pricing model, such as the random walk or capital asset pricing model—this is the well known joint hypothesis problem which poses a significant obstacle to making conclusive statements about market efficiency.

\(^{33}\) The origins of the efficient market asset pricing theory can in fact be traced back to well before the independent analysis of Samuelson and Fama in 1960s—much of the financial mathematical toolkit was revealed in Bachelier’s (1900) seminal Ph.D. thesis, Théorie de la Spéculation, which sat idle in a Parisian
From these findings a rich asset pricing theory began to emerge on the basis that investors made rational decisions and markets fully and correctly incorporated all relevant information, culminating in Fama’s (1970, 1991) seminal efficient markets hypothesis. Importantly, the framework implied rational arbitrageurs acted as a stabilizing influence on financial markets by swiftly moving to correct valuation misalignments (Friedman, 1953a; Fama 1965a). As such, speculative bubbles were virtually impossible. Backward induction also ruled out bubbles in a world of a finite number of investors and where asset prices had already run up strongly, as no rational agent would put themselves in a position to be the last in line, “getting stuck with a hot potato” (Tirole, 1982, p. 1180). Asset markets would thereby converge toward market efficiency over time, as the impact of competition and natural selection worked to weed out irrational investors with a systematic tendency to buy high and sell low (such investors cannot lose money forever). Jensen’s (1978, p. 95) claim that there was “no other proposition in economics which has more solid empirical evidence supporting it than the Efficient Market Hypothesis” arguably heralded the high watermark of classical asset pricing theory.

library for decades. Shortly after reviewing Bachelier’s thesis, Henri Poincaré asserted it would not be sensible to take the Brownian motion model as a basis for analyzing financial markets because individuals who monitor one another (as in a market) do not take independent decisions, but rather tend to ‘herd.’ Thus, Poincaré anticipated one of the most critical features of asset pricing long before modern economists (Kirman, 2009).

34 The classic taxonomy of information sets, developed by Roberts (1967), distinguishes weak-form efficiency (where the information set includes only the history of prices), semi strong-form efficiency (the information set includes all publicly-known information), and strong-form efficiency (the information set includes all known information). Subsequent theories of asset price determination have challenged each of these assumptions.

35 The following key assumptions came to feature in the efficient markets theory: all investors make fully rational decisions (they correctly calculate expected values as the probability-weighted sum of potential outcomes, and make decisions fully consistent with these estimates); all investors are equally and fully informed in that “they have access both to the correct specification of the ‘true’ economic model and to unbiased estimators of its coefficients” (Friedman, 1979, p. 38); all investors share the same beliefs and risk preferences; and markets are frictionless (i.e., there are no binding constraints on short sales or leverage).

36 In early ‘rational bubble’ models, bubbles were shown to exist where there were no close substitutes for an infinitively-lived asset, and the bubble component expanded at the same rate as its required return. A bubble growing at a slower rate than its required return has a present value of zero; one growing at a rate faster than its required return has an undefined present value (infinity). In the rational setting, the process of backward induction ruled out the prospect that assets with a finite life could exhibit bubble-like characteristics: in the case of assets with a known maturity and an upper bound on price (i.e., a bond maturing at par at a fixed future date), if all agents have the same information and agree that a bubble exists today but will burst at time $t$, then why would it not burst at time $t-1$, since at that time no investor would be willing to buy the asset at an inflated price? The same logic can be extended to time $t-2$, $t-3$, and so on, up until the present. While Kreps (1977) argued rational bubbles could be ruled out if it was commonly known that the initial allocation in an economy was Pareto efficient (there were no gains to be made from trade), Allen and others (1993) showed that a rational bubble could exist for a finite-lived asset when common knowledge about the existence of a bubble was absent and short sale constraints were binding.
From the late 1970s however, competing theories of asset price determination (and bubble formation) that questioned the underpinnings of the efficient markets paradigm began to emerge. As Shleifer and Summers (1990, p. 19) noted, “if the efficient markets hypothesis was a publicly traded security, its price would be enormously volatile.” Drawing upon a long empirical history of booms and busts, the stylized five phase Kindleberger-Minsky model emphasized the procyclical tendencies of investor psychology as rational boom gave way to irrational bubble. More formally, Grossman and Stiglitz (1980) were among the earliest to emphasize “the impossibility of informationally efficient markets” in the sense that there must exist “an equilibrium degree of disequilibrium”—if markets were perfectly efficient, there would be no profit to gathering and processing information, in which case there would be no incentive to trade and markets would collapse (Milgrom and Stokey, 1982). Shiller (1981) and Leroy and Porter (1981) questioned the efficiency of markets on the basis that stock market volatility seemed far in excess of levels justified by volatility in fundamentals.

Just as the statistical validity of these findings was called into question (Kleidon, 1986; Marsh and Merton, 1986), the stock market plunge in October 1987 on virtually no new fundamental news dealt a further blow to the standard asset pricing model. If markets were fully efficient, the market crash had to reflect either a very large increase in risk premia because the economy suddenly became a lot riskier, or a significant decline in the expected future growth rate of dividends. Yet the news flow prior to the crash cannot adequately explain a 22 percent one-day decline in the valuation of the American corporate sector. Invoking the Japanese Heisei bubble, Allen (2001) argued that the standard asset pricing paradigm could not possibly explain how in 1989, the few hundred acres of Tokyo land on which the Imperial Palace is located could have been equivalent to the value of all of the land in Canada or California. A review of studies of speculative investment behavior led Krugman (1993, p. 17) to assert, “Nobody has found any reasonable way to save the efficiency hypothesis within the data. This is devastating in its impact on our research. What we know how to model are efficient markets; what we apparently confront are inefficient ones.”

---

37 The five phase model begins with an initial displacement, such as a new technology, institutional innovation or other positive shock. This leads to a rational boom phase where prices increase more or less in line with improving fundamentals. A period of steadily rising prices then gives way to a phase of euphoria, where the price action and trading volumes turn explosive, fuelled by leverage. Expectations of future price gains become widespread and self-fulfilling. Insider profit-taking eventually emerges as those with superior information (such as company management or sophisticated valuation-conscious investors) begin to sell down their exposure to late entrants. As the price action begins to deteriorate, fewer and fewer new participants are drawn in, and once margin calls begin to prompt fire sales, widespread panic follows.

38 French (1988) argued there was some news in the week leading up to the crash: the announcement of a larger-than-expected trade deficit, news that the Congress would support the elimination of the tax benefits on leveraged buyouts, and press speculation that the Federal Reserve would raise interest rates. Nonetheless, the efficient market theory cannot reconcile the magnitude of the crash on this basis.
After conceding that long-term asset returns displayed some degree of predictability, the focus of efficient market studies shifted to arguing that movements in risk premia rationally reflected changes in aggregate consumption and wealth, rather than irrational investor psychology. Risk premia were seen as rationally lower when the economy was strong (Fama and French, 1989; Campbell and Cochrane, 1999; Lettau and Ludvigson, 2001, Cochrane 2011). This helped to rationalize the direction of asset returns, but their magnitude through an economic cycle has remained difficult to align with the fundamentals they supposedly reflect. Over the past century for instance, three-quarters of the present value of the U.S. stock market has derived from cash flows discounted more than seven years into the future. By implication, a recession or boom in any given year should be a trivial event for long duration assets like stocks, whose future cash flows derive from stable estimates of long-term GDP growth. As Grantham (2012, p. 2) observed, “the short term will always be exaggerated, and the fact that a corporation’s future value stretches far into the future will be ignored … the market trades as if all value lies within the next 5 years, and sometimes 5 months.”

