Economic Perspectives
Number 3, 2025
https://doi.org/10.21033/ep-2025-3

Why bubbles occur: Revisiting the rationality debate

Gadi Barlevy, senior economist and economic advisor

Introduction and summary[1]

In 1978, economic historian Charles Kindleberger published his seminal book *Manias, Panics, and Crashes: A History of Financial Crises*. The book chronicled the history of financial crises from the birth of modern financial markets to the present era. Kindleberger pointed out that many of these financial crises occurred in the wake of a rapid surge in the price of some asset category that eventually reversed—a scenario commonly known as an asset bubble that inflated and then burst. Financial intermediaries that either invested in these assets or lent to those who did incurred large losses when asset prices fell. These intermediaries then cut back on the credit that producers and consumers rely on to finance their spending, leading to contractions in economic activity until the financial sector recovered.

In referring to these episodes as manias and panics, Kindleberger (1978) reflected a view that asset boom-and-bust episodes are rooted in market psychology. According to this view, some traders get swept up in the initial euphoria when asset prices start to rise. These traders drive asset price growth even higher, until eventually asset prices slow down and traders panic and dump their asset holdings. Some of these historical episodes featured mind-boggling price movements. For example, Frehen et al. (2013) documented that the share prices of the Mississippi and South Sea companies increased about tenfold in mere months in 1719 and 1720, respectively, followed by a collapse of between 70% and 90% of the peak value over several months.

The notion that asset booms and busts are driven by irrational beliefs is widespread. Consider the following description from economist Eugene Fama's 2014 Nobel Prize lecture:

> The stock market run-up to 2007 and subsequent decline is often called a "bubble." Indeed, the word "bubble," applied to many markets, is now common among academics and practitioners. A common policy prescription is that the Fed and other regulators should lean against asset market bubbles to preempt the negative effects of bursting bubbles on economic activity.
>
> Such policy statements seem to define a "bubble" as an irrational strong price increase that implies a predictable strong decline. This also seems to be the definition implicit in most recent claims about "bubbles." (Fama, 2014, p. 1475)

Some support for the view that irrational behavior is an important driver of these boom-and-bust dynamics came from early theoretical research on asset bubbles. As I explain in more detail later, economic theorists typically define bubbles not as boom-and-bust episodes but as situations in which the price of an asset is objectively too high compared with the asset's intrinsic worth. Specifically, economists define a bubble as a case in which an asset price exceeds the discounted value of the cash flows that the asset is expected to pay out over its lifetime. Overvaluation and boom-and-bust dynamics are distinct but related phenomena. An asset price might end up too high after the price surges. Conversely, an asset whose price is too high might be vulnerable to a crash if there is a risk that traders switch to an equilibrium in which the price of an asset is equal to the intrinsic value of the asset.

Early theoretical work on the possibility of asset bubbles found that overvaluation (and undervaluation) can be ruled out in various models where all agents are rational. Subsequent research demonstrated that overvaluation is in fact possible in economies populated by rational agents, but only under certain conditions. However, these scenarios tend to be fragile and rely on empirically questionable assumptions.

At the same time, models in which at least some of the traders in asset markets hold irrational beliefs did seem able to generate robust and empirically plausible examples of overvaluation. Such models include settings in which agents are overconfident about the precision of the information they receive and settings in which agents extrapolate from recent past returns and believe that future returns will be similar to those that they experienced before. The conclusion that many drew from these findings is that overvaluation might be more plausible when at least some agents are irrational.

In this article, I propose a different interpretation of these results: I argue that rather than a contest between models with exclusively rational agents and models where at least some traders hold irrational beliefs, the theoretical literature on overvalued assets is better understood as a contest between models where rational agents have symmetric information and models where agents—whether rational or irrational—hold different beliefs from one another. Early work arguing that asset bubbles are either impossible or exist only under empirically implausible assumptions when agents are rational also assumed these rational agents share the same information. Later research on models in which rational agents hold different beliefs showed that these models can give rise to empirically plausible and more robust overvaluation. Moreover, the bubbles that arise in models where rational agents hold heterogeneous beliefs mirror the bubbles that arise in models where some agents hold irrational beliefs.[2]

The fact that agents hold different beliefs from one another does not mean that they cannot all be rational. Fully rational agents might be exposed to different sources of information that lead them to hold different beliefs. Alternatively, rational agents facing scenarios that are outside the scope of what they experienced in the past might start with subjective beliefs about their environment, and there is no reason that rational agents facing environments they have no experience with must start out with identical subjective beliefs. Rationality does impose that agents should update their beliefs consistently with Bayes' rule[3] in light of what they observe, and it may imply that agents must eventually converge on the same correct beliefs. But rationality does not require that agents must agree with one another at each and every date.

When agents hold different beliefs from one another, even if only temporarily, they may have an incentive to pay more for an asset than they believe it is worth if they also believe that they might be able to sell the asset later for an even higher price to others holding different beliefs. This gives rise to a different type of bubble than those that arise with rational agents who all share identical beliefs. In the latter case, a bubble can only occur if trading an overvalued asset can benefit all agents. By contrast, the bubbles that can arise when rational agents hold different beliefs allow agents to expect to profit at the expense of others, which is more similar in spirit to the bubbles that arise when some traders hold irrational beliefs.

In the remainder of this article, I review the relevant literature to explain why the key implications of some of the early work on bubbles in models where all agents are rational should really be understood as implications of models where agents hold the same beliefs. I begin by discussing how economists define asset bubbles. I then review some of the key results from early work on rational bubbles. I discuss the limitations of these models and how these limitations are addressed in models where agents hold irrational beliefs. I then argue that models where agents are rational but hold heterogeneous beliefs can also address these limitations. I conclude with some remarks on the importance of rationality given that similar bubbles can arise whether agents are rational and hold different beliefs or whether agents hold irrational beliefs that lead them to expect to earn profits from trading assets.

Defining a bubble

A natural starting point for discussing the theoretical literature on bubbles is how to define the term *bubble*. Kindleberger (1978) defined a bubble as a rapid and unsustainable increase in the price of an asset that eventually implodes. That definition is motivated by the various episodes of speculative frenzies he cites in his book in which asset prices surged and then collapsed.

While Kindleberger's (1978) definition captures the key features of episodes that are typically described as bubbles, it raises several issues. The fact that the episodes in his book featured rapid asset price surges that were largely reversed does not prove that the original price increases were unsustainable. Asset prices may have increased when one shock raised demand for assets and then fell when a different shock lowered demand for these same assets. Defining the increase as unsustainable essentially imposes a particular interpretation of how these episodes must have unfolded. More generally, it is not obvious that asset booms must give way to busts. Greenwood et al. (2019) combed through data on industry portfolios in the United States between January 1928 and March 2012 to isolate situations where both the raw return and the net after-market return exceeded 100% over a two-year period. To eliminate cases in which industries were recovering from poor previous performance, they further restricted attention to cases in which cumulative returns in the previous five years exceeded 50%. These criteria identified 40 booms. In 21 of these cases, the price of the portfolio fell by at least 40% within two years of the surge. In the remaining 19 episodes, asset prices continued to rise, with an average return of 21% in the next year and 46% in the next two years. The notion that large and rapid asset price increases are inherently unsustainable is thus far from obvious.

Given these issues, economists have focused on a different definition of bubbles from the one Kindleberger proposed. This alternative definition focuses not on asset price dynamics but on whether an asset price is objectively too high (or, in principle, too low). By this definition, a bubble is a situation in which the price of an asset differs from the present discounted value of the dividends that the asset is expected to pay out over its remaining lifetime. The latter is commonly known as the fundamental value of the asset. Intuitively, the value that an asset provides society as a whole—without regard to who owns the asset—derives from the dividends it pays out. The present discounted value of dividends that an asset yields over its lifetime thus seems like a natural benchmark for the price at which an asset should trade. Deviations from this value are of independent interest from the question of whether an asset's price might collapse if the price were somehow too high. For example, a price above fundamentals can incentivize the wasteful creation of assets that cost more to produce than the value these assets provide to society. Conversely, a high price suggests that society has managed to create value out of thin air without needing to back it with dividends, implying society as a whole is richer because of the bubbles that were created.[4]

While overvaluation is a distinct notion from boom-and-bust dynamics, the price of an overvalued asset may be particularly vulnerable to a collapse. This is because the fundamental value of an asset is often a natural candidate for the price at which the asset might trade in equilibrium, and an exceptionally high price (or low price) may prove temporary before the price switches to the fundamental value. That said, the notion of overvaluation does not presuppose that the asset price must rise rapidly or collapse quickly. Consistent with this, Hirano and Toda (2025) derived conditions under which asset bubbles must occur in equilibrium and a bubble cannot disappear. Given the finding in Greenwood et al. (2019) that many booms do not end in busts, it seems reasonable to define bubbles in a way that does not presuppose that the price of an overvalued asset will necessarily collapse. At the same time, since one of the reasons for studying bubbles is to understand historical asset boom-and-bust episodes, models in which overvaluation can be associated with a collapse are of inherent interest.

