Bubbles and mispricing

Sam Wycherley*

A note on some models of mispricing, including but not limited to bubbles. Up until slow moving capital, the choice of papers covered here takes a lot of inspiration from Markus Brunnermeier's slides.

1 Overview and stylized facts

- A *mispricing* is a situation where a security's price deviates from the security's "fundamental value". In a frictionless world, this would present a profitable arbitrage opportunity.[1]
- A *bubble* is a situation where a security's price exceeds the security's fundamental value because investors expect to be able to resell the security at an even more favourable price (this definition goes back to at least Keynes (1931)). Thus a bubble is a form of mispricing driven by *pure speculation*. Bubbles are typically characterized by a period of explosive price behaviour (typically followed by a collapse in price) though, in general, bubbles can also be non-explosive.
- *Some famous historical bubbles.* While systematically identifying bubbles is difficult, it is easy to find clear examples of bubbles.
  - *Tulip Mania, 1634-1637.* This is probably the most famous historical bubble. The tulip was introduced to Europe in the 16th century. In the Netherlands, it became a luxury item, and there was a booming trade in propagating new varieties. Tulip bulbs infected with the tulip breaking virus, which causes multicoloured petals, were the most desirable. Increasing demand led to the price of forward contracts for bulbs infected with the virus increasing, which from 1634 led to speculators entering the forwards market. The price increased dramatically, peaking in the winter of 1636-1637. The market ubruptly collapsed in February 1637, with a 99.999% price decline. Forwards contracts written in this winter period were not honoured in practice – in Holland, the legislature voided all contracts and in Haarlem, the city eventually allowed cancellation of contracts for a 3.5% cancellation fee.

*Email: wycherley@stanford.edu.

[1] Throughout, we usually interpret the fundamental value of a security to be the present value of its discounted future cashflows (under the risk-neutral measure). In some cases, it is reasonable to consider convenience yields to be part of a security's fundamental value.

- *South Sea Bubble, 1720.* The *South Sea Company* was a joint-stock company created by the British government in 1711 century by converting government debt into equity in the new company. As part of the Treaty of Utrecht in 1713, Spain granted the South Sea Company a monopoly on the slave trade with its American colonies. The company also had a limited quota to sell goods in the Spanish Americas, which it supplemented with smuggling. From 1718, the War of the Quadruple Alliance disrupted the company's trade. From 1719, the company began to purchase British government debt funded through new equity issues. The company offered loans to investors to fund share purchases. The price of shares in the company increased dramatically throughout 1720, from £125 per share in January to £950 per share in July. By December, it had collapsed to £185. Company stock had been used as collateral for other loans, and so the collapse in share price lead to wider bank failures.
- *Dot-com bubble, 1995-2000.* Named after all the US internet start-ups that included ".com" in their name, and either collapsed or were acquired. The NASDAQ index increased 800% over the period 1995-2000, driven primarily by the increasing share prices of technology companies – for example, Qualcomm's share price increased by over 2600% in 1999. Having peaked in March 2000, the NASDAQ index subsequently fell 78% through to October 2002.
- Many other examples: *Mississippi bubble* (France, 1720, stock); *Amsterdam banking crisis of 1763* (Netherlands, banking crisis brought on by collapsing commodity bubble); *Bengal bubble* (Britain, 1769, East India Company stock); *Railway mania* (Britain, 1840s, rail company stock); *Florida land boom* (US, 1920-1926, real estate); *Poseidon bubble* (Australia, 1970, Poseidon Nickel stock); *Japanese asset price bubble* (Japan, 1986-1991, real estate and equities); *Uranium bubble* (2007, uranium ore).

• *Some notable mispricings.* Other kinds of mispricing can be easier to identify, because they sometimes appear to be obvious arbitrages.