Other empirical challenges to efficient markets theory have arisen over the past quarter century, including (i) behavioral studies suggesting systematic departures from rational decision making (Kahneman, 2011); (ii) the negative correlation between rational ‘objective’ measures of expected returns, such as yield-based risk premia, and ‘subjective’ survey readings of expected returns (after a sharp run up in prices, investor expectations of future returns tend to rise while rational objective measures fall; Ilmanen, 2011; Greenwood and Shleifer, 2014); (iii) the tendency for periods of unusually low risk premia to be followed by asset busts and recessions, suggesting systematic forecast errors (Jones, 2014); (iv) the positive correlation between trading volumes and prices during bubbles (Cochrane, 2011; Jones, 2014); and (v) the host of documented asset pricing ‘anomalies’ (such as the sustained success of simple momentum strategies) which exploit widely available data at the time of portfolio formation and cannot be easily explained by ‘rational compensation for risk’ arguments (Fama and French, 2008; Ilmanen, 2011; Asness and Liew, 2014).

---

39 Results available from the author upon request.

40 Reflecting on the lessons for policy makers stemming from the limitations of economic models to adequately capture asset dynamics, Williams (2013, p. 2) argued, “A cursory reading of the academic literature on asset prices reveals a litany of puzzles, conundrums, paradoxes, and anomalies. Much of the research on asset prices continues to rely on highly stylized models with identical agents, rational expectations, and optimizing behavior. According to the prevailing view, asset price surges that many would perceive to be bubbles are not really so … Adherents of this view may admit that bubbles have occurred in the past—like the dot-com boom and bust. And they may even be willing to accept that bubbles are something to worry about in the future—say, in financial supervision. But, in practice, they are never willing to find a bubble in the present. There’s always a reason why what looks like a bubble, walks like a bubble, and quacks like a bubble is not actually a bubble.”

41 The ‘value premium’ first reported in Graham and Dodd (1934), a strategy of buying cheap assets and selling expensive assets, is another widely-cited anomaly, but is easier than others to justify with rational explanations.
While there were no more hotly contested topics in financial economics prior to the financial crisis, its aftermath has seen critiques of the efficient markets paradigm become more prominent both inside and outside policy making circles. In testimony to the House Committee of Government Oversight and Reform, Greenspan (2008) concluded that of the “paradigm (that) held sway for decades … the whole intellectual edifice collapsed.” Volcker (2011) asserted, “it should be clear that among the causes of the recent financial crisis was an unjustified faith in rational expectations and market efficiencies.” In line with Shiller’s (1992, p.8) claim “the efficient markets hypothesis represents one of the most remarkable errors in the history of economic thought … remarkable in the immediacy of its logical error and in the sweep and implications of its conclusions,” Krugman (2009) argued, “the belief in efficient financial markets blinded many if not most economists to the emergence of the biggest financial bubble in history. And efficient-market theory also played a role in inflating that bubble in the first place. Now that the undiagnosed bubble has burst, the true riskiness of supposedly safe assets has been revealed and the financial system has demonstrated its fragility.”

Drawing lessons from the crisis for fellow policy makers, Williams (2013, p. 7) urged an overhaul of the policy response toolkit: “The lesson from history is clear: asset price bubbles and crashes are here to stay. They appear to be a consequence of human nature. And the events of the past decade demonstrate the enormous human costs of asset price bubbles and crashes. To understand the past and avoid a recurrence of the devastating events we lived through so recently, we need to acknowledge that investors and financial markets do not behave the way rational asset price theory implies … It’s essential if we want to design policies that foster robust economic performance in the future.”

B. Competing Models of Bubble Formation and Persistence

With the broad consensus in policy circles having recently shifted toward a more activist approach to leaning against asset bubbles, a firm understanding of the process by which they grow and persist has assumed increasing importance. The related literature can be broadly distilled into four classes of models, each with different assumptions and explanations for the persistence of bubbles (Table 2): the limits to learning; frictional limits to arbitrage; institutional limits to arbitrage; and irrational behavioral models. The policy implications

42 Grantham (Nocera, 2009) was more vigorous still: “The incredibly inaccurate efficient market theory was believed in totality by many of our financial leaders, and believed in part by almost all. It left our economic and government establishment sitting by confidently, even as a lethally dangerous combination of asset bubbles, lax controls, pernicious incentives and wickedly complicated instruments led to our current plight. ‘Surely, none of this could be happening in a rational, efficient world,’ they seemed to be thinking. And the absolutely worst part of this belief set was that it led to a chronic underestimation of the dangers of asset bubbles breaking.”

43 The first three classes of models are examples of New Generation Rational Models, in that they retain the assumption of rational decision making, but allow for bubbles by relaxing assumptions related to information, incentives and frictions. For recent surveys, see Brunnermeier and Oehmke (2012) and Scherbina (2013).
associated with these competing explanations has not been a prominent feature of this analysis, though each entails a quite distinct policy response (Section IV).

Table 2. Stylized Summary of Asset Pricing/Bubble Models

<table>
<thead>
<tr>
<th>Assumptions</th>
<th>Fundamental View on Bubbles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rational</td>
<td>Homogenous</td>
</tr>
<tr>
<td>Expectations</td>
<td>Beliefs and Preferences</td>
</tr>
<tr>
<td>I. Neoclassical model (EMH)</td>
<td>√</td>
</tr>
<tr>
<td>II. &quot;Limits to Learning” model</td>
<td>√</td>
</tr>
<tr>
<td>III. &quot;Frictional” Limits to Arbitrage model</td>
<td>√</td>
</tr>
<tr>
<td>IV. &quot;Institutional” Limits to Arbitrage model</td>
<td>x or √</td>
</tr>
<tr>
<td>V. (Irrational) Behavioral models</td>
<td>x</td>
</tr>
</tbody>
</table>

Source: Author

Limits to Learning

‘Limits to learning’ models suggest that while investors process information rationally (a key assumption in efficient markets theory), large pricing errors can lead to bubbles because of incomplete and/or inaccurate information. Even if investors engage in ‘Bayesian updating’ as they rationally process new incoming information, it may not be possible to ever have complete knowledge over the (constantly changing) structure of the economy and all the factors capable of impacting the investment opportunity set. Put differently, there are limits to learning over the true data generating process which make ‘rational bubbles’ possible.44 Markets are particularly vulnerable to bubbles where there is a low level of financial literacy among participants, and common knowledge over the existence of a bubble is absent.45

44 These are alternatively referred to as “rational structural uncertainty” models (for a related survey, see Brav and Heaton, 2002). As Kurz (1994) points out, in the rational expectations setting, rational investors make optimal statistical decisions in a world about which they have all relevant structural knowledge.