While the definition of a bubble is easy enough to state, inferring the fundamental value of an asset in practice is often difficult, if not downright impossible. Consider the present discounted value of cash flows for a residential property at a particular location. The cash flows from such an asset are the rents that the owner of the property can charge whoever uses the property. But it is hard to predict the prevailing rent payments at a given location in 20, 50, or 100 years in light of how evolving circumstances might dictate where people will want to or be able to live. Nevertheless, the concept of a bubble can be useful even if a bubble cannot be directly observed. There are various economic analyses that similarly focus on unobservable quantities. For example, a key concept in many models of monetary policy is *potential output*, or the output that would be produced if goods prices were fully flexible instead of being rigid as empirical evidence suggests they are in practice. While potential output is not directly observable and central banks routinely grapple with the question of what the level of potential output might be, the theory of what policy would be appropriate if the economy were above or below potential is viewed as a useful guide for monetary policy implementation. Similarly, models that give rise to bubbles can offer guidance on what policymakers should do if they thought that an asset might be overvalued without quite being sure.

The impossibility of bubbles in simple models with rational agents

After settling on a definition for bubbles, I now turn to some of the early theoretical results on the impossibility of bubbles in settings featuring rational agents. Tirole (1982) collected some of these results and derived several new ones. The first scenario he discussed in which bubbles can be ruled out involved models with a finite number of trading periods, meaning there is a last period beyond which the asset will not trade.

The argument for why a bubble can be ruled out in such settings is as follows. If there were a last period in which the asset could trade, the only reason to buy the asset in that final period is to collect the dividends that this asset will pay out over its remaining lifetime. No rational agent would pay more for the asset than the present value of these dividends. The asset thus cannot be overvalued in the last period in which it can trade. This same restriction also limits what agents would pay for the asset before the final trading period, since if the price of the asset in the last period cannot exceed the present discounted value of remaining dividends, buying the asset before the last period and selling it in the last period yields the same expected payoff as buying the asset and holding it indefinitely. The asset will never be overvalued in equilibrium.

I can flesh out this intuition using a simple deterministic economy with no uncertainty. Assuming a deterministic economy simplifies the argument but is not essential. For reasons I discuss later, the feature that rules out bubbles is not the absence of uncertainty, but that all agents agree about the asset and value it equally. In a deterministic setting, there is nothing for agents to disagree about, which contributes to them valuing the asset equally.

Suppose there are only two assets that agents can hold: 1) one-period debt that pays a constant return of r each period and 2) a long-lived asset that pays a stream of dividends over time. Let d_t denote the value of dividends in period t. Absent uncertainty, agents can perfectly anticipate d_t. Let p_t denote the price of the asset at date t and let T denote the last period of trade. While the asset cannot trade past date T, it can keep paying dividends past T.

At date T, any agent who buys the asset knows that they will have to hold the asset over its remaining lifetime. The interest rate r tells us how agents should trade off payments over time: One unit of wealth today can be saved and converted into $(1 + r)^t$ units of wealth in t periods. Since the present discounted value of dividends corresponds to the fundamental value of the asset, I denote this value by f_t. The fundamental value at date T is given by

$$f_T = \sum_{s=1}^{\infty} \frac{d_{T+s}}{(1+r)^s}.$$

No rational agent would be willing to pay more than f_T to buy the asset at date T. If the price p_T were higher than f_T, an agent could take the amount p_T that they would have used to buy an asset at date T and use it to buy debt instead. After one period, this would give them $(1 + r)\, p_T$, which exceeds $(1 + r)\, f_T$ when $p_T > f_T$. The expression for f_T implies

$$(1+r)\, f_T = d_{T+1} + f_{T+1},$$

where $f_{T+1} = \sum_{s=1}^{\infty} \frac{d_{T+1+s}}{(1+r)^s}$ denotes the fundamental value of the asset at date $T + 1$. Buying debt would thus leave the agent with more resources after one period than if they held the asset. The agent could consume this extra amount, leaving themselves with wealth f_{T+1}, which they can keep using to buy debt and ensure the same flow of payments (d_{T+2}, d_{T+3}, …) that they would have received after date $T + 1$ had they bought the asset. No agent would be willing to hold the asset at date T given this alternative, so the supply of the asset will exceed demand for it.

At the same time, the equilibrium price p_T cannot be lower than the fundamental value f_T. If this were so, agents who buy the asset could make themselves strictly better off, since the dividends offered by the asset yield a higher payoff than holding on to the amount needed to buy the asset and using it to buy debt. This means that anyone who holds the asset would keep holding on to it while others would try to buy the asset, and demand for the asset at date T would exceed supply. It follows that the asset price at date T must equal the fundamental value at date T.

Moving back to period $T - 1$, I can similarly establish that the price must equal fundamentals in this period. An agent who buys the asset at date $T - 1$ will receive dividends d_T at date T and must then decide whether to keep the asset or sell it. Since $p_T = f_T$, agents should be indifferent between keeping the asset and selling it at date T. This means that the most a rational agent would pay for the asset in period $T - 1$ is f_{T-1}, since any agent who buys the asset that period should be willing to hold the asset indefinitely. Using a similar argument as for date T, I can rule out the case where the price is below the fundamental. Hence, $p_{T-1} = f_{T-1}$. The same logic can be applied all the way back to period 0. The asset price will not exceed the fundamental value of the asset in any period.

Impossibility of a bubble in a setting with rational agents and infinite trading horizons

The argument for the impossibility of bubbles with rational agents in the finite horizon setting is based on no rational agent being willing to overpay for the asset in the last trading period. But what if there was no last trading period? While one would not be able to apply backward induction, other arguments could be used to rule out the possibility of a bubble in this case, at least under certain conditions. Once again, a key assumption in the argument is that all agents agree about the asset and value it equally.

The argument is somewhat technical. Here I provide a heuristic sketch that gives the essential idea. Consider the same environment as in the finite horizon case in which rational agents can trade only two assets and there is no uncertainty. In equilibrium, the two assets must offer the same return in equilibrium for agents to be willing to hold both. This implies

$$1+r=\frac{d_{t+1}+p_{t+1}}{p_t}.$$

I can rearrange this equation as

$$\left(1+r\right)p_t=d_{t+1}+p_{t+1}.$$

Next, since the fundamental value is just the present discounted value of dividends, I have

$$f_t=\frac{d_{t+1}+f_{t+1}}{1+r}.$$

This equation can be rearranged to read

$$\left(1+r\right)f_t=d_{t+1}+f_{t+1}.$$

I define the bubble component b_t as the difference between the price and the fundamental, that is, $b_t = p_t - f_t$. Taking the difference between the equations that govern p_t and f_t, respectively, reveals that in equilibrium, the bubble term must satisfy

$$b_{t+1} = (1 + r)b_t.$$

Hence, the bubble term will grow at the same rate as the interest rate. Assuming the interest rate r is positive, this means that the bubble will grow without bound over time. If an agent held a bubble asset indefinitely, they would amass a growing amount of wealth. Even if an agent kept selling the asset and buying it back so they held the asset infinitely often but not every period, they would still need to amass wealth without bound so they could keep affording to buy the asset. Accumulating wealth this way cannot be optimal: Since wealth yields no direct utility, a household would be better off selling off some of the asset and increasing their consumption. Formally, this argument is known as the transversality condition for optimality in infinite horizon maximization problems. As long as the marginal utility of consumption is positive, the discounted wealth that agents optimally accumulate should asymptotically converge to zero.

Given that someone must hold the asset each period, if the number of agents in the economy is finite, there will necessarily be an agent who will hold the bubble asset infinitely often. Such an agent would violate their transversality condition. At the same time, equilibrium requires that all agents act optimally. This means that a bubble cannot be an equilibrium. For a more formal discussion of when and how the transversality condition rules out bubbles, see Kamihigashi (2006).

Intuitively, rational agents would naturally be reluctant to pay more for an asset than its fundamental value if they intended to hold on to that asset indefinitely. In both finite horizon settings and infinite horizon settings with a finite number of agents, there must be an agent who either buys the asset and never sells it or who holds the asset infinitely often. Neither of these would be in their best interest when the asset is overvalued. However, as I next explain, bubbles can arise in alternative settings where there are either infinitely many agents or where agents do not all value the asset equally as in the two examples I just discussed.

Rational bubble models

To appreciate why bubbles might be possible even when all agents are rational, note that while rational agents would refuse to pay more than the fundamental value of an asset to buy and hold that asset indefinitely, they might be willing to buy an overvalued asset if they expected to sell it. This point was emphasized by Blanchard and Watson (1982). The observation that it can be rational to buy an overvalued asset to sell it for a higher price later is sometimes described as the *greater-fool theory* of bubbles: Agents are willing to do something foolish such as paying more for an asset than it is inherently worth if they believe that there is some greater fool who is willing to buy the asset for an even higher price. When the trading horizon is finite or when the trading horizon is infinite but there are finitely many agents, the existence of a bubble requires that there be some ultimate fool who agrees to hold the asset indefinitely, even though this is not in their best interest. No rational agent would agree to do so. Thus, agents cannot all reasonably expect to profitably sell the overvalued asset under these circumstances—and a bubble will not exist.