  - *Post-GFC deviations from covered interest parity.* Since the Global Financial Crisis of 2008-9, covered interest parity has ceased to hold (Du, Tepper & Verdelhan, 2017; Du & Schreger, 2021).
  - *The TIPS-Treasury bond puzzle.* Fleckenstein, Longstaff & Lustig (2014) find US Treasury bonds are generally overpriced relative to Treasury Inflation-Protected Securities (TIPS). By buying TIPS and making inflation swaps, one can replicate a Treasury bond's cashflows precisely, at a lower cost.
  - *The on-the-run/off-the-run Treasury bond mispricing.* The most recent issue of Treasury bonds (such bonds are said to be "on-the-run") trades at a premium relative to the second-most recent issue (old issues are said to be "off-the-run"), beyond what can be explained by the term premium (Krishnamurthy, 2002).

- In a bubble, the volume of trade typically increases significantly, along with price. For example, some of the Dutch tulip forward contracts were changing hands up to five times. Volume of stock trading increases markedly in stock market bubbles (Scheinkman & Xiong, 2003; Liao & Peng, 2019, etc.) and in housing bubbles.
- Bubbles are often accompanied by consumption booms.
- As some of the examples above indicate, many bubbles have generated broader risk, for example when the asset involved has been used to collateralize lending, or where risks have been magnified through the use of derivative instruments. The US real estate bubble and the financial crisis of 2008 is a good example.
- Note dramatic growth in the price of a bubble is not necessarily evidence of a bubble. Indeed, Greenwood, Shleifer & You (2017) examine episodes where stock prices in a US industry have increased by at least 100% over two years and find this does not, on average, predict lower subsequent returns, although it does predict significantly higher probability of a subsequent crash. For example, Amazon's share price in October 2014 was around $15. Over the subsequent 2 years, it increased by over 180%, and at the end of 2023 was 300% above its October 2016 price.
- A bubble or mispricing is not *ipso facto* inefficient. For example, fiat money bubbles in overlapping generations models can improve the efficiency of equilibrium by redirecting oversaving in capital to an unproductive asset.

There are four main species of models of bubble formation/mispricings:

1. *Rational expectations equilibria with symmetric information*: Tirole (1982); Blanchard & Watson (1982). *In OLG models*: Tirole (1985); Woodford (1990); Martin & Ventura (2012).
2. *Rational expectations equilibria with asymmetric information*: Allen, Morris & Postlewaite (1993); Allen & Gorton (1993); Morris, Postlewaite & Shin (1995).
3. *Limits to arbitrage*: DeLong, Shleifer, Summers & Waldmann (1990); Shleifer & Vishny (1997); Krishnamurthy (2002); Abreu & Brunnermeier (2002, 2003); Mitchell, Pedersen & Pulvino (2007); Brunnermeier & Pedersen (2008); Duffie (2010); Duffie & Strulovici (2012); Dow, Han & Sangiorgi (2021). Note many of these are explanations of mispricings more generally, and not necessarily bubbles.
4. *Heterogeneous beliefs*: Harrison & Kreps (1978); Scheinkman & Xiong (2003); Ofek & Richardson (2003).

Finally, there is a fifth species of models that study bubble-shaped mispricing but with a non-bubble explanation (currently not covered here):

5. *Lumpy information aggregation*: Romer (1993); Avery & Zemsky (1998); Lee (1998).

2 Rational expectations bubbles

2.1 Symmetric information

Consider a discrete time rational expectations setting, fixing probability space $(\Omega, \mathcal{F}, P)$ and a filtration $\mathbb{F} = \{\mathcal{F}_t \mid t \in \mathbb{N}\}$. There is riskless borrowing with short rate process r and a risky security with price process S and dividend process δ. If π denotes the stochastic discount factor, then under no-arbitrage, we have pricing equation,

$$\pi_t S_t = \mathbb{E}_t \left[\pi_{t+1}(S_{t+1} + \delta_{t+1}) \right],$$

or letting Q denotes the risk neutral measure, we have

$$S_t = \mathbb{E}_t^Q \left[\frac{1}{1+r_t}(S_{t+1} + \delta_{t+1}) \right].$$

If the security has a finite maturity with terminal payoff X at date T, then we must have $S_T = X$ and thus S_t is uniquely pinned down for $t \leq T$, so the existence of a bubble is not possible.