45 Greenwood and Nagel (2008) document that young, inexperienced investment managers were far more active participants in the technology bubble of the late 1990s relative to their older brethren.
Frictional Limits to Arbitrage

Bubbles can also arise in a rational setting because of frictional limits to arbitrage, such as trading costs and the incompleteness of markets.\(^46\) If differences in beliefs over fundamental value for an asset wash out in aggregate (i.e., there are optimists and pessimists in roughly equal number with equal conviction), but pessimists are constrained in their ability to short or sell assets due to costs or other frictions (such as availability of inventory or instruments), then prices will disproportionately reflect the beliefs of optimists and thus rise above fair value (Miller, 1977; Chen and others, 2002).\(^47\) Residential housing is a notable example. When short sale restrictions bind and common knowledge of a bubble is absent, even finite-lived assets can be subject to rational bubbles (Allen and others, 1993). More generally, arbitrage in the context of ‘stabilizing speculation’ is only riskless under restrictive assumptions: short sales are as seemless to implement as purchases (markets are complete); there are no funding, margin or liquidity constraints; information and transaction costs are negligible; and assets are perfectly substitutable (where two assets are highly correlated, an overpriced asset can be sold, and an underpriced asset purchased, so as to remove overall market risk).\(^48\) However in a practical context, none of these conditions typically hold, especially during systemic liquidity shocks (as demonstrated in 1998 and 2008).

Another source of friction capable of amplifying bubbles stems from the ‘captive buying’ of securities in momentum-biased market capitalization-weighted financial benchmarks. Underlying constituents that rise most in price will see their benchmark weights increase irrespective of fundamentals,\(^49\) inducing additional purchases from fund managers seeking to minimize benchmark tracking error.\(^50\) As a case in point, the 1980s Heisei bubble saw Japan’s share of the MSCI World equity market capitalization soar from 21 percent in 1983 to 51 percent by 1989, while during the 1990s technology bubble, the technology sector weighting in the S&P500 rose from 5 percent in 1993 to 34 percent by 2000 (Jones, 2012).

---

\(^{46}\) Important early contributions in what has become a large literature include De Long and others (1990a); Shleifer and Summers (1990); Lakonishok and others (1992); and Shleifer and Vishny (1997).

\(^{47}\) There is a long history of aversion to short sales: England banned short selling for much of the eighteenth and nineteenth centuries, while in 1803, Napoleon declared short sellers to be enemies of the state.

\(^{48}\) In what has become known as ‘Samuelson’s dictum’ (Jung and Shiller, 2005), ‘macro-inefficiency’ (where overall markets are incorrectly priced relative to fundamentals) is generally considered more difficult to arbitrage (owing to a lack of close substitutes) than ‘micro-inefficiency’ (where market-neutral long/short relative value trades within an asset class can be implemented).

\(^{49}\) In the case of debt markets, an unfortunate and well-known side-effect of market-capitalization based weights is that the largest issuers of debt will attract the largest weights.

\(^{50}\) In the case of downward momentum in price movements, institutions can become forced sellers of securities that have already fallen heavily in price if their market capitalization declines below permissible thresholds.
Importantly, the captive buying phenomena is unlikely to abate given the (passive) benchmark-tracking exchange traded product industry has expanded at a decade-long annual growth rate of more than 20 percent (to US$2.5 trillion), a much faster rate than for other relatively non-benchmark constrained investors (i.e. alternative or absolute return managers).

**Institutional Limits to Arbitrage**

Institutional limits to arbitrage can also give rise to bubbles owing to the nature of incentives and principal-agent issues in the rapidly growing fund management industry. In the United States for instance, the proportion of stocks held by institutional investors hovered around 5 percent in the first half of the twentieth century; from 6 percent in 1952, the institutional share then rose to 45 percent in 1991, and 67 percent by 2010 (Allen and Gorton, 1993; Blume and Keim, 2012). While the growing presence of institutional investors might intuitively suggest markets would become more efficient over time (and thus less prone to bubbles), business risk and compensation practices in the delegated portfolio management industry can encourage institutional herding and ‘rational’ bubble riding. At the very least, incentives for investment managers can dissuade them from leaning against bubbles. Although much of the theory of asset pricing has until recently ignored the real world complication that individuals (asset owners) delegate virtually all their involvement in financial markets to professional intermediaries (asset managers), the institutional investment industry is anything but a frictionless pass-through to asset markets and the economy (Woolley, 2010). The associated principal-agent problems are still to be fully addressed.

Acting as agents, institutional fund managers are only likely to lean against bubbles where they have confidence in their ability to withstand redemptions from principals whilst waiting for a fundamental mispricing to correct. But because principals have imperfect knowledge about the ability of their investment agents—reflecting that managers operate in a high-noise, low-signal environment which makes it difficult to distinguish skill from luck—they may interpret a period of short-term underperformance against peers as a sign of manager incompetence. As the relative underperformance grows, principals are likely to conclude their manager is unskilled and thus terminate them in favor of peers demonstrating

51 When the capital asset pricing model (CAPM) was developed in the mid-1960’s, around 85 percent of US equities were then directly held by domestic households (Brennan and Li, 2008). Davis and Steil (2001) report that the share of household wealth managed by financial institutions has increased sharply in recent decades, in particular in the Anglo-Saxon countries, Europe and Japan.

52 See also Allen and Gorton (1993); Allen (2001); Bank for International Settlements (2003); Stracca (2005).

53 While such action has been justified on the basis of ‘rational learning,’ this is unpersuasive given the amount of noise in short-term estimates of both absolute and relative manager performance.
outperformance by holding securities with strong momentum.\textsuperscript{54} The high turnover that reflects short-term performance chasing appears to have increased over time: since the 1950s, the average holding period for a U.S. equity mutual fund has declined from 16 years to around 3 years, while annual portfolio turnover has increased from 16 percent to more than 100 percent (Bogle, 2011). As Jin (2005) points out, the “output short-termism” of investment agents can largely be explained by the “heat” they feel from the “input short-termism” of principals. Ironically, the pressure from principals to redeem on rational arbitrageurs who are long (or overweight) cheap but underperforming securities and short (or underweight) expensive but outperforming securities will typically be most intense just when the mispricing (and hence, prospective return) is greatest.\textsuperscript{55}

The impact of earlier price changes can be amplified by other institutional practices such as window dressing and fund name changes. Since holding poorly performing securities can be difficult to justify to impatient asset owners, regular reporting intervals encourage agents to ‘window dress’ their portfolios so as to appear to have shown good judgment in selecting popular securities (Lakonishok and others, 1991; Chevalier and Ellison, 1997). In order to attract client inflows, mutual funds have also resorted to name changes so as to give the impression of a close association with a fashionable sector (Cooper and others, 2005).

DeMarzo and others (2008) show that if utility is derived not on the basis of one’s own wealth, but rather relative wealth, then investors will herd into what they knowingly perceive to be a bubble for as long as others do; relative wealth concerns make them unwilling to trade against the crowd. Institutional investors may also infer information from the prior trades of highly-informed and respected peers and invest in the same direction—they disregard their private information due to the reputational risk of acting differently from other managers (Scharfstein and Stein, 1990; Cai and others, 2012). ‘Synchronization risk’ can further dissuade rational arbitrageurs from pricking a bubble because each investor is uncertain about when others will begin selling: since a single investor cannot bring down the market, coordination among a vast pool of arbitrageurs is required, and this might prove elusive for a long period (Abreu and Brunnermeier, 2002, 2003). At a minimum, these considerations reduce the amount of non-consensus risk that rational arbitrageurs might be willing to take.\textsuperscript{56}

\textsuperscript{54} Short-term performance chasing has been established among mutual and hedge funds. For mutual funds, see Ippolito (1992); Chevalier and Ellison (1997); Sirri and Tufano (1998); Sapp and Tiwari (2004); Bailey and others (2011); and Lou (2012). For hedge funds, see Agarwal and others (2002, 2005); Baquero and Verbeek (2005); Dichev and Yu (2011).