However, there are scenarios in which rational agents can all reasonably expect to profitably sell a bubble asset after buying it—that is, where there is no ultimate fool who holds the asset indefinitely against their own interest. Blanchard and Watson (1982) referred to such scenarios as *rational bubbles*. This terminology deliberately sought to distinguish these scenarios from explanations of bubbles rooted in market psychology. In rational bubble models, traders who buy overvalued assets are fully cognizant that the price of the asset they are buying exceeds the underlying fundamental value of the asset and are not swept up in a mania about the asset. Rather, these agents intend to sell the overvalued assets they buy, and as such, they are willing to buy these assets despite knowing full well their intrinsic value.

One situation in which a bubble can arise with rational agents is if the population grows over time and new agents keep entering the asset market. With a growing population, it is no longer the case that there must be some agent who holds the asset infinitely often. That is, if the number of agents who can ever trade the asset is infinite, it will be possible for every agent to hold the asset for a finite period and sell it to someone who has not held the asset before.

Tirole (1985) formalized this argument using the overlapping-generations model introduced by Samuelson (1958) and modified by Diamond (1965) to incorporate production. The overlapping-generations model features an infinite sequence of cohorts, each of whom lives for a finite number of periods. Because agents from different cohorts overlap, old agents can always find young agents to whom they can sell their assets. In this setting, an intrinsically worthless asset that pays no dividends can sell for a positive price under certain assumptions. Agents will buy these assets when they are young and then sell them when they turn old. Tirole considered the possibility of deterministic asset bubbles that survive indefinitely. Subsequent work, starting with Weil (1987), studied stochastic bubbles in which the asset price can switch from exceeding the fundamental value of the asset to equaling the fundamental value.

Since agents can always hold whatever other interest-bearing assets are available in the economy, the return on the bubble asset must be at least as high as the prevailing interest rate on other assets so that agents agree to hold it. Since agents must be able to keep buying the bubble asset for a bubble to be possible, the economy must grow at least as fast as this interest rate to ensure there are enough resources to keep buying the bubble asset.

For a bubble to be possible in Tirole's (1985) model, not only must the growth rate be at least as high as the interest rate, but the supply of the bubble asset must also be limited. When an asset is overvalued, agents will have an incentive to create more such assets and sell them, since they can earn more from selling an asset than they would have to pay out in dividends over the life of the asset. Some restriction on both the ability to create assets and the ability to sell them short is required to sustain a bubble in equilibrium. For example, Tirole (1985) assumed that the supply of bubble assets is constant over time and that short-selling is not allowed.

While assuming an infinite number of agents offers a way to get around the nonexistence of bubbles I discussed earlier, there are also circumstances in which a rational bubble can arise when the number of agents is finite. Recall the argument for ruling out bubbles I previously laid out assumed both that there is a finite number of agents and that all these agents value the asset equally. Consider a setting in which each period, some fraction of agents have a temporary urge to consume or a temporary opportunity to produce. Different agents are hit with these shocks at different times. To act on these shocks, agents might need to borrow against their future income to purchase the consumption goods they desire or the inputs they need to produce. If agents are limited in how much they can borrow, they might want to sell assets to fund such activities. At the same time, agents who anticipate being borrowing constrained in the future might value the asset more than their currently constrained counterparts because they can sell that asset in the future when they become constrained. Those agents who buy the asset will not do so for the dividends the asset offers but for its help in relaxing future borrowing constraints. In fact, they may be willing to pay a positive price for an asset that pays no dividends. Although some agent will necessarily hold the asset infinitely often when there are finitely many agents, doing so will not be suboptimal for them: That agent will keep buying the asset in anticipation of being constrained and then sell it later to an agent who is not constrained and as such values the asset more than them. There is no ultimate fool who holds the overvalued asset for infinitely many periods but would be better off if they did not do so.[5] Kocherlakota (1992) and Santos and Woodford (1997) analyzed these types of models.

The examples of rational bubbles here are all adapted from earlier work exploring the foundations for noncommodity money, that is, work on why intrinsically worthless paper might have value and trade for goods. Tirole (1985) built on the overlapping-generations model of money in Samuelson (1958), and the models of bubbles with borrowing constraints in Kocherlakota (1992) and Santos and Woodford (1997) built on the Bewley (1980) and Townsend (1980) models of money with infinitely lived agents. Not surprisingly, these rational bubble models involve bubble assets that serve as money-like objects. In the Tirole (1985) model, different cohorts buy an intrinsically worthless asset to use as a means of storage. This allows them to save some of the income they earn when they're young and exchange it for consumption goods when they turn old. In Kocherlakota (1992) and Santos and Woodford (1997), agents with temporarily low income would like to borrow while agents with temporarily high income would like to lend, and they end up trading an intrinsically worthless asset to replicate this desired allocation given limits on borrowing.

Some economists have seized on the connection between these rational bubble models and models of money to interpret digital currencies as examples of rational bubbles. If there is some goal that households aim for but cannot satisfy with existing assets—for example, saving at a high rate of return, relaxing borrowing constraints, or finding a convenient store of value—households might use intrinsically worthless digital coins that pay no dividends to achieve their desired goal. Since money-like objects can be associated with booms and busts, these models might offer insight on why the prices of digital currencies can be so volatile.

Limitations of rational bubble models

One well-known feature of the monetary models that the rational bubble models I have discussed thus far build on is that money can only have value if in each period, agents expect it to keep circulating in the next period with positive probability. If agents expected money to stop circulating by some fixed date, they would not accept money one period earlier—and the equilibrium in which money has value would unravel. This implies that while rational bubbles featuring money-like objects might crash, they must also be able to potentially survive indefinitely. This is at odds with Kindleberger's (1978) proposed definition of bubbles as unsustainable asset price increases that must necessarily implode.

The implication that rational bubbles must be potentially sustainable indefinitely is not necessarily problematic. Recall Greenwood et al. (2019) found that nearly half of the asset booms in their sample kept growing after the initial rapid surge in prices. However, since a bubble can in principle last indefinitely, the existence of bubbles crucially relies on there being infinitely many trading periods, traders, or states of the world.[6] Such bubbles are not possible in finite environments. The notion that the existence of bubbles hinges on an environment that is infinite along some dimension seems hard to defend. Do agents really distinguish between very long and infinite horizons? Is it plausible that whether a bubble is possible depends on whether the universe will end in finite time? The fact that rational bubbles hinge on such assumptions led Santos and Woodford (1997, p. 19) to conclude that such rational bubbles are "relatively fragile." In particular, Santos and Woodford (1997, p. 48) reported that their "main results show the nonexistence of asset pricing bubbles under fairly general assumptions," and "these results suggest that known examples of pricing bubbles depend upon rather special circumstances."

Giglio et al. (2016) proposed an empirical test for this class of rational bubbles that focuses on whether agents distinguish between very long and infinite horizons. They collected data from the United Kingdom and Singapore, countries in which a nonnegligible number of housing units are attached to extremely long lease contracts that extend hundreds of years into the future. Such leases are issued by several long-lived entities, including the British royal family, large corporations, colleges, and governments. Many of these lease contracts are associated with ordinary housing units that are indistinguishable from neighboring units that can be freely bought and sold. For very long lease contracts that expire only hundreds of years in the future, buying a leasehold should be a close substitute to buying a house free and clear: Whoever owns the lease contract on a unit can rent it out to others, sell the lease contract to others, make improvements to the property, and bequeath the lease contract to their immediate offspring. A long lease thus governs housing services well past the planning horizon of those who lease a given property, but the rights to these housing services that it stipulates do not last indefinitely.