Bubbles can thus only exist under no-arbitrage in this setting if there is an infinite horizon. In this case, we can have a situation where agents expect the price to blow up as $t \to \infty$. Suppose we decompose the price of the risky security into a "fundamental value" component S^f and a "bubble" component S^b, with $S_t = S_t^f + S_t^b$. The fundamental value component is the discounted sum of future dividends,

$$S_t^f = \mathbb{E}_t^Q \left[\sum_{i=1}^{\infty} R_{t,t+i} \delta_{t+i} \right],$$

where $R_{t,t+i} = \prod_{j=0}^{i-1} \frac{1}{1+r_{t+j}}$. It follows that

$$S_t^d = \lim_{T \to \infty} \mathbb{E}_t^Q \left[R_{t,T} S_T \right]$$

and so $\mathbb{E}_t^Q \left[\frac{S_{t+1}^d}{S_t^d} \right] = 1 + r_t$. Thus if there is a bubble component to the security price, then under the risk neutral measure Q, it must grow at the same rate in expectation as the money market account. This is unsurprising – it is immediate from the deflated security price process being a Q-martingale.

Example 1 (Blanchard & Watson, 1982)**.** Suppose the short rate is a constant r and the agent is risk neutral. Then a deterministic bubble involves a bubble component $S_t^b = (1+r)^t S_0^b$.

A simple non-deterministic example of a bubble here involves the bubble collapsing with probability $p \in (0,1)$ each period t (in which case $S_t^b = 0$), and persisting otherwise. Then we have $S_{t+1}^b = \frac{1+r}{1-p} S_t^b$ if the bubble persists. Note if τ is the stopping time at which the bubble collapses, then $P(\tau < \infty) = 1$.

Note that in principle, the bubble term S_t^b could be negative. This implies the price of the security would eventually become negative, at least for some states of the world. However, if there is free disposal, then such "negative bubbles" cannot occur.

So far, we have focused only on no-arbitrage. Under stronger conditions, we can rule out bubbles. Namely, Tirole (1982) shows that in any fully dynamic rational expectations equilibrium, bubbles cannot exist. There is a risky security paying dividends with fixed net supply $\bar{x}$. Suppose there is a finite set of n infinitely lived traders, each of whom is risk neutral, discounts at rate r and has a common prior. Each trader i receives a signal process σ^i from set Σ^i, and $\sigma_t = (\sigma_t^1, \ldots, \sigma_t^n)$ (in set Σ_t) is the profile of signals agents receive at t. A *forecast function* at time t is a function $\Phi_t : \Sigma_t \to \mathbb{R}_+$ that associates to the profile of signals σ_t a price $S_t = \Phi_t(\sigma_t)$. A *fully dynamic rational expectations equilibrium* is a sequence of forecast functions $\Phi = (\Phi_t)_{t \geq 0}$ such that there each agent i has a trading strategy x so that (i) at each time t, every agent i's demand $x_t^i(\sigma_t^i, S_t)$ maximizes i's expected present discounted gain conditional on i's information $(\sigma_t^i, \Sigma_t, S_t)$, (ii) markets clear, i.e. $\sum_i x_t^i(\sigma_t^i, S_t) = \bar{x}_t$ for all signal profiles $\sigma_t \in \Sigma_t$, all prices S_t, across all times t.

Regardless of whether shorting is permitted, there can be no price bubbles in a fully dynamic REE. See Proposition 6 in Tirole (1982) for a proof. Intuitively, a speculator entering the market after a bubble has started has formed can only expect to realize a gain if she expects the price to continue to increase and she can find another investor to sell the security to down the line. But at this point, we have a negative-sum game because some traders have already realized profits (at a minimum, some must have sold to the entering speculator). If there are finitely many agents, then some trader is going to eventually find they have no-one to sell to, and so we must have $\lim_{T\to\infty} S_T^b = 0$, which then implies $S_t^b = 0$ for all t.