\textsuperscript{55} Compounding matters for the rational arbitrageur will likely be margin calls (if leverage is employed) and possibly binding internal stop-loss or value-at-risk rules which force the manager out of a position at the worst possible time. Goyal and Wahal (2008) document that investment managers who lose a larger fraction of their assets tend to have higher post-termination returns.

\textsuperscript{56} See for instance, De Long and others (1990a); Scharfstein and Stein (1990); Maug and Naik (1996); Shleifer and Vishny (1997); Chevalier and Ellison (1999); Jin (2005); and Vayanos and Woolley (2013).
Indeed relative performance has been established as the most important factor explaining fund survival and the continued employment of portfolio managers. In particular, because investors have a quicker tendency to ‘fire than hire,’ avoiding underperforming one’s peers in the short-term is more important than outperforming in the long-term if one’s aim is simply to survive as an institutional investment manager.\textsuperscript{57} Feroli and others (2014, p. 10) emphasize that the high costs of underperforming peers introduces an important element of short-termism that generates negative spillover effects: “the concern of relative rankings in the payoff by the delegated agents injects an element of coordination in their portfolio choice that has the outward appearance of herding behavior.”

If principals appraise their agents on the basis of short-term performance vis-à-vis a momentum-biased benchmark, it can also be entirely rational—from the perspective of business and compensation risk—for agents to herd and knowingly participate in bubbles. Even if they suspect their benchmark to be overvalued, managers can take comfort in the belief they will be excused from poor absolute performance so long as their benchmark and peer group also perform poorly (Scharfstein and Stein, 1990; Rajan, 2005; Porter and Trifts, 2014). Herding during a bubble allows one to “share the blame” in the event of its subsequent implosion, in the same way that Keynes (1931) outlined the career risk associated with exercising prudent, non-consensus judgment: “A sound banker, alas, is not one who foresees danger and avoids it, but one who, when he is ruined, is ruined in a conventional and orthodox way along with his fellows, so that no one can really blame him.”\textsuperscript{58} The attendant distortion of agent incentives can lead to highly inefficient and socially undesirable portfolios (Brennan, 1993; Maug and Naik, 1996; Stein, 2004; Brennan and Li, 2008).

Even in the ‘absolute return’ (benchmark agnostic) hedge fund industry, aggregate returns can increasingly be explained solely by the standard benchmark index for world equities. The correlation of hedge fund and world equity returns has almost doubled on a trend basis over the past two decades (gray line, Figure 3), while the excess returns to hedge funds (those not explained by the world equity benchmark, alternatively known as ‘alpha’) has collapsed on a trend basis from a high of 17 percent in 1996 to just 2 percent in recent years (black line, Figure 3). As the composition of investors (principals) in hedge fund vehicles has evolved from largely benchmark-agnostic high net worth investors in the 1990s to benchmark-conscious institutional investors over the past decade, so too have hedge fund returns more closely mirrored returns to benchmarks.

\textsuperscript{57} Related studies of employment characteristics in the institutional investment management industry can be found in Khorana (1996); Chavalier and Ellison (1999); Brown and others (2001); Lynch and Musto (2003); Baquero and Verbeek (2005); and Porter and Trifts (2014).

\textsuperscript{58} In subsequent work, Keynes (1936, Chapter 12) spent considerable time addressing the institutional incentives to herd, including the well-known dictum, “worldly wisdom teaches it is better for reputation to fail conventionally than to succeed unconventionally.”
A key factor motivating herding by investment managers is that their performance is usually assessed, and fees paid, on an annual rather than multi-year basis—a time horizon where momentum tends to trump valuation as a driver of returns. Bubbles are often left unattacked by rational arbitrageurs operating open-ended funds (a fund structure favored by principals) because valuation-based arbitrage requires a long investment horizon that only closed-end funds can accommodate (Stein, 2004). Given that value-based strategies have often been found to take up to 3 to 5 years to generate meaningful outperformance (in contrast to momentum strategies which are most profitable with holding and look-back periods of 6 to 12 months), abstaining from bubble riding might simply entail too much business risk (Lakonishok and others, 1992; Lux, 1995; Dasgupta and others, 2011; Vayanos and Woolley, 2013). As Grantham (2012, p.1) observes from the perspective of a value-centric manager:

“The central truth of the investment business is that investment behavior is driven by career risk. The prime directive, as Keynes knew so well, is first and last to keep your job. To do this, he explained that you must never, ever be wrong on your own. To prevent this calamity, professional investors pay ruthless attention to what other investors in general are doing. The great majority ‘go with the flow,’ either completely or partially. This creates herding, or momentum, which drives prices far above or far below fair price. There are many other inefficiencies in market pricing, but this is by far the largest ... Ignoring (long-term fundamental value) may be the correct response on the part of most market players, for ignoring the volatile up and down market moves and attempting to focus on the slower burning long-term reality is simply too dangerous in career terms. Missing a big move, however unjustified it may be by fundamentals, is to take very high risk of being fired. Career risk and the resulting herding it creates are likely to always dominate investing.”
In a different context, the asymmetry associated with limited liability and incentives in delegated fund management can also impart a rational bias for bubble riding. The convex payoff structure associated with limited liability (i.e., limited downside, unlimited upside) can generate a preference for risk taking, selling disaster insurance and riding bubbles (Allen and Gorton, 1993; Allen and Gale, 2000; Rajan, 2005; Stein, 2013b).\(^{59}\) If fund managers participate in the formation of a bubble, they can earn large fees on both absolute performance and the management of a swelling asset pool. If the bubble deflates while they remain invested, managers generally do not have to repay the earlier performance fees, and will continue to earn management fees (albeit off a lower base).\(^{60}\) Alternatively, managers may chose to close down the existing fund and re-open another, thus lowering the hurdle level on which future compensation will be based (analogous to ‘re-striking’ a call option).\(^{61}\)

Finally, financial sector agents involved in the dissemination of information may also have an incentive to propagate partial truths (downplaying adverse information relative to market-friendly information) which can fuel bubbles, if their business model is based on transaction volumes which rise in a strong market. Prominent examples include deal originators, securities analysts, brokers and credit rating agencies. For instance, securities analysts might fear issuing downbeat assessments of the issuers they cover for any number of reasons (i.e., they could lose favor with management and be frozen out of future communication), and both the trading and investment banking business typically stand to benefit more from bullish than bearish recommendations due to the relative difficulties in monetizing downbeat recommendations.\(^{62}\) Further, since issuers of securities have both the ability and incentive to ‘ratings shop’ for the highest possible credit ratings, and rating agencies are overwhelmingly paid by issuers, rating agencies have an incentive to upwardly bias their credit assessments.\(^{63}\)

---

\(^{59}\) Brunnermeier and Nagel (2004) suggest hedge funds actively rode what they perceived to be a bubble in technology stocks in the 1990s. Jurek and Stafford (2011) argue that the selling of protection against low probability (but catastrophic) outcomes accounts for much of the measured alpha in hedge fund returns.

\(^{60}\) In this sense, the worst case scenario is a manager foregoes earning a performance fee as the bubble unwinds, until such time that cumulative returns recover back above some high watermark.

\(^{61}\) Some reputational risk is entailed in closing an underwater fund, only to reopen another shortly thereafter. But if performance has dropped so far below the previous high watermark (upon which performance-linked compensation will be based) and redemptions have also lowered assets under management (further increasing the percentage returns needed to restore fund value back to the high watermark), it may still be rational to do so.

\(^{62}\) In the United States, the Securities and Exchange Commission adopted Regulation Fair Disclosure in August 2000 in an attempt to stamp out selective disclosure.