Giglio et al. (2016) found that these leaseholds traded at the same price as similar houses that were traded freely, including when house prices boomed in both the UK and Singapore in the mid-2000s. Under the logic of the rational bubble models I described previously, this finding would mean that one could reject the prospect that a rational bubble in housing markets can potentially explain the boom in house prices in these two countries. Formally, for any reasonable discount rate, the present value of housing services beyond the maximum duration of the lease contract will be negligible. Hence, the present discounted value of housing services over the duration of the lease contract is approximately equal to the present discounted value of housing services for a freely trading house. The fundamental value of long leaseholds and freely trading houses should therefore be nearly identical. By the logic of rational bubble models, there cannot be a bubble for leaseholds given the lease contract expires after a finite amount of time. Hence, the price of long leaseholds must equal the fundamental value of leaseholds, which in turn is approximately the same as the fundamental value of freely trading houses. Since leaseholds and freely trading houses trade at the same price, it follows that there cannot be a bubble in housing. This suggests that according to rational bubble models, there could not have been a bubble in the housing markets in the UK and Singapore.[7]

Irrational beliefs as an alternative

The evidence against rational bubbles as an explanation for boom-and-bust episodes prompted researchers to explore settings in which not all agents are rational. For example, Barberis et al. (2018) cited the Giglio et al. (2016) results as a reason for skepticism about rational bubble models. They proposed an alternative model in which some agents naively expect that future returns on assets will be similar to returns on that same asset in the past. Specifically, they assumed that a fraction of traders in the market for assets form extrapolative beliefs. A high realization for dividends leads these agents to expect high payoffs in the future, inducing them to try to buy assets from agents with rational beliefs. This pushes up asset prices, confirming these agents' expectations for high returns. Eventually, if their optimism about assets is not matched by high dividends, optimistic traders will lack the income to keep buying more assets. At that point, asset prices stop rising. The same agents who were previously bullish about returns will turn bearish. Expecting the low returns to continue into the future, they start selling their assets to agents with rational beliefs. Thus, asset prices predictably rise and fall in the wake of a temporary shock to dividends.

The logic of the Barberis et al. (2018) model does not hinge on an infinite horizon. Agents buy assets because they expect a high payoff next period, regardless of what happens to the asset in the long run. As such, the model is consistent with the finding in Giglio et al. (2016) that traders do not seem to distinguish between infinitely lived assets and those that expire hundreds of years into the future. The Barberis et al. (2018) model is also consistent with Kindleberger's (1978) definition of a bubble as an unsustainable increase in asset prices that must eventually implode, since a positive but temporary shock to dividends can lead to a temporary boom that must eventually collapse. At the same time, the Barberis et al. (2018) model does not imply that every asset price boom must end in collapse. A shock that leads to permanently higher dividends will correspond to a boom without a bust, in which high dividend payments allow buyers to continue to spend more on the asset.

Bordalo et al. (2021) considered an alternative framework in which agents hold diagnostic beliefs rather than extrapolative beliefs. Diagnostic beliefs assign more weight to more salient events, which may include events that happened recently. Agents who hold diagnostic beliefs will eventually correct their beliefs as the information that led them to turn optimistic becomes stale. The higher weight given to recent patterns leads agents to behave similarly to the agents with extrapolative beliefs in Barberis et al. (2018). Favorable news can thus lead to a boom. However, the reason a boom gives way to a bust is different than in Barberis et al. (2018). Rather than the price falling when optimistic agents can no longer afford to keep buying assets, the price falls when agents update their views as dividends fail to live up to their expectations. Agents eventually abandon their optimism as the facts fail to match their expectations. Although Bordalo et al. (2021) assumed all agents hold diagnostic beliefs, they also assumed that different agents receive different signals. Agents with more favorable signals buy the asset from those with less favorable signals. The result is a boom-and-bust dynamic in the wake of a temporary positive shock to fundamentals where some agents temporarily bid up the price of the asset.

Other work has explored different psychological biases. For example, Scheinkman and Xiong (2003) looked at the asset pricing implications of overconfidence. In their model, all agents receive equally informative private signals about an asset's fundamentals. These signals are assumed to be independent across agents. Each agent incorrectly believes that their particular signal is more precise than the signals of others. At the same time, they understand that others will focus on their own signals. One implication of this model is that the agent who receives the most favorable signal about an asset's fundamentals will bid up its price toward their beliefs. The price will therefore exceed the fundamental value of the asset given equally weighted signals, in line with the objectively true precision. In addition, the agent with the highest current assessment of the fundamentals would be willing to pay more than what they believe the asset will pay out in dividends. This is because they understand that in the future some other agent might be more optimistic about the asset than they will be, and in that event they can sell the asset to that agent for a higher price than they believe can be justified by its dividends. That is, there is a speculative motive for buying an asset in the hope of selling it to a more optimistic agent later in the future.

Overconfidence can thus push up the asset price above even what the most optimistic agent believes the asset will pay out in dividends.

While overconfidence can explain why an asset might be overvalued, it does not necessarily explain why the price of an overvalued asset should crash. Agents who buy the asset for speculative purposes count on selling the asset to others who rely on their own positive signals and are, hence, bullish about the asset. More recent work has shown how a combination of overconfidence and leverage can lead to crashes. Examples of such arguments include Fostel and Geanakoplos (2008), Geanakoplos (2010), and Simsek (2013). If the agents who are most optimistic about an asset borrow funds to buy more of it, they may be forced to sell the asset if a negative shock leaves them unable to pay their debt obligations. The agents who end up with the asset after a negative shock will be those with less favorable beliefs, and they will require a lower price to hold the asset. A shock that negatively impacts optimists will trigger a crash if those optimists had borrowed against the assets they hold.

Although leverage can help explain why asset booms might be followed by busts, not all historical boom-and-bust episodes were associated with high leverage. For example, the U.S. dot-com boom in the late 1990s did not feature significant borrowing against technology stocks, yet the price of equity shares in technology companies eventually crashed. Hong et al. (2006) offered a different explanation for crashes when agents are overconfident that does not rely on leverage. They observed that the decline in equity prices at the end of the dot-com boom occurred at the same time that insiders working at technology firms were allowed to begin selling their equity holdings. Hong et al. (2006) considered a version of a model with overconfidence in which traders are risk averse. In that case, the most optimistic trader will be reluctant to hold all the assets out of concern that they will be exposed to too much risk from a particular asset class. In their model, when some agents receive favorable signals about an equity stake in a firm, they will buy more (but not all) of the asset from less optimistic agents and drive up its price, creating a boom. This higher price will encourage insiders at that firm, who are assumed to hold correct beliefs about the asset's fundamental value, to sell some of their asset holdings when they are not constrained. For the market to absorb this increase, the share of equity stakes held by less optimistic agents will have to rise given the reluctance of the most optimistic agents to take on more risk. Since these less optimistic agents require a higher expected return to agree to hold the asset, its price will fall.

Limitations of models with irrational beliefs

One issue for models of bubbles in which some agents hold irrational beliefs is that such bubbles might not survive if agents with rational beliefs were allowed to also trade assets. In principle, agents with rational beliefs should have an incentive to buy undervalued assets and to sell or short overvalued assets. That would tend to push asset prices toward fundamentals. Separately, since agents who hold irrational beliefs tend to make worse investment decisions than agents who hold rational beliefs, one might expect the former agents to have relatively little wealth to invest. This means irrational agents should have limited ability to drive asset prices away from fundamentals.

One reason that rational agents might not be able to fully eliminate overvaluation (or undervaluation) is that trading against irrational agents can be costly. Gromb and Vayanos (2010) surveyed several potential explanations for why agents may find it too costly to fully offset deviations in asset prices from fundamentals. One reason is constraints on short sales, which recall are necessary to sustain bubbles even when all agents are rational. But there are other explanations for why agents may not act to push prices back to fundamentals. Taking large arbitrage positions may require agents to borrow large amounts, which might be costly or impossible to do. Risk-averse agents may also be reluctant to take large positions if the returns on the overvalued assets are stochastic. This particular feature plays an important role in the Barberis et al. (2018) model, which allowed rational agents to trade with the agents holding extrapolative beliefs. In their model, rational agents actually want to take similar positions as those of agents with extrapolative beliefs early on, since the former set of agents expects the latter to drive up asset prices in the near future, implying a high expected capital gain from buying the asset. Thus, constraints on rational agents are not necessary early on during the boom. DeLong et al. (1990) first derived conditions under which rational agents amplify rather than offset the trades of irrational agents. Abreu and Brunnermeier (2003) found a similar result in a different setting. However, in the Barberis et al. (2018) model, rational agents will eventually want to exit the market and short the asset once they become worried about price declines. Limits to arbitrage then become necessary to sustain the bubble in equilibrium.

The separate argument that agents holding irrational beliefs are unlikely to have enough wealth to drive asset prices away from fundamentals, at least not for very long, was originally advanced by Friedman (1953). Sandroni (2000) and Blume and Easley (2006) subsequently showed that under fairly general conditions, long-run wealth will indeed be concentrated among the most patient agents whose beliefs are closest to the truth, consistent with Friedman's intuition. However, later work has found alternative scenarios in which this result does not hold. For example, Kogan et al. (2006) showed that the share of wealth held by agents with irrational beliefs may not tend to zero in a growing economy, and Borovička (2020) showed that the result may not hold when agents have time-nonseparable preferences, meaning that how much agents enjoy consumption today depends on how much they consumed in the past or expect to consume in the future. Kogan et al. (2006) also showed that even when the wealth share of certain agents does tend to zero, the impact of this group on asset pricing may not vanish.

More generally, the claim that the share of wealth held by irrational agents will vanish in the long run requires that the agents who hold rational beliefs remain consistently rational over time. If rational agents who accumulate wealth periodically switch to holding irrational beliefs, such as when they face new scenarios they have not previously encountered, or if their heirs inherit their wealth but not their rational beliefs, the wealth held by agents holding irrational beliefs at any given point in time may be enough to keep asset prices away from fundamentals.