Put another way, if the interim allocation is Pareto efficient and agents behave optimally with respect to their information, then there can be no bubbles. Note fully dynamic REE is a very strong solution concept.

2.1.1 Overlapping generations

If new investors keep arriving, then it is possible for each speculator to always find a 'greater fool' to sell a security on to. This is the case in *overlapping generations models*, as analysed by Tirole (1985).

Suppose consumers each live for two periods. In the first period, the consumer is young and inelastically supplies a unit of labour at wage w_t. In the second period, the consumer is old and enjoys retirement. The population of young consumers $L_t = (1+n)^t$ at time t grows at a rate n, with initial young population normalized to 1. There is a single consumption good, and the utility of a consumer born at t given consumption levels c_t^t while young and c_{t+1}^t while old is $u(c_t^t, c_{t+1}^t)$. Young consumers can save, earning an interest rate r_{t+1} at time $t+1$ on savings decided at time t, and can consume savings when old. Aggregate savings are $A_t = (1+n)^t s(w_t, r_{t+1})$.

On the production side, the consumption good is produced competitively using a constant returns production technology, $Y_t = L_t f(k_t)$, where $k_t = \frac{K_t}{L_t}$, with K_t being the capital stock and L_t being the labour stock at t, and f satisfying $f(0) = 0$, $f'(0) = \infty$ and decreasing marginal rates of substitution. Then the real interest rate at time t is $r_t = f'(k_t)$ and the real wage at time t can be written as $w_t = \phi(r_t)$, where ϕ is a decreasing function. At time t, the representative firm can invest at the interest rate r_{t+1}, so in equilibrium, $r_{t+1} = f'((s(w_t, r_{t+1}) - a_t)/(1+n))$, where $a_t = \frac{A_t - K_t}{L_t}$ denotes the difference between savings per capita and the level of capital stock per capita. From this, we assume there is a function ψ so that $r_{t+1} = \psi(w_t, a_t)$ with $\psi_w < 0$ and $\psi_a > 0$. Making some technical assumptions on the curves $\{w = \phi(r)\}$ and $\{r = \psi(w, 0)\}$ and letting $\bar{r} = \psi(\Phi(\bar{r}), 0)$, and taking $a_t = 0$ for all t, Diamond (1965) shows there is a unique equilibrium with interest rates converging to $\bar{r}$. Now, the optimal capital level per capita is the golden rule level k^* solving $f'(k^*) = n$, and consumer optimization implies $\frac{u_1(c_t^t, c_{t+1}^t)}{u_2(c_t^t, c_{t+1}^t)} = 1 + r_{t+1} = 1 + f'(k_{t+1})$. Thus if $\bar{r} < n$, then the capital level per capita converges to $\bar{k} > k^*$, which is Pareto inefficient since the golden rule level is feasible (there is oversaving in equilibrium). If, on the other hand, $\bar{r} > n$, then the capital level per capita converges to $\bar{k} < k^*$, which is Pareto efficient.

Tirole (1985) considers introducing a bubble asset with no fundamental value, such as *fiat money*. The aggregate value of the bubble asset per capita at time t is b_t. The bubble asset can be freely disposed of so $b_t \geq 0$. There is no uncertainty, so the bubble asset is riskless and therefore no arbitrage implies

$$b_{t+1} = \frac{1 + r_{t+1}}{1+n} b_t.$$

It is now possible for savings to be unproductive, i.e. not allocated to production, because they can be saved via the bubble asset. The per capita difference between savings and capital level is $a_t = b_t$. Now, for the Pareto efficient equilibria ($\bar{r} > n$), the unique equilibrium is bubbleless, i.e. $b_t = 0$ for all t. On the other hand if $\bar{r} \in (0, n)$, so the bubbleless equilibrium is inefficient, there is a maximum feasible initial bubble $\hat{b}_0 > 0$ such that for all $b_0 \in [0, \hat{b}_0)$, there is a unique equilibrium associated with initial bubble b_0. These bubbles vanish per capita as time grows large – i.e. while the bubble $B_t = (1+n)^t b_0$ is always growing, it grows more slowly than the real resources of the economy. There is also a unique equilibrium associated with initial bubble $\hat{b}_0$, in which case the bubble per capita converges to $\hat{b}$ solving $n = f'((s(\phi(n), n) - \hat{b})/(1+n))$. Bubbles here can improve efficiency by crowding out investment in capital, reducing overinvestment. The existence of a bubble expands the set of possible trading strategies, and thus the equilibrium allocation.