\(^{63}\) Nationally recognized statistical rating organizations (NRSROs) operating under the issuer-pay (rather than buyer-pay) model account for approximately 99 percent of all outstanding NRSRO credit ratings in the United States. Of the ten NRSROs, only Egan Jones Ratings operates under the buyer-pays model. This emerged as a significant issue in the lead-up to the global financial crisis as structured financial products became significant contributors to the overall profitability of rating agencies (Segoviano and others, 2013).
(Irrational) Behavioral Models

The central issue in behavioral finance has been to explain why investors make systematic and irrational errors of judgment. Rationality presupposes that investors calculate the expected utility of each investment as a probability-weighted sum of utility outcomes, and then choose the option with the highest expected utility. But if investors are afflicted by cognitive biases and do not process information rationally, behavioral models may offer a better guide to understanding asset price dynamics (Thaler, 2005). As far back as the 1930s, Keynes (1936, p.144) asserted that rather than rational “mathematical expectation” forming the basis of asset price dynamics, it was instead a multitude of irrational emotional influences: “Our decision making … can only be taken as a result of animal spirits—of a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities.” Yet for many years behavioral insights into asset pricing vanished from view under the influence of the efficient markets hypothesis and its macroeconomic cousin, rational expectations.

Behavioral explanations for market inefficiency and speculative bubbles derive from the principal that, due in part to their evolutionary origin, expectational errors do not cancel out across investors—they are systematic rather than idiosyncratic. Market prices can become chronically distorted by widely shared, self-reinforcing misconceptions (Soros, 1987; De Bondt, 2012). Akerlof and Shiller (2009) argue the recurring nature of bubbles and panics through history suggests an irrational social-psychological explanation. Invoking the assertion by Keynes (1936) that “the facts of the existing situation enter disproportionately into the formation of our long-term expectations; our usual practice being to take the current situation and project it into the future,” Greenspan (2002, p. 6–7) noted that, “Investors too often exaggerate the extent of the improvement in economic fundamentals. Human psychology being what it is, bubbles tend to feed on themselves, and booms in their later stages are often supported by implausible projections of potential demand.” In their sweeping study of booms and busts spanning eight centuries, Reinhart and Rogoff conclude (2009b, p. 80): “The pervasive view that ‘this time is different’ is precisely why this time usually is not different, and catastrophe eventually strikes again.” The actions of investors falling prey to behavioral errors are typically painted as consistent with financially ruinous behavior.

---

64 Psychologists have documented a dichotomy between two modes of thought: System 1, which is fast, instinctive, intuition-based and emotional; and System 2, which is slower, more deliberative, reasoned and logical (Kahneman, 2011). In this context, much of the debate about whether investors are rational (and markets efficient) reduces to whether the actions of investors are most consistent with System 1 or System 2 thinking.

65 Homo sapiens are believed to have developed around 50,000 years ago, yet investors have been trading securities on listed exchanges for just 400 years. As such, the proposition that investors are not biologically ‘hardwired’ for optimal investment decision making would not seem to constitute a herculean leap. Lo (2004) presents an encouraging attempt at applying evolutionary principles to asset pricing applications.
In recent times, systematic departures from rationality have been established along two complementary lines in financial economics: experimental studies and investor surveys. Experimental studies suggest the rationality of judgment is bounded or constrained, and emphasize the strong behavioral tendency of investors to herd and display overconfidence after prior vindication. Separately, survey-based studies of investor return expectations reveal a strong tendency toward extrapolation of the recent past, with direct implications for speculative bubbles and the ‘momentum effect’ in asset prices.

Among the vast number of documented belief biases, the most important in the context of speculative bubbles relate to the tendency for market participants to herd, extrapolate, and fall victim to overconfidence. A typical process is as follows. The impact of a positive news shock to fundamentals generates a rational increase in asset prices. But because of anchoring and the tendency for investors to dispense with priors only slowly in the face of new information, the price increase is initially gradual and diffused, until such time the smooth upward trend attracts the attention of so-called ‘positive feedback traders.’ The high Sharpe ratio generated by the low-volatility uptrend begins to be extrapolated with an increasing number of positive feedback traders piling onto the trend (Cutler and others, 1991; Barsky and Delong, 1993; Lakonishok and others, 1994). With the serial correlation in returns now firmly in place, new entrants to the market with less than full confidence in their market timing ability opt to move with the herd in the belief that competitors must be acting upon superior information to their own (Scharfstein and Stein, 1990; Shiller, 2002). Reflecting expectations of a structural shift, economic commentators extrapolate the recent improvement in underlying fundamentals to persist into the indefinite future. As prices continue to rise, validation breeds overconfidence for traders and analysts alike. In certain circumstances, the positive price momentum can even reinforce the improving fundamentals that prices are supposed to reflect, such as a tightening in credit spreads lowering the interest charge (thus raising the profitability) for indebted companies. In this sense, investors can create their own sense of reality in a “reflexive” feedback process (Soros, 1987).

Behavioral errors are typically categorized in two ways: biased beliefs and nonclassical preferences, with the role of biased beliefs receiving most attention. Beliefs represent the probabilities we attach to certain outcomes, while preferences relate to how we make choices based on our attitudes toward risk (see Kahneman and Tversky, 1979). In turn, biased beliefs, or judgmental errors, come in two broad forms. The first comprises heuristic simplifications and shortcuts, including the availability, anchoring, and representativeness biases (the phenomena where we are more likely to overweight recent information, or that which is attention grabbing or easy to recall, or where we fall back on stereotypes). The second form of biased beliefs relate to self-deception, which includes our tendency towards mistaken overconfidence and the confirmation bias (seeking information that supports our prior). Studies have also documented that experiences early in life can exert a significant long-term impact on return expectations and risk attitudes (Malmendier and Nagel, 2011). The behavioral field has ballooned into a vast literature over the past two decades. Surveys include Thaler (1993, 2005), Barberis and Thaler (2003), Montier (2007), Kahneman (2011), De Bondt (2012), and Scherbina (2013).

This example draws on the initial under-reaction/subsequent over-reaction literature.
As serial correlation in asset returns has been documented across various asset classes and time periods, it is likely that positive feedback behavioral dynamics play at least some role in asset price determination over short to medium time horizons (Cutler and others, 1990; De Long and others, 1990b; Barberis and others, 1998, Hong and Stein, 1999). Unlike most other anomalies, there are no candidate explanations from the rational expectations school which do a sound job of explaining its pervasive existence; for this reason, momentum has been declared the “premier anomaly” in asset markets (Fama and French, 2008, p. 1653).68 Surveys of expected returns also reveal that investors tend to anticipate strong returns in the future when returns have been strong in the recent past—despite objective mathematical estimates of expected returns (yields) having been compressed by the same rapid price rises.69 In other words, after a boom in asset prices, subjective expectations of returns tend to be high while objective expected returns tend to be low (a finding established across stocks, bonds, exchange rates, and real estate).70 On balance, evidence from surveys of investors’ expected returns appears “directly at odds with the implications of standard asset price theory” (Williams, 2013, p. 6). Cochrane (2011, p. 1068) similarly suggests surveys of expected returns pose an “unsettling” challenge to rational market efficiency. More constructively, Barberis and others (2013, p. 1) conclude, “Survey evidence does not need to be seen as an inconvenient obstacle to understanding the market; on the contrary, it is consistent with the facts about prices and returns, and may be the key to understanding them.”