In short, while models where some traders hold irrational beliefs can give rise to asset bubbles, these models require that rational agents find it too costly to trade in a way that offsets deviations from fundamentals. These models also require some mechanism to ensure that enough wealth flows to agents holding irrational beliefs so that they can end up pushing asset prices significantly above fundamentals. Irrationality on its own may not suffice to give rise to an asset bubble when some market participants are rational.

Rationality or informational symmetry?

Up to this point, I have shown that rational bubble models adapted from monetary models do allow asset bubbles to arise, but those bubbles are fragile and are at odds with data on asset prices for long-lived assets. By contrast, models in which agents assign excessive weight to information that is recent or salient or which they themselves produced can generate plausible and robust bubbles, provided their trades are not offset by agents with rational beliefs.

I now argue that irrational beliefs are not essential for generating robust and empirically plausible asset bubbles. The rational bubble models I have discussed so far assumed not only that all agents are rational but also that all agents hold identical beliefs. In particular, an important (and often unstated) assumption in these models is that any information agents receive is common knowledge across agents. This means agents are fully coordinated in their beliefs: They all observe the same information, they know that others observe the same information, they know that others also know that they all observe the same information, and so on.

When information is common knowledge, agents have no reason to believe that others around them will value the asset any differently as a result of the information those agents receive. As such, there is no scope for a greater-fool bubble in which agents buy an asset expecting to profit at the expense of lesser informed agents. Instead, a bubble can emerge only if trading the asset creates mutually beneficial gains for all the agents and there is no ultimate fool who ends up stuck with the asset against their own self-interest. For example, in overlapping-generations models, a bubble can only occur if the economy grows at least as fast as the interest rate that would prevail if there were no bubble. When this condition holds, agents would be at least as well off by committing to a scheme in which they all transfer a share of their earnings to the previous generation and in return receive a share of the earnings of the next generation. Since income grows at least as fast as the interest rate, this scheme would give agents a return that is at least as high as they could earn from investing. Trading an intrinsically worthless asset implements these transfers and makes all cohorts better off. Similarly, in models of bubbles based on borrowing constraints, trade in overvalued assets allows agents to relax their borrowing constraints and direct more resources from those who are not currently borrowing constrained to those who are.

By contrast, when information is not commonly shared across agents, rational agents might agree to buy an overvalued asset because they believe they can profit at the expense of lesser informed agents—and not because they believe that trading the asset benefits all agents. I now argue that rational bubbles when agents hold different beliefs are more similar to bubbles that arise when agents hold irrational beliefs and buy overvalued assets because they expect these assets will yield high short-run returns.

Rational bubbles without common knowledge

The first example of rational bubbles where agents do not share the same beliefs involves settings in which the information agents receive about assets is not common knowledge. The first paper to show that this scenario could allow for bubbles was Allen et al. (1993). That paper presented a finite-horizon model in which there is one state of the world where all agents know that an asset is worthless and there is another state of the world where only some of the agents know the asset is worthless. Agents who know that the asset is worthless are not sure which state they are in, and hence, are unsure what the other agents know. Allen et al. (1993) showed that agents might be willing to pay a positive price to buy an asset they know is worthless to try to sell it to another agent for a higher price. Conlon (2004), Doblas-Madrid (2012), and Awaya et al. (2022) generalized this idea and developed more robust examples of such bubbles.

To be sustainable, bubbles of this type rely on several key assumptions. First, there must be at least two rounds of trading: an early period in which agents know that an asset is overvalued and a later period in which whoever held the overvalued asset in the early period can potentially sell it for a higher price. Second, as Liu et al. (2023) explained, there must be at least three particular states of the world that each serve particular functions. There must be one state of the world in which all agents know that the initial price of the asset exceeds its fundamental value. This is the state of the world in which the asset is a bubble. There must also be another state of the world in which there is some agent who does not know that the asset is overvalued and is willing to buy it for a higher price than the one that prevailed in the initial period. Importantly, the agent who knows the asset is overvalued in the initial period cannot tell these two states apart at that time. This is what makes that agent willing to gamble on trying to sell it later for a higher price. Finally, there must be a third state of the world in which whoever buys the asset in the later period is better off for having bought the asset. Moreover, that buyer cannot distinguish in the later period between this state and the second state in which they buy an asset that is overvalued. Unless the buyer can profit from buying the asset in some state, they would be reluctant to buy the overvalued asset when another trader tries to sell it to them.[8]

As an example of such a bubble, consider the Doblas-Madrid (2012) model that is inspired by the earlier work of Abreu and Brunnermeier (2003). Agents in the Doblas-Madrid model can trade what are essentially equity shares in new technology firms whose ultimate profitability is uncertain. In the Doblas-Madrid model, the price of these equity shares rises initially as firms refine the new technology and render it more productive. Eventually, the maximum technological potential from the technology is realized and the fundamental value of the asset stops growing. Doblas-Madrid (2012), following Abreu and Brunnermeier (2003), assumed the news that the asset's fundamental value stopped growing is communicated to agents slowly and incompletely. As a result, when agents learn that profits have already peaked, they will not know exactly when the technology peaked and whether other agents are already aware of this fact. Agents might be willing to hold on to equity shares at a price they know exceeds the fundamental value of the asset so that they can try to sell those equity shares later on at a higher price to others who learned that the asset is overvalued after they did.

For a bubble to be an equilibrium, the agents who hold an asset that they know is overvalued must believe there is some chance they will be able to sell it for a higher price in the future. This in turn requires that there are willing buyers for the asset despite the risk that more knowledgeable traders are trying to take advantage of buyers by selling them an overvalued asset. Doblas-Madrid (2012) assumed that some agents periodically have a desperate need for cash that makes them willing to sell their asset at any price, regardless of what they know. Agents who buy equity shares thus cannot be sure whether they are buying from someone with a pressing need for liquidity who might be selling the asset at a low price or an informed agent trying to take advantage of them by selling it at too high of a price.

While the Doblas-Madrid (2012) model is elaborate and features many states of the world, it contains the three essential states of the world needed to sustain a bubble. First, if agents hold on to the asset for a long enough period after learning that the asset is overvalued, there will be a state in which all agents would have learned that the asset is overvalued and its price remains above fundamentals. However, this is not common knowledge: The last agent to learn that the fundamentals stopped growing does not know whether they are the first or the last to learn, so this agent cannot be sure whether others have learned that the fundamentals stopped growing. Second, there is a state in which the agent who holds an overvalued asset can profit from selling it, namely, if they learn that the fundamentals stopped growing before others did and manage to sell the asset before their decision to sell lowers the asset price and reveals to all agents that the fundamentals stopped growing. Finally, there is a state of the world in which those who buy the asset profit from doing so, namely, if they buy the asset before the price collapses from someone who is desperate for liquidity and willing to sell it at any price.

Unlike rational bubble models with symmetric information, rational bubble models with private information allow bubbles to arise in completely finite settings. To see why these bubbles will not unravel as in a symmetric information setting, observe that in the final period of trading, no agent would be willing to buy the asset in the state of the world where all agents know that the asset is overvalued. But there is another state of the world in which some agents do not know that the asset is overvalued in the last period and would be willing to buy it at a price that exceeds the objectively correct fundamentals. One period earlier, agents might be willing to buy the asset at a price they know is overvalued because they are unsure which of these two states they are in and they choose to gamble that they can sell the asset at an even higher price in the last period to an agent who is unaware that the asset is overvalued. The impossibility of overvaluation in the last period in one state of the world does not preclude the possibility of overvaluation in that same state of the world earlier.

In the state of the world where all agents know that the asset is overpriced, the asset price will have to eventually crash. This is because in the last trading period, no agent is willing to buy an overvalued asset. At this point, the asset price will collapse if it hasn't already done so earlier. In that sense, the price of an overvalued asset must ultimately implode, consistent with Kindleberger's (1978) definition for a bubble. At the same time, not every boom is a bubble, so not every boom ends in a bust. In certain states of the world, some agents do not know that the asset is overvalued and the price does not crash. This is similar to what happens in the models where agents hold extrapolative or diagnostic beliefs. In those models, the occurrence of a crash depends on whether the shock to dividends that starts the boom is permanent or temporary. With a permanent shock, the asset price rises and remains high, and the agents' beliefs are correct. With a temporary shock, agents overvalue the asset, and the price of the asset will eventually collapse in finite time. The pattern in both of these models offers a stark contrast to rational bubble models where all agents have the same information. In rational bubble models, a bubble must potentially last indefinitely to exist, and it is impossible for a boom that is specifically associated with a bubble to necessarily collapse in finite time.