One issue with the Diamond-Tirole framework above is that bubbles tend not to be deterministic (rather, they usually burst), and are associated with consumption booms, heightened capital stock and temporarily higher output.

Martin & Ventura (2012) introduce investor shocks and capital market frictions to reconcile an OLG model with such stylized facts. Suppose now that there is no population growth and each consumer i maximizes their expected old-age consumption, $U_{it} = \mathbb{E}_t[c_{i,t+1}]$. That is, consumers are risk-neutral and do not consume only when old. Young consumers can save, but some are more skilled in saving than others. Namely, a fraction ϵ of young workers can convert a unit of output into a unit of capital, while the complement can convert a unit of output into only $\delta \in (0,1)$ units of capital. Production of the consumption good is competitive using a Cobb-Douglas production function $F(K_t, L_t) = K_t^\alpha L_t^{1-\alpha}$ for some $\alpha \in (0,1)$. Since $L_t = 1$, the real wage and interest rate are $w_t = (1-\alpha)K_t^\alpha$ and $r_t = \alpha K_t^{\alpha-1}$. Since young consumers do not consume, they save all their labour income, which is a constant fraction $1-\alpha$ of output. There is a capital market friction that prevents the less skilled investors lending to the more skilled investors, so aggregate efficiency of investment is $A := \epsilon + (1-\epsilon)\delta$. If capital is the only available asset, then aggregate savings and the capital stock are equal and so $K_{t+1} = (1-\alpha)AK_t^\alpha$.

Now suppose there are also bubbles, which are created at random and without cost. A bubble in this model is effectively a pyramid scheme, with purchasers buying a right to a share of the next period's purchasing revenues (and the original creator receiving an initial windfall equal to the initial bubble price). Let b_t denote the market value of all old bubbles, and let b_t^p and b_t^u denote the market value of all new bubbles created by the productive and unproductive investors.

Equilibria here are stochastic, driven by shocks to investor sentiment. Take any nonnegative stochastic process $(b_t, b_t^p, b_t^u)_{t\geq 0}$ for the bubbles, define $h_t = (b_t, b_t^p, b_t^u)$ to be the realization of the shock and $h^t = (h_0, \ldots, h_t)$ to be the history of shocks up to t, and let H_t be the set of all histories h^t, for each t. Call the stochastic process $(b_t, b_t^p, b_t^u)_{t\geq 0}$ an *equilibrium bubble* if (i) $b_t + b_t^p + b_t^u > 0$ (i.e. a bubble exists) for some t and $h^t \in H_t$, and (ii) there is a nonnegative capital stock process $(K_t(h^t))_{t\geq 0}$ such that all agents maximize their expected utility and markets clear, for all times t and histories $h^t \in H_t$.

Bubble owners are either old, who previously acquired bubbles when young, or young, who created new bubbles. Owners sell to the young, who are the only buyers.

Let $c_t = \frac{b_t + b_t^p}{(1-\epsilon)(1-\alpha)K_t^\alpha}$. The first order conditions for investors' portfolio problem imply that in equilibrium,

$$\mathbb{E}_t\left[\frac{b_{t+1}}{b_t + b_t^p + b_t^u}\right] \begin{cases} = \delta\alpha K_{t+1}^{\alpha-1} & \text{if } c_t < 1, \\ \in \left[\delta\alpha K_{t+1}^{\alpha-1}, \alpha K_{t+1}^{\alpha-1}\right] & \text{if } c_t = 1, \\ \alpha K_{t+1}^{\alpha-1} & \text{if } c_t > 1. \end{cases}$$

This reflects that the marginal investor of a small bubble is an unproductive investor, whereas the marginal investor of a large bubble is productive. In equilibrium, we also have that $0 \leq b_t \leq (1-\alpha)K_t^\alpha$ – that is, bubbles are nonnegative and cannot exceed the aggregate savings of the young.