These two behavioral concepts—extrapolation and the negative correlation between subjective surveys and ‘objective’ yield-based measures of expected returns—are illustrated in Figure 4. The left hand side panel depicts the relationship between realized annual U.S. stock returns over the prior five years (vertical axis), and expectation of annualized returns from a survey of U.S. based Chief Financial Officers (horizontal

68 Much of the academic literature on momentum, ignited by the single stock findings in Jegadeesh and Titman (1993), has focused on cross-sectional momentum. Geczy and Samonov (2013) provide evidence of momentum in U.S. stocks from 1801 to 2012 in what the authors declare to be, “the world’s longest backtest.” Asness and others (2014) distill the stylized facts associated with the ‘momentum effect’ in equity markets, while recent cross asset class studies include Blitz and Van Vliet (2008), and Asness and others (2013). Yet time series (or absolute) momentum has equally strong ‘anomaly’ credentials (see Faber, 2007; Moskowitz and others, 2012).

69 Historically, survey data were often dismissed as noisy, unrepresentative, and overly-sensitive to language, framing and interpretation. But the benefits of survey data, such as timeliness and the fact they can be measured directly from actual investors rather than derived from a complex theoretical model (long after the fact), are gradually swaying researchers in this direction (particularly as survey techniques become more refined).

70 For stocks, see De Bondt (1993); Fisher and Statman (2000); Shiller and others (1996); Amromin and Sharpe (2008, 2012); and Greenwood and Shleifer (2014). In the case of bonds, see Bacchetta and others (2009). For exchange rates, see Dominguez (1986); Frankel and Froot (1987); Froot and Frankel (1989); and De Bondt (1993). For real estate, see Case and others (2012); Williams (2013); and Jurgilas and Lansing (2013).
These data suggest high (low) past returns are associated with high (low) expectations of future returns. The right hand side panel shows that an ‘objective’ earnings yield-based measure of expected returns (vertical axis) is negatively correlated to survey-based expectations of returns (horizontal axis): when objective measures of future stock returns are low, investors expect high returns in the future (and vice versa). Given the ease with which extrapolative expectations (including among experienced investment professionals) can lead to speculative bubbles, the broader findings from survey data that ‘extrapolative expectations trump rational expectations’ has potentially important policy implications.

Figure 4. Survey-based (Subjective) Expected Returns vs. Past Returns (left) and ‘Objective’ Expected Returns (right)

Source: IMF Staff
Notes: Survey return expectations for the U.S. stock market based on an average of 352 U.S. Chief Financial Officers, across 51 quarterly observations from September 2000 to March 2013, as reported in Graham and Harvey (2013). The earnings yield is shown as the inverse of the cyclically adjusted price-earnings ratio.

71 To demonstrate the concept that investors tend to extrapolate events of the recent past far into the future, the lookback (or reference) period is set to half of the long-term forecast horizon of ten years.

72 Results are robust to the use of the dividend yield instead of the cyclically-adjusted earnings yield.

73 The correlation between survey-based ten year expected returns and subsequent realized ten-year returns is -0.91, but +0.92 for the relationship between the objective earnings yield measure and subsequent realized ten-year returns. Given there were only 15 ‘out of sample’ observations on which it was possible to assess the accuracy of ten-year return projections, the precision of these estimates needs to be interpreted with caution. However the qualitative takeaways are entirely consistent with the aforementioned literature.
IV. Policy Implications

The four classes of models earlier discussed give rise to asset bubbles in different ways. Accordingly, the policy prescriptions that follow from these models vary widely. A brief summary is presented in Table 3.

Table 3. Mapping Policy Responses to Bubble Models

<table>
<thead>
<tr>
<th>Policy Prescriptions</th>
<th>Rational Measures</th>
<th>Irrational Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time-Invariant</td>
<td>&quot;Limits to Learning&quot;</td>
<td>&quot;Frictional&quot; Limits to Arbitrage</td>
</tr>
<tr>
<td></td>
<td>More accurate, timely and widely available information</td>
<td>More complete markets (more securities, easier ability to short, etc.)</td>
</tr>
<tr>
<td></td>
<td>Not relevant</td>
<td>Not relevant</td>
</tr>
<tr>
<td>Counter-Cyclical</td>
<td></td>
<td>Macroprudential policy in the case of a credit boom; monetary policy in the case of unjustifiably low risk premiums and/or a credit boom with systemic implications</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Full suite of of time-varying and time-invariant measures if irrationally low risk premiums pose systemic risk</td>
</tr>
</tbody>
</table>

Source: Author

Limits to Learning

The main policy implication from the limits to learning-based rational bubble model is that investors should have ready access to all the information they need in assessing investment opportunities; no time-varying countercyclical policy response is warranted. As Stout (1997) points out, rational investors have little need for a paternalistic capital market policy, but require only more accurate and timely information. Authorities should therefore concentrate their efforts on strengthening corporate reporting standards and prosecuting cases of fraud and the dissemination of misleading information. Bernanke (2002a) suggests that policy makers should focus efforts not on leaning against the wind, but rather in designing initiatives to “improve the financial literacy and competence of investors.” Kirchner (2009, p. 6) similarly argues that inefficiencies associated with limits to learning provide “an argument for improving the information content and completeness of markets, not for displacing or overriding them” with countercyclical policy settings.
While few would dispute the notion that more accurate, timely and readily accessible information constitutes a societal good, the evidence that it has or could in the future offer a robust defense against asset bubbles is unpersuasive. There is little to suggest that the incidence or severity of asset booms and busts has diminished over time, despite (i) a meaningful improvement in the information disclosure practices in the financial sector; (ii) technological advances in communication which have facilitated the rapid and widespread dissemination of information throughout society; and (iii) the growth in the relative share of financial assets managed by institutional investors, who can fairly easily overcome constraints in obtaining timely and accurate information. Shiller (2000) even suggests that the free flow of information—aided and abetted by the news media which he labels “generators of attention cascades”—which acts to intensify (rather than mitigate) self-reinforcing feedback loops, investor herding, financial instability and asset bubbles. Ensuring that information is more accurate, timely and widely available information would seem a natural place to begin, but not finish, a policy of ‘bubble defense.’

Frictional Limits to Arbitrage

If rational bubbles are largely the result of frictional limits to arbitrage, authorities should focus on fostering an environment in which product offerings and tradable securities (including hedging and shorting instruments) can accommodate the needs of investors with widely varying objectives and risk appetites. A system in which investors have a greater array of instruments through which to bear and diversify risk is clearly desirable, particularly in countries where economic efficiency is hampered by a lack of depth in asset markets. Tradable securities that are simple, transparent, have little embedded leverage, directly hedge economic risk (Shiller, 2003), and are amenable to trading on listed exchanges in the full glare of regulators and the investing public, have a valuable role in this regard. However, events leading up to and during the global financial crisis cautions that the introduction of complex, opaque and highly leveraged financial products traded on over-the-counter markets and designed primarily for speculative purposes (rather than the hedging of economic risk) can be more a hindrance than help to economic growth and financial stability. It is worth recalling that record levels of trading activity and capital market issuance prior to 2008 provided no insurance against the crisis. This suggests that while initiatives designed to ‘complete markets’ and eliminate frictions are welcome in principal, quality rather than quantity considerations need to be at the forefront of the financial deepening process.

74 The debate over whether marked-to-market accounting conventions contribute to or help ameliorate financial stability concerns is closely related.