Rational bubbles with risk shifting

In the Doblas-Madrid model of bubbles I just discussed, the existence of a bubble is itself not common knowledge: While all agents know that the asset is overvalued, they are not sure if others also know it is overvalued. This contrasts with symmetric information models of bubbles, where it is common knowledge that an asset is overvalued. However, there are private information models of rational bubbles where all agents know that an asset is overvalued. The key feature of those models is that agents have asymmetric information not on the payoff from the asset but on how exposed individual agents are to the asset. More precisely, these models feature wealthy agents who rely on intermediaries to carry out investments on their behalf, but they cannot monitor or control what these intermediaries do with the funds they receive. Wealthy agents hope that intermediaries invest in safe assets. However, if intermediaries can shift losses to their investors, they may prefer to invest at least some of the funds they receive in risky assets in the hope that these assets pay more than they promise to pay their investors. If demand for risky assets from intermediaries is sufficiently high, the price of such assets can rise above fundamentals. All agents would then know there are assets that are overvalued, but those whose wealth is used to buy those assets are unsure what assets the intermediaries they invested with bought.

The first model of risk-shifting bubbles was developed by Allen and Gorton (1993). Their model required elaborate assumptions to ensure that information about what investors do is not revealed to others. Subsequent work by Allen and Gale (2000) presented a simpler model that conveyed the same mechanism. Barlevy (2014), Dow and Han (2015), and Allen et al. (2022) further explored these types of models.

In addition to the assumption that wealthy agents must rely on others to invest on their behalf, the existence of bubbles in these models requires limits on contracting that prevent wealthy agents from incentivizing managers to act in the interest of the investors whose wealth they handle. For example, in Allen and Gale (2000), wealthy agents must lend to others using fixed-rate credit contracts with limited liability for the borrower. This precludes lenders from using a menu of contracts to screen out borrowers prone to buying risky assets or from designing contracts to punish borrowers who did buy risky assets. Absent these safeguards, borrowers will have an incentive to buy risky assets and gamble that they pay off more than what they promised lenders. This incentive may be strong enough that borrowers would be willing to buy risky assets even when the asset price exceeds the expected payoff from the asset.

More generally, bubbles can arise even if there is a single risky asset rather than a mix of risky and safe assets. For example, Barlevy and Fisher (2021) developed a model of bubbles in which the only asset that agents can hold is housing, but borrowers have preferences for housing services that are unobservable to lenders. Some borrowers greatly value homeownership and would not default on their loans if house prices fell. Others do not particularly value homeownership and borrow for speculative reasons, hoping to profit if house prices grow and intending to default if they fall. In that case, demand for housing from speculators might drive up its price above the expected value of housing services when these services are allocated to those who value homeownership most. Lenders know that the houses they are lending against are overvalued, but they do not know how exposed they are to default risk for any given borrower given they do not know their borrower's preferences.

Risk-shifting models of bubbles can arise in completely finite settings, so these models are not subject to the same critiques as rational models of bubbles with symmetric information among all agents. In fact, the Allen and Gale (2000) model can give rise to a bubble when there is only one trading period. Agents can profit from buying a risky asset even without selling it by gambling on whether the realized dividends on the asset are high or low. Risk-shifting models are also robust in the sense that small changes in the environment will not eliminate overvaluation. However, these models do rely crucially on contracting frictions that prevent lenders from screening out or discouraging speculation among their borrowers.

One difference between risk-shifting models of bubbles and models of bubbles in which information is not common knowledge is that the collapse of a bubble is not inevitable in risk-shifting models as it is in models without common knowledge. In risk-shifting models, agents buy overvalued risky assets to gamble. The gamble succeeds when the asset is valuable, for example, if the realized dividend on the asset is high. In that case, the asset's price will rise. The gamble fails when the realized dividend on the asset is low. In that case, the asset's price will fall. Therefore, the fact that the asset is overvalued does not imply that its price must eventually collapse. By contrast, in models of bubbles where information is not common knowledge, if the asset is a bubble and all agents know it is overvalued, agents will try to sell the asset to others who know that the asset is overvalued, at which point the asset price must collapse.

Subjective beliefs

While rational agents might hold different beliefs from one another if they are exposed to different information, another reason their beliefs might differ is that they might start with different subjective beliefs when confronted with situations they never encountered before. In this case, a rational agent might agree to buy an asset that they think is overvalued to sell it later to someone they know will be more optimistic about it in the future. While this bears some similarity to models of bubbles where information is not common knowledge, the motivation for agents to trade when agents hold heterogeneous subjective beliefs is different. In particular, when agents start with the same priors but receive different private information, they might be willing to buy an overvalued asset to gamble that their information is better than the information of others around them. If their information proves worse, they will incur a loss. Since agents are aware of this risk, they will only buy an overvalued asset if they expect its price to rise should their gamble succeed. The profits they earn if the price rises compensate them for the losses they would incur should the price fall. When agents hold different subjective beliefs, they will not necessarily perceive their speculation as risky in the same way given they treat the subjective beliefs of others as incorrect. As such, they may not require the price of the asset to rise in order to be willing to hold it. Overvaluation will not necessarily imply rising asset prices as is true in models with private information.

Models in which agents hold heterogeneous subjective beliefs are similar to the Scheinkman and Xiong (2003) model of overconfidence that I discussed earlier. Recall in that model, agents believe their signals are more precise than the signals others receive, even though all signals are in fact equally precise. In that sense, one could say that agents hold irrational beliefs and are prone to be overly swayed by the particular information they receive. However, if agents begin with different subjective priors about situations they are unfamiliar with, it need not follow that these agents must be irrational. Rationality requires that agents update their beliefs in a way that is consistent with Bayes' rule, and it arguably requires that agents not put zero probability on the objectively true state of the world. But rationality does not impose that agents must start out with the same subjective beliefs, especially when facing situations for which there is no objective historical basis to guide beliefs.

Work by Harrison and Kreps (1978) showed that when agents have different subjective beliefs and the identity of the agent who values the asset most can change over time, the equilibrium price of the asset can exceed what each of the agents believes the asset could pay out as dividends. This is the same result as in Scheinkman and Xiong (2003), precisely because the two models have a similar structure. The key difference between the two models is the reason agents hold different beliefs. Scheinkman and Xiong assumed that some agents are overconfident, while Harrison and Kreps assumed that agents start out with different subjective beliefs. Results for models where agents are overconfident can therefore arise in settings where agents are rational. As in models in which agents are overconfident, generating crashes when agents start with different subjective beliefs may require additional features such as borrowing or risk aversion.

One critique of the Harrison and Kreps (1978) framework is that it assumed agents maintain fixed subjective beliefs over time. If dividends were determined by some fixed stochastic process, agents should in principle update their beliefs and learn the true distribution over time as dividends are realized. Since Harrison and Kreps assumed agents hold fixed beliefs over time, these agents appear to be dogmatically and irrationally sticking to their prior beliefs, much in the same way that agents in the Scheinkman and Xiong (2003) model dogmatically and irrationally believe that their signals are superior to those of others. Morris (1996) considered a version of the Harrison and Kreps model in which agents do update their beliefs. He showed that at least under certain conditions, the equilibrium price will exceed what all agents expect the asset will pay out in dividends early on, even though all agents' beliefs eventually converge to the true value. The price of an asset can exceed what each agent believes it will pay out even when agents learn and update their beliefs in light of the data they observe.

Rationality and irrationality

Models in which rational agents do not all share the same beliefs give rise to qualitatively different bubbles from the bubbles that can arise when rational agents share the same beliefs. If agents hold different beliefs, there can be overvaluation in finite settings, and agents might be willing to buy assets they believe are overvalued in the hope of profiting at the expense of others, just as in models where some agents hold irrational beliefs. This suggests that irrational beliefs are not essential for generating robust and empirically plausible asset bubbles. Even rational agents who understand that the assets they buy are overvalued might be willing to buy them so that they can profit from other agents who hold different beliefs, much as agents who are overly optimistic or overly confident in their own information are willing to buy assets that they expect will deliver high returns or can be sold later to more optimistic agents. That raises an important question: If bubbles in models where agents hold irrational beliefs are similar to bubbles in models where agents are rational but hold different opinions, does it matter whether agents are rational?

One reason why it might matter whether agents are rational or not involves policy. When agents hold irrational beliefs that are at odds with the underlying objective truth, there is scope for policymakers to intervene to correct the actions of agents that policymakers might view as misguided. For example, while agents who are optimistic about an asset would be willing to pay a high price for it, a policymaker might view such a trade as going against that agent's own self-interest. Such an intervention is inherently paternalistic, since it involves imposing the preferences and beliefs of a policymaker on agents whose preferences and beliefs differ from those of the policymaker. By contrast, with rational agents, there may be reasons for policymakers to intervene, but these rely on correcting an inefficiency that the agents should agree is harmful. That is, a policymaker should be able to get rational agents to agree that the proposed interventions would leave them all better off. For example, rational agents who trade in an overvalued asset might not take into account that their trade encourages excessive asset creation in states of the world where the asset is overvalued and discourages asset creation in other states of the world where the asset is undervalued. Since the policymaker's beliefs and preferences are the same as those of the agents, the intervention merely aligns incentives so that agents internalize all of the effects of their actions. Conlon (2015) and Dow and Han (2015) described private-information models of rational bubbles in which there is scope for intervention to avoid misallocation when assets are created. This is a different reason for intervention than when policymakers intervene paternalistically to correct the erroneous beliefs that agents hold.