The existence of bubbles here improves investment efficiency through two channels. First, bubbles attracts unproductive investors first, since these investors cannot command as high returns on capital as their productive counterparts, so are willing to purchase the bubble at a higher price. Average efficiency of capital investment increases as the composition of capital investors becomes more dominated by the productive investors. This is a crowding-out effect. Second, there is a crowding-in effect as the productive young who own newly created bubbles sell these to the unproductive young.

This generates a windfall for these productive young, which they use to invest in capital. This again raises average investment efficiency.

A few other remarks:

- Santos & Woodford (1997) point out the bubble result in the OLG model is quite fragile. For example, as soon as we introduce some infinitely lived agent (the government might carry this interpretation in an OLG model), the bubble result disappears (as does dynamic inefficiency).
- In the empirical literature, the existence of dynamic inefficiency has some support – Geerolf (2018) finds that no OECD economies fulfil the Abel et al (1989) sufficient conditions for dynamic efficiency, and Japan and South Korea unambiguous exhibit oversaving so are not dynamically efficient.

2.2 Asymmetric information

In the case of asymmetric information, Allen, Morris & Postlewaite (1993) consider the necessary conditions for a bubble to exist in a finite horizon setting. If agents are unaware of other agents' beliefs, then bubbles are possible even with a finite horizon.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. There is a risky security with price process S that pays a terminal dividend $d(\omega)$ at time T in state ω, and no interim dividends. There are n agents who can trade at $t = 0, 1, \ldots, T$. Each agent i has information given by a filtration $\mathbb{F}^i = (\mathcal{F}_{it})_{t=0}^T$ (naturally, $\mathcal{F}_{it} \supset \sigma(\{S_u \mid u \leq t\})$, so the agent knows the price history). Agent i consumes only after the final period, and has expected utility $\mathbb{E}[u(c_i(\omega)) \mid \mathcal{F}_{it}]$ for some concave Bernoulli function u, where $c_i(\omega)$ is i's final consumption in state ω. Finally, there are short sale constraints limiting the extent to which any agent can take a short position.

Allen, Morris & Postlewaite (1993) draw a distinction between two types of bubble. Let $\bar{d} = \sup\{d(\omega) \mid \omega \in \Omega\}$. There is a *strong bubble* if $P(\{\omega \in \Omega \mid S_t(\omega) > \bar{d}\}) > 0$ for some time t, i.e. with positive probability, the price of the security can exceed all possible realizations of the dividend at some point in time. This is a very strong notion – if there are not restrictions on short sales, then clearly a strong bubble would constitute an arbitrage, since shorting in the strong bubble at time t the security would deliver profit at least $S_t - \bar{d} > 0$.

A weaker notion is an *expected bubble*. In this case, with positive probability, the price S_t of the security exceeds every agent's marginal valuation, at some time t.

Expected bubbles occur at a state ω and time t in a rational expectations equilibrium only if the initial allocation is interim Pareto inefficient and for each agent i, the short-selling constraint binds at some $t' \geq t$ with positive probability (conditioning on $\mathcal{F}_{it}$). The idea here is that if at t, if an agent i's (conditional) expected marginal utility selling the security at $t + 1$ exceeds the expected marginal utility if selling the security at t, then that agent expects to profit from buying the security at t and selling it at $t + 1$. If the agent is not at the short-sale position limit at t and the agent's expected utility from selling the security now exceeds that of selling the security at $t+1$, the agent expects to profit from shorting at t until the short-sale position limit binds and closing the short position at $t+1$. Thus if the short-selling does not bind for agent i at t, the price must equal agent i's expected marginal utility from selling the security at t, conditional on $\mathcal{F}_{it}$. By induction, the agent must also expect that at all future prices, the price does not exceed the agent's marginal utility from selling the security at t. Since the price at T is d, it follows that the agent i's marginal valuation does not exceed d. Thus an expected bubble requires that for all agents, the short sales constraint might bind at some point in the future.