75 For a description of the pronounced changes in both the quantity and quality of structured credit product offerings leading up to the crisis, see Šegoviano and others (2013).
Institutional Limits to Arbitrage

A broader suite of policy responses is likely to be required if, as I argue in this paper, incentives in the institutional investment management industry are a key factor in the growth and persistence of asset bubbles. While time-varying responses (macroprudential and monetary policy) are applicable tools once a bubble is underway, preventing their emergence in the first instance will likely necessitate time-invariant measures which alter the underlying incentives of institutional investment managers. In this context, reforms to contract and financial benchmark design could well be the natural place to start.

- **Contract design.** Reform proposals should reflect the notion that agents respond to the incentive structure imposed on them by principals. It is unreasonable to expect investment managers to act countercyclically (in a manner that appeals to the broader public good) when their incentive structure can instead reward, or at a minimum, fail to discourage, short-term herding behavior. A number of issues could be considered in the formulation of principal-agent contract design:

  - A stronger emphasis should be placed on long-term performance appraisal, including incentives for countercyclical investment (for instance, by weighting manager performance more heavily in difficult market conditions). This emphasis would encourage agents to lean against, rather than with, large mispricings (which can take considerable time to correct), and also better align the interests of agents and principals with the public good;
  - To discourage redemptions over short time horizons, investment fund exit sales loads could be formulated on the basis of a sliding time scale where loads reduce as a function of investment holding period. Current practice (which makes no contribution to financial stability) is that sales loads are staggered exclusively on the basis of account size;
  - Manager performance should be assessed in ways that do not exclusively involve price momentum-biased benchmarks. For instance, pension and insurance funds could place more emphasis on the degree to which a manager hedges their liability structure and thus their true source of economic risk;
  - The asymmetry that arises in agent compensation (where agents own the upside but not downside) could be addressed with more thorough use of multi-year clawback provisions (such as those now in effect in the banking industry) and the use of high watermarks;
  - In the case of financial information providers whose business model thrives on issuing upbeat assessments of issuer prospects (i.e., brokers and credit rating agencies), there should be stronger emphasis on accountability.\(^6\)

\(^{6}\) In June 2013, the European Commission enacted new provisions to ensure credit rating agencies can be held liable by investors for damages in the event their analysis is found to be grossly negligent.
• **Benchmarks.** Initiatives currently aimed at reforming financial benchmarks (in the wake of interest rate and foreign exchange fixing scandals) could be broadened to examine the financial stability implications of price-based market capitalization benchmarks. These impart a procyclical bias to financial markets by forcing benchmark-constrained managers to increase (reduce) holdings of securities which have recently shown strong (weak) price momentum, irrespective of fundamentals. In the case of debt markets, market-capitalization based benchmarks also force captive buying of issuers that have recently increased the amount of tradable debt outstanding. Alternative non-price based benchmark formations, for instance where objective measures of risk premia (Jones, 2013) or company fundamentals (Arnott and others, 2008) form the basis of constitute weights, can mitigate herding and be less prone to excessive concentration arising from market fads which result in heavy exposure to individual companies, sectors or countries. By forcing rebalancing to be conducted in a countercyclical manner, these approaches raise the prospect of ameliorating (rather than contributing to) financial instability, and possibly generating superior longer-term returns compared to price-based benchmarks (Figure 5).

---

**Figure 5. Relative 10-year Annualized Outperformance of Fundamental-based Indices vs. Industry Standard Market Capitalization-based Indices**

Source: IMF Staff, Research Affiliates

Notes: Average annualized return differential between fundamental-based indices and industry standard market capitalization-weighted benchmarks. Fundamental Index constituents are weighted using a composite of fundamental factors, including cash payouts, free cash flow, sales and book value of asset (prices and market values are not determinants of index weights). The 10-year performance assessment period is up to July 2014.

---

77 From a broader financial stability perspective, automatic rating triggers that induce fire sales of downgraded assets should also be used with much care. Many investors use a single rating threshold to determine the eligibility of issuers and counterparties. A rating corridor with different thresholds for investing and divesting reduces the probability that a portfolio will have to be adjusted excessively in a procyclical fashion. Better still would be to use a rating change as a trigger for reviewing the risk and expected return of the entire portfolio, rather than simply inducing an automatic asset fire sale or purchase (Papaioannou and others, 2013).
• **Transaction Taxes.** Securities transaction taxes (STT) in various guises have long been proposed to both discourage speculative investment and increase public revenues in a fair and progressive manner.\(^{78, 79}\) While lower transaction costs are undeniably welcome in markets that efficiently intermediate long-term savings and investment, the case becomes less clear in markets dominated by short-term herding and speculation. Woolley (2010, p. 133) suggests that in inefficient markets, “lowering the frictional costs of trading opens the door to short-termism and momentum trading which distorts prices … both contribute to the loss of social utility.”\(^{80}\) However, all STT will impede price discovery to some degree by discouraging new information from being incorporated in prices, thus reducing the informational efficiency of markets and raising the cost of capital for issuers. Rogoff (2011) argues the latter impact will in turn lower corporate investment (with no concomitant increase in financial stability) and thus prove counterproductive. More specifically, a blanket STT does not distinguish motives or holding periods (i.e. short-term momentum-based speculation from long-term countercyclical investment or standard portfolio rebalancing) and so is a crude response to mitigating the threat of asset bubbles. Nor is there empirical evidence to suggest that such measures reduce short-term price volatility (see Habermeier and Kirilenko, 2003; and Matheson, 2011).\(^{81}\) Taxes based on holding periods and the frequency of portfolio turnover are more directly targeted toward discouraging short-term speculation (for instance, taxes can be set as a declining function of holding period, and a rising function of portfolio turnover).\(^{82}\) Taxes on derivatives (which have built-in leverage) that are used for non-

---

\(^{78}\) Upon the request of G-20 leaders following the September 2009 Summit, the IMF (2010) examined the costs and benefits of various financial sector taxes. It recommended: (i) the adoption of levies on financial institutions to pay for the resolution of troubled institutions in the event of future failures and crises, and (ii) the use of a “financial activities tax” levied on the sum of financial institutions profits and wages to be preferable to a financial securities transaction tax. A financial transactions tax proposal supported by eleven EU member states was approved in the European Parliament in 2012, and by the European Union Council in 2013.

\(^{79}\) The intellectual basis for STT dates back to Keynes (1936, p. 160) who argued, “It is usually agreed that casinos should, in the public interest, be inaccessible and expensive. And perhaps the same is true of stock exchanges … a tax on all transactions might prove the most serviceable reform available, with a view to mitigating the predominance of speculation over enterprise.”

\(^{80}\) The utility value of high frequency trading is now an active area of investigation for policy makers.

\(^{81}\) On the implementation of securities financial transactions, Brondolo (2011) suggests that while taxing exchange traded instruments would be relatively easy and efficient, taxing over the counter transactions would prove more costly and challenging. According to Matheson (2011), estimates of the revenue potential of a low-rate (0.5–1 basis point) multilateral transaction tax on the four major trading currencies suggest that it could raise about $20–40 billion annually, or roughly 0.05 percent of world GDP. A one basis point STT on global stocks, bonds and derivatives is estimated to raise approximately 0.4 percent of world GDP.

\(^{82}\) For over thirty years the UK tax statutes have contained a clause withdrawing tax exemptions for any foreign fund deemed to be “trading” rather than “investing,” though Woolley (2010) points out this designation has rarely been implemented.
hedging purposes could be more thoroughly examined, though instruments that explicitly target leverage, such as higher margin and collateral requirements, and more generally, the removal of tax subsidies on corporate and financial sector balance sheet debt, could be more effective in addressing leverage and procyclicality (Matheson, 2011). Many tax codes allow interest payments to be deducted as a business expense, yet there is no equivalent allowance for payments to equity investors. The application of tax breaks on debt but not equity is counterintuitive from the perspective of financial stability.

- **Time-varying Macroprudential and Monetary Policies.** If incentives and principal-agent frictions in the institutional investment industry give rise to unjustifiably low risk premia and sharp increases in leverage which could pose a systemic threat to economic stability, macroprudential policies and tighter monetary policy will have a role to play in warding off bubbles. Given they can be highly targeted, macroprudential measures should offer the first line of defense. But just as monetary policy can contribute to easing financial conditions by compressing risk premia—a central aim of unconventional monetary policy in recent years—logic dictates that monetary policy can (and should) also lean in the opposite direction when circumstances warrant (tighter monetary policy is the only instrument that can instantaneously “get into all the cracks”). Nevertheless my reading of the evidence suggests that in the absence of fundamental reforms to incentives, contracts and benchmarks, both macroprudential and monetary policies will be left to deal with the symptoms (rather than underlying causes) of deeper flaws in the financial system. If the institutional contribution to asset bubbles is left unaddressed, and the institutional investment industry continues to grow in absolute terms and as a share of total investment activity, this raises the unappealing prospect of indefinitely recurring waves of boom and bust in the future.

**Behavioral Models**

Behavioral models offer tremendous promise in enriching our understanding of various aspects of economic life. And survey data strongly suggests that policy makers should stay attuned to the possibility of large scale behavioral errors in the pricing of assets. To the extent that (i) behavioral errors combine with other explanations to amplify bubbles, and (ii) there is good reason to suspect that a reversal in irrationally low risk premia could exert a large negative impact on the economy, a full suite of policy responses (the proverbial “kitchen sink”) may be required. These would include the aforementioned time-invariant

---

83 Countercyclical adjustments in stamp duties on residential property have also been deployed where exchange rate arrangements constrain monetary policy (for instance in countries such as Singapore and Hong Kong), though the practical feasibility of widespread cyclical changes in taxes could be challenging and has not been rigorously examined.
measures to address information asymmetries, the incompleteness of markets, and incentives and principal-agent frictions in the institutional fund management industry. If fickle and irrational behavior is the driving force behind asset prices, it also raises questions over whether entities should be able to lever up or consume out of unrealized capital gains. On top of such measures, an even stronger emphasis on countercyclical macroprudential and monetary measures would likely be required in order to temper the impact of extrapolation and overconfidence errors that can characterize late cycle bubbles.

Nevertheless, when viewed in isolation, behavioral models of widespread irrationality tend to leave open key issues. For instance: if all investors are irrational due to our biological hardwiring, who is on the other side of the irrational trading activity? If widespread irrationality leads to financially ruinous behavior, shouldn’t the process of natural selection wipe out such investors, and thus bias markets toward a more rational and efficient state over time? If irrationality is so pervasive, shouldn’t it be straightforward to outperform the market? Finally, how is it possible to defend against the same criticism leveled at ‘rational’ explanations of asset pricing that the pliability of behavioral finance is both its key strength and Achilles heel? Because behavioral explanations have proliferated in recent times, they can be retrospectively selected (from what is now a large menu) to justify nearly any empirical finding.

On balance, these considerations suggest that widespread investor irrationality is unlikely to be the only factor at play in the formation and persistence of asset bubbles.

V. CONCLUDING REMARKS AND FUTURE RESEARCH

By process of deduction, this paper seeks to offer an explanation for why asset bubbles have posed—and will likely continue to do so—a threat to economic stability despite financial markets being characterized by more complete information, a greater array of securities through which to express views, and a more pronounced impact of sophisticated institutional investors, than ever before. The arguments advanced here suggest other candidate explanations for the persistence of bubbles, such as limits to learning, frictional limits to arbitrage, and behavioral errors, are unsatisfactory (by themselves at least) as they are inconsistent with the aforementioned trends sweeping across global capital markets. By contrast, investment manager incentives, the nature of the principal-agent relationship, and the growing presence of institutional investors, are all entirely consistent with the persistence

---

84 The most enduring empirical finding in support of the efficient markets hypothesis has been the difficulty encountered by active investment managers in consistently outperforming market-capitalization based benchmarks. This result has held, almost uncontested, ever since the earliest studies of Sharpe (1966) and Jensen (1968). To the extent that persistence in active manager excess returns has been documented, much of it can be explained away by systematic exposure to standard risk factors (see for instance, Carhart, 1997).
of financial bubbles. Importantly, this explanation does not require a baseline assumption of widespread irrationality in the conventional sense. Simply put, it can be entirely rational—from the perspective of business and compensation risk—for asset managers to knowingly ride bubbles because of benchmarking and the short-term performance appraisal periods often imposed on asset managers by asset owners.

To the extent that this diagnosis is close to the mark, it has potentially important policy implications. First, it suggests that bubble episodes will be at least as frequent as in the past, and quite possibly more so, as institutional assets under management continue to increase in both absolute and relative terms (particularly in the emerging markets).

Second, it suggests remedial policies need to be multifaceted in nature. While countercyclical monetary and macroprudential policy may be best placed to lean against conventional leverage-driven asset booms (notably in real estate markets), they are not particularly well suited to dealing with the challenges posed by the rapidly growing asset management industry. Contemporary discussions of financial instability are heavily conditioned by the build up of leverage that culminated in the 2008 crisis, and understandably, measures aimed at avoiding a similar scenario in the future have been front and center in subsequent policy initiatives. However, subdued leverage is not a sufficient condition for financial stability—if systemic risk, and activity in the wider economy, is shaped importantly by large shifts in risk premia owing to the ‘rational herding’ motivations of asset managers (even in the absence of leverage), then this traditional focus may be too narrow (Feroli and others, 2014; Stein, 2014; Haldane, 2014). Moreover, as risk-taking migrates out of the formal banking sector, policy makers must guard against the risk of ‘fighting the last war.’ Put differently, while they will always comprise a key component of the policy making toolkit, there are likely to be limits to what countercyclical policies can achieve alone. Time-invariant policies related to the design of principal-agent contracts and financial benchmarks—addressing the cause, not simply the symptoms, of institutional behavior—have a key role to play in mitigating the impact of institutional frictions on financial stability.

More broadly, while the behavior and failure of banks has been studied for centuries, similar analysis of the institutional asset management industry is, by contrast, a greenfield site—yet the risks and opportunities presented by asset management could be every bit as important (Haldane, 2014). Whether it be asset management industry characteristics (along the dimensions of size, concentration, and interconnectivity), or asset management activities (duration mismatches, securities lending, fire sales and herding), there are a variety of channels by which the industry could pose a threat to financial stability. Care will need to be taken in distinguishing the degree to which asset managers passively transmit risk (in passing through the decisions of asset owners) as distinct from originating new sources of risk themselves. Nonetheless, if present indications are any guide, these will serve as fertile grounds for future research aimed at safeguarding the international financial system.
References


Bogle, John C., 2011, Don’t Count on It! (New Jersey: John Wiley & Sons, Inc.).


Krugman, Paul, 1993, “What Do We Need to Know About the International Monetary System?”, *Essays in International Finance*, No 190, (Princeton, NJ: International Finance Section, Department of Economics, Princeton University)