Whether agents are rational also matters for the type of policy interventions that policymakers might undertake. In models where agents form beliefs that are influenced by prices, such as models with extrapolative beliefs, policymakers can potentially influence outcomes by changing asset prices and affecting agents' beliefs. Farhi and Werning (2020) studied interventions of this type in a setting where goods prices are rigid. In such models, aggregate demand for goods depends on wealth, which in turn depends on asset prices. If asset prices fall, possibly in response to a shock to the realized dividends of the asset, aggregate demand will fall as agents feel poorer. When goods prices are rigid, this fall in aggregate demand will result in an inefficient contraction in output. This can in principle be addressed using monetary policy to stimulate aggregate demand. But if there are limits on monetary policy, such as a zero lower bound on nominal rates, there may be scope for making society better off by mitigating the drop in asset prices in response to a negative shock. If agents form beliefs about future returns by extrapolating from past returns, one way to mitigate the drop is by making agents less pessimistic after such a shock. Reducing the price of the asset during the boom would mean that the same low asset price after the shock will be associated with a less negative expected return. Agents whose expectations of future returns are extrapolated from past returns would then be more optimistic about assets and thus more willing to buy them. Policy can improve outcomes by shaping the beliefs agents hold.

When rational agents hold different subjective beliefs that are not influenced by historical returns, there may still be a role for policymakers to intervene if goods prices are rigid and monetary policy is unable to respond after a negative shock to realized dividends. This scenario is analyzed by Caballero and Simsek (2020). They showed that preventing optimists from borrowing to buy assets during the boom may improve welfare. However, the purpose of this intervention is not to lower the price during the boom and shape agents' beliefs. Rather, the purpose is to prevent optimists from becoming overly leveraged in a way that would force them to liquidate their assets after a negative shock so that less optimistic agents end up with the assets. Restricting credit this way may reduce the price of assets before the negative shock, but the primary goal of the intervention is to stimulate demand for the asset after the shock, not to dampen it before the shock and shape beliefs.

Finally, the question of whether agents are rational or not matters for how to interpret the evidence on the existence of bubbles. Recall how Fama (2014) described a bubble as "an irrational strong price increase that implies a predictable strong decline." Fama went on to argue that the average return on assets whose price previously surged is in practice no different than the average return on assets whose price did not surge. This is consistent with what Greenwood et al. (2019) found: Roughly half of asset booms ended in a crash within two years, while the other half continued to rise, so the average returns on such assets were no different than for the market as a whole. While this evidence suggests that price increases do not predict subsequent low average returns, it does not mean that overvaluation isn't possible. Indeed, rational bubble models imply declines should not be predictable. Even if a bubble—a state of the world in which all agents know that the asset is overvalued—must end in a collapse, not all asset booms do. The only reason that rational agents would be willing to hold an asset they know is overvalued and whose price might crash is precisely because they expect to profit from selling it in some states of the world. Evidence that asset booms do not predict lower returns may rule out the possibility of bubbles driven by particular irrational beliefs, but not bubbles where agents are rational.

Conclusion

In this article, I have argued that some of the main issues with early work on rational bubbles are a consequence of the assumption in those models that agents hold the same beliefs rather than agents' rationality *per se*. Models where agents are both rational and hold different beliefs can give rise to robust and empirically plausible bubbles. It is not the case that robust and empirically plausible bubbles can only arise when agents get swept up in unwarranted optimism about an asset. Empirically plausible bubbles can arise even when all agents have a sober assessment of the market and understand that the asset they are trading is overvalued. This is not to argue that participants in financial markets are always rational and that market psychology has no role in explaining asset booms in practice. Rather, the point is that one cannot dismiss explanations of asset bubbles that arise when rational agents trade assets they understand are overvalued but believe they can trade profitably. Therefore, the exact role of market psychology remains an open question—one that can matter for whether and what type of policy interventions might be appropriate during asset booms.

While this article has emphasized how models of rational bubbles in which agents hold heterogeneous beliefs can overcome some of the issues with rational bubble models in which agents share the same beliefs, I do not wish to dismiss rational bubble models with symmetric information. Such models may be relevant in some situations. Consider the example of cryptocurrencies. Many (but not all) cryptocurrencies offer no dividend payments. It seems reasonable to think that agents know that the underlying asset is intrinsically worthless, yet these digital currencies trade at positive prices. Since rational bubble models with symmetric information among agents build on models of money, such models are natural starting points for studying an asset billing itself as a purely digital substitute for traditional forms of money. At the same time, agents might disagree about other aspects of cryptocurrencies, such as how widely they might circulate or how valuable these assets might be for those with limited savings alternatives. Rational models of bubbles in which agents disagree and traders buy assets so that they can sell them to someone else for a profit might therefore be relevant for studying some aspects of cryptocurrencies. But models of rational bubbles where bubbles are money-like objects may also be useful frameworks for studying this phenomenon.

Notes

[1] This article is based on a chapter I have prepared for the *Handbook of Bubbles and Manias*, which is edited by Manuel Santos.

[2] The observation that similar bubbles arise in models with rational agents holding heterogeneous beliefs and models with irrational agents is similar in spirit to the point in Brav and Heaton (2002) that certain anomalous patterns in asset prices can arise in both behavioral models where agents hold irrational beliefs and rational models where agents are uncertain. However, they focused on underreaction and overreaction rather than bubbles and on learning rather than heterogeneous beliefs.

[3] Bayes' rule is a statistical formula governing how to update the likelihood of a given event in the face of new information that is relevant about the event.

[4] These two statements are not in logical conflict because one concerns the marginal benefit of an additional asset while the other concerns the average benefit from existing assets. The average benefit from existing assets may exceed the cost of producing them even if the marginal benefit from creating an additional asset exceeds the marginal cost of creating it.

[5] Formally, the transversality condition is not a necessary condition for optimality when agents face binding borrowing constraints.

[6] Most rational bubble models feature infinitely many trading periods. Kamihigashi (1998) showed that a rational bubble can also arise with only two periods but infinitely many states of the world. The reason the bubble doesn't unravel despite the finite trading horizon is that expected utility falls to negative infinity if an agent sells any of the asset. This result relies on utility being unbounded below, allowing for states with arbitrarily negative utility. Kamihigashi noted that his example relies on nonstandard preferences and that the point of his example is not to explain real-world booms but to show that the infinity required to sustain a bubble need not necessarily involve conditions on the trading horizon.

[7] A different argument against rational bubble models is that some of these models hinge on dynamic inefficiency, or the idea that there is excessive capital accumulation and society would be better off consuming more instead of creating more capital. Abel et al. (1989) devised an empirical test for dynamic inefficiency and found little support for it. However, later work raised theoretical and empirical issues with their test. Moreover, as I pointed out before, rational bubbles can occur in environments with other frictions, such as borrowing constraints—which would hold true even in dynamically efficient economies. See Barlevy (2025) for a discussion of these issues.

[8] Tirole (1982) discussed the nonexistence of bubbles when there is no motive for agents to trade beyond the private information they receive.

References

Abel, Andrew B., N. Gregory Mankiw, Lawrence H. Summers, and Richard J. Zeckhauser, 1989, "Assessing dynamic efficiency: Theory and evidence," *Review of Economic Studies*, Vol. 56, No. 1, January, pp. 1–19. Crossref

Abreu, Dilip, and Markus K. Brunnermeier, 2003, "Bubbles and crashes," *Econometrica*, Vol. 71, No. 1, January, pp. 173–204. Crossref

Allen, Franklin, Gadi Barlevy, and Douglas Gale, 2022, "Asset price booms and macroeconomic policy: A risk-shifting approach," *American Economic Journal: Macroeconomics*, Vol. 14, No. 2, April, pp. 243–280. Crossref

Allen, Franklin, and Douglas Gale, 2000, "Bubbles and crises," *Economic Journal*, Vol. 110, No. 460, January, pp. 236–255. Crossref

Allen, Franklin, and Gary Gorton, 1993, "Churning bubbles," *Review of Economic Studies*, Vol. 60, No. 4, October, pp. 813–836. Crossref

Allen, Franklin, Stephen Morris, and Andrew Postlewaite, 1993, "Finite bubbles with short sale constraints and asymmetric information," *Journal of Economic Theory*, Vol. 61, No. 2, December, pp. 206–229. Crossref

Awaya, Yu, Kohei Iwasaki, and Makoto Watanabe, 2022, "Rational bubbles and middlemen," *Theoretical Economics*, Vol. 17, No. 4, November, pp. 1559–1587. Crossref

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2018, "Extrapolation and bubbles," *Journal of Financial Economics*, Vol. 129, No. 2, August, pp. 203–227. Crossref

Barlevy, Gadi, 2025, *Asset Bubbles and Macroeconomic Policy*, Cambridge, MA: MIT Press.