For the existence of strong bubbles, more conditions are necessary, in addition to the above. First, in a strong bubble, every agent has to have private information. Second, there cannot be common knowledge of agents' trades – such common knowledge "negates" private information since each agent i can then make inferences about everyone else's information by observing their trades.

Note these necessary conditions are not sufficient. For example, Morris, Postlewaite & Shin (1995) provide weaker mutual knowledge conditions that rule out strong bubbles.

3 Limits to arbitrage

Limits to arbitrage models can accommodate bubbles by relaxing the no-arbitrage assumption. The traditional concept of an arbitrage involves no risk on the part of the arbitrageur and no capital. If there is an arbitrage, then there exist two portfolios of securities that generate the precise same future cashflows but trade at different market valuations today – the arbitrageur merely needs to take a long position on the lower valuation portfolio and short the higher valuation portfolio, profiting the difference.

In practice, arbitrage operations are risky and involve committing capital. For example, consider two identical futures contracts listed on different exchanges (Shleifer & Vishny make the following point using the example of Bund futures contracts listed on the LIFFE (London) and DTB (Frankfurt) exchanges, which really tells you when that paper was published (1997)). Say the two futures contracts trade at different prices. Absent any frictions, an arbitrageur can sell the more expensive futures contract and buy the cheaper futures contract to realize a riskless profit. However, in practice, the arbitrageur will need to post good faith money for both transactions, to signal to the counterparty that she intends to follow through on the agreed contracts – this will be returned when closing her position. At each trading date, as the price of the contracts change, each exchanges charges whichever of the two parties in each contract has netted a loss on that contract (either the arbitrageur or her counterparty) their loss and transfers this to the gaining party. If the price of the two contracts converge, the arbitrageur nets a gain and can close her position. However, in the short term, the price of the two contracts might diverge further, in which case the capital cost to the arbitrageur grows. Even though the price of the two contracts are guaranteed to converge in the long run, in the short run, if the arbitrageur does not have access to sufficient capital, she may be forced to close early at a loss.

Here, we go through some risks and costs arbitrageurs face, which can limit the ability of arbitrageurs to correct mispricing.

3.1 Trading costs

If arbitrageurs face trading costs, then they may not find it profitable to close arbitrage opportunities that would exist in the absence of such costs. There can therefore be persistent pricing differences between securities that deliver similar cashflows, for example.

A notable example is the persistent spread between on-the-runs and off-the-runs in the market for US Treasury bonds, analysed by Krishnamurthy (2002). Consider the market for 30-year bonds. These are typically issued at 6 month intervals. When the Treasury issues a new 30-year bond, it is said to be "on-the-run", and all previously-issued bonds are said to be "off-the-run". On-the-runs are used as a benchmark and command a significant premium relative to most recently issued off-the-run (both have a similar, distant maturity date).[2] Suppose an arbitrageur engages in the following convergence trade. She shorts the new on-the-run at issuance and buys the most recently issued off-the-run, when the spread between these bonds is high. Then at the next issuance auction, she closes out her position. At this point the spread between the two bonds is much lower (for example, in February 2001, the spread between a newly issued on-the-run 30-year and the most recently issued off-the-run was 12 bps, whereas the spread between the most recently issued off-the-run and the previously issued off-the-run was 3 bps (Fig. 1, Krishnamurthy, 2002)).

However, Krishnamurthy (2002) finds that the average profit for an arbitrageur engaging in this convergence trade is close to zero. The reason is because on average, shorting costs eat up the profit. To enter the short position on the new on-the-run, the arbitrageur has to borrow the on-the-run bond on the repo market via a reverse-repo agreement, and then sell the bond in the Treasury market. An overnight reverse-repo agreement typically requires that the borrower of the bond deposits a cash amount P_t equal to the value of the bond with a counterparty. The counterparty lends the bond to the borrower, and pays an overnight financing rate f_t on the deposited cash. This transaction requires that a bond is actually available to loan, so there is a supply constraint on the number of reverse-repo agreements that can be entered into. When the supply constraint does not bind, the overnight financing rate f_t equals the overnight risk-free rate r_t, since in equilibrium any lender is indifferent between holding the bond herself and lending it and investing the deposited cash at the riskless rate. If the supply constraint binds, however, then if the lenders were to offer financing rate f_t, the reverse repo market would not clear – hence the lenders set the overnight financing rate $f_t < r_t$ so that the market clears, and can extract a rental price $s_t = r_t - f_t$ (a measure of *specialness*) for the reverse-repo.[3]

Krishnamurthy (2002) finds that on average, the specialness of the reverse-repos extinguish the profitability of the convergence trade – when the spread between the on-

[2]The standard explanation for this premium is liquidity. The question (which Krishnamurthy (2002) resolves) is why that premium cannot be arbitraged away.

[3]A reasonable question here is: if the suppliers can earn rents, why wouldn't we expect entry on the supply side until the rents vanish? For a discussion of that question, see Duffie (1996).

the-run and off-the-run tends to be highest, so are the rents extracted by lenders in the repo market. Thus arbitrageurs earn a lower interest on the capital they must put up for shorting than the risk-free rate. This is a carry cost that on average, eats up the arbitrage profits that would otherwise be available. While the convergence trade can be profitable, profitability is a question of timing the market well – poor timing results in losses.

3.2 Noise trader risk

Noise traders are traders who have false, stochastic beliefs, treating random "noise" as if it were informative. The terminology dates back to Kyle (1985) and Black (1986). Friedman (1953) and Fama (1965) asserted noise traders cannot influence prices persistently because any price impact will be arbitraged away by more sophisticated investors. Moreover, noise traders that do move prices will incur losses on average and be driven out of the market. It turns out this argument is not very watertight, because if risk-averse arbitrageurs have short horizons, the risk that noise traders will push the price of a security in an unfavourable direction acts to limits arbitrage activity and potentially lets the arbitrageurs earn high average returns.

The presence of noise traders pushing the price of a security away from its fundamental value does not only expose arbitrageurs to noise trader risk. Arbitrageurs may also be exposed to fundamental risk, in the sense that the dividend on the security is uncertain, and so if arbitrageurs are risk averse, their arbitrage activity may be limited by their preference to avoid taking on too much risk. Figlewski (1979), Shiller (1984) and Campbell & Kyle (1987) lay out a lot of the groundwork for this class of models. However, noise trader risk models cover similar ideas but are more interesting, because the source of risk is endogenous to the model.

3.2.1 DeLong, Shleifer, Summers & Waldmann (1990)

DeLong, Shleifer, Summers & Waldmann (1990) highlight how introducing noise traders into the market can expose arbitrageurs to risk. If arbitrageurs are risk averse and have short horizons, then the risk that noise traders push prices in an unfavourable direction will limit arbitrageur activity. Remarkably in their setting, noise traders can earn higher returns than sophisticated investors by having greater exposure to noise trader risk, even though there is no fundamental risk, providing a neat argument for why noise traders can survive in the market.

Fix probability space $(\Omega, \mathcal{F}, P)$ and filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t\in\mathbb{N}}$. The setting is an overlapping generations model. Agents live for 2 periods, earn some fixed income $y > 0$ in the first period which they invest in a portfolio, consuming their wealth only in the second period. There are two securities – a safe security and an unsafe security. Both securities pay an identical fixed dividend r. The safe security is perfectly elastically supplied, so one unit of the safe security can be converted into one unit of the consumption good, and is the numeraire. The unsafe security is inelastically supplied – supply of this security is fixed at one unit. Let S be the price process for the unsafe security.

There are two types of agent. There is a mass $\mu \in (0,1)$ of noise traders and a mass $1-\mu$ of sophisticated traders. Sophisticated traders are rational and correctly understand the (normal) distribution of prices for the unsafe security. Noise traders have random, generally incorrect beliefs about the expected