Barlevy, Gadi, 2014, "A leverage-based model of speculative bubbles," *Journal of Economic Theory*, Vol. 153, September, pp. 459–505. Crossref

Barlevy, Gadi, and Jonas D. M. Fisher, 2021, "Why were interest-only mortgages so popular during the U.S. housing boom?," *Review of Economic Dynamics*, Vol. 41, July, pp. 205–224. Crossref

Bewley, Truman, 1980, "The optimum quantity of money," in *Models of Monetary Economies*, John H. Kareken and Neil Wallace (eds.), Minneapolis: Federal Reserve Bank of Minneapolis, pp. 169–210, available online.

Blanchard, Olivier J., and Mark W. Watson, 1982, "Bubbles, rational expectations, and financial markets," in *Crises in the Economic and Financial Structure*, Paul Wachtel (ed.), Lexington, MA: Lexington Books, pp. 295–316.

Blume, Lawrence, and David Easley, 2006, "If you're so smart, why aren't you rich? Belief selection in complete and incomplete markets," *Econometrica*, Vol. 74, No. 4, July, pp. 929–966. Crossref

Bordalo, Pedro, Nicola Gennaioli, Spencer Yongwook Kwon, and Andrei Shleifer, 2021, "Diagnostic bubbles," *Journal of Financial Economics*, Vol. 141, No. 3, September, pp. 1060–1077. Crossref

Borovička, Jaroslav, 2020, "Survival and long-run dynamics with heterogeneous beliefs under recursive preferences," *Journal of Political Economy*, Vol. 128, No. 1, January, pp. 206–251. Crossref

Brav, Alon, and J. B. Heaton, 2002, "Competing theories of financial anomalies," *Review of Financial Studies*, Vol. 15, No. 2, January, pp. 575–606. Crossref

Caballero, Ricardo J., and Alp Simsek, 2020, "Prudential monetary policy," National Bureau of Economic Research, working paper, No. 25977, revised March 2020 (originally issued June 2019). Crossref

Conlon, John R., 2015, "Should central banks burst bubbles? Some microeconomic issues," *Economic Journal*, Vol. 125, No. 582, February, pp. 141–161. Crossref

Conlon, John R., 2004, "Simple finite horizon bubbles robust to higher order knowledge," *Econometrica*, Vol. 72, No. 3, May, pp. 927–936. Crossref

DeLong, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, "Positive feedback investment strategies and destabilizing rational speculation," *Journal of Finance*, Vol. 45, No. 2, June, pp. 379–395. Crossref

Diamond, Peter A., 1965, "National debt in a neoclassical growth model," *American Economic Review*, Vol. 55, No. 5, part 1, December, pp. 1126–1150, available online.

Doblas-Madrid, Antonio, 2012, "A robust model of bubbles with multidimensional uncertainty," *Econometrica*, Vol. 80, No. 5, September, pp. 1845–1893. Crossref

Dow, James, and Jungsuk Han, 2015, "Contractual incompleteness, limited liability and asset price bubbles," *Journal of Financial Economics*, Vol. 116, No. 2, May, pp. 383–409. Crossref

Fama, Eugene F., 2014, "Two pillars of asset pricing," *American Economic Review*, Vol. 104, No. 6, June, pp. 1467–1485. Crossref

Farhi, Emmanuel, and Ivan Werning, 2020, "Taming a Minsky cycle," Massachusetts Institute of Technology, unpublished manuscript.

Fostel, Ana, and John Geanakoplos, 2008, "Leverage cycles and the anxious economy," *American Economic Review*, Vol. 98, No. 4, September, pp. 1211–1244. Crossref

Friedman, Milton, 1953, "The case for flexible exchange rates," in *Essays in Positive Economics*, Chicago: University of Chicago Press, pp. 157–203.

Frehen, Rik G. P., William N. Goetzmann, and K. Geert Rouwenhorst, 2013, "New evidence on the first financial bubble," *Journal of Financial Economics*, Vol. 108, No. 3, June, pp. 585–607. Crossref

Geanakoplos, John, 2010, "The leverage cycle," in *NBER Macroeconomics Annual 2009*, Daron Acemoglu, Kenneth Rogoff, and Michael Woodford (eds.), Vol. 24, Chicago: University of Chicago Press, pp. 1–66. Crossref

Giglio, Stefano, Matteo Maggiori, and Johannes Stroebel, 2016, "No-bubble condition: Model-free tests in housing markets," *Econometrica*, Vol. 84, No. 3, May, pp. 1047–1091. Crossref

Greenwood, Robin, Andrei Shleifer, and Yang You, 2019, "Bubbles for Fama," *Journal of Financial Economics*, Vol. 131, No. 1, January, pp. 20–43. Crossref

Gromb, Denis, and Dimitri Vayanos, 2010, "Limits of arbitrage," *Annual Review of Financial Economics*, Vol. 2, December, pp. 251–275. Crossref

Harrison, J. Michael, and David Kreps, 1978, "Speculative investor behavior in a stock market with heterogeneous expectations," *Quarterly Journal of Economics*, Vol. 92, No. 2, May, pp. 323–336. Crossref

Hirano, Tomohiro, and Alexis Akira Toda, 2025, "Bubble necessity theorem," *Journal of Political Economy*, Vol. 133, No. 1, January, pp. 111–145. Crossref

Hong, Harrison, José Scheinkman, and Wei Xiong, 2006, "Asset float and speculative bubbles," *Journal of Finance*, Vol. 61, No. 3, June, pp. 1073–1117. Crossref

Kamihigashi, Takashi, 2006, "Transversality conditions and dynamic economic behavior," Kobe University, Research Institute for Economics & Business Administration, mimeo, January, available online.

Kamihigashi, Takashi, 1998, "Uniqueness of asset prices in an exchange economy with unbounded utility," *Economic Theory*, Vol. 12, No. 1, July, pp. 103–122. Crossref

Kindleberger, Charles P., 1978, *Manias, Panics, and Crashes: A History of Financial Crises*, New York: Basic Books.

Kocherlakota, Narayana R., 1992, "Bubbles and constraints on debt accumulation," *Journal of Economic Theory*, Vol. 57, No. 1, June, pp. 245–256. Crossref

Kogan, Leonid, Stephen A. Ross, Jiang Wang, and Mark M. Westerfield, 2006, "The price impact and survival of irrational traders," *Journal of Finance*, Vol. 61, No. 1, February, pp. 195–229. Crossref

Liu, Feng, Joseph S. White, and John R. Conlon, 2023, "A three-state rational greater-fool bubble model with intertemporal consumption smoothing," *International Economic Review*, Vol. 64, No. 4, November, pp. 1565–1594. Crossref

Morris, Stephen, 1996, "Speculative investor behavior and learning," *Quarterly Journal of Economics*, Vol. 111, No. 4, November, pp. 1111–1133. Crossref

Samuelson, Paul A., 1958, "An exact consumption-loan model of interest with or without the social contrivance of money," *Journal of Political Economy*, Vol. 66, No. 6, December, pp. 467–482. Crossref

Sandroni, Alvaro, 2000, "Do markets favor agents able to make accurate predictions?," *Econometrica*, Vol. 68, No. 6, November, pp. 1303–1341. Crossref

Santos, Manuel S., and Michael Woodford, 1997, "Rational asset pricing bubbles," *Econometrica*, Vol. 65, No. 1, January, pp. 19–57. Crossref

Scheinkman, José A., and Wei Xiong, 2003, "Overconfidence and speculative bubbles," *Journal of Political Economy*, Vol. 111, No. 6, December, pp. 1183–1220. Crossref

Simsek, Alp, 2013, "Belief disagreements and collateral constraints," *Econometrica*, Vol. 81, No. 1, January, pp. 1–53. Crossref

Tirole, Jean, 1985, "Asset bubbles and overlapping generations," *Econometrica*, Vol. 53, No. 6, November, pp. 1499–1528. Crossref

Tirole, Jean, 1982, "On the possibility of speculation under rational expectations," *Econometrica*, Vol. 50, No. 5, September, pp. 1163–1181. Crossref

Townsend, Robert M., 1980, "Models of money with spatially separated agents," in *Models of Monetary Economies*, John H. Kareken and Neil Wallace (eds.), Minneapolis: Federal Reserve Bank of Minneapolis, pp. 265–303, available online.

Weil, Philippe, 1987, "Confidence and the real value of money in an overlapping generations economy," *Quarterly Journal of Economics*, Vol. 102, No. 1, February, pp. 1–22. Crossref

Economic Perspectives is published by the Economic Research Department of the Federal Reserve Bank of Chicago. The views expressed are the authors' and do not necessarily reflect the views of the Federal Reserve Bank of Chicago or the Federal Reserve System.

ISSN 0895-0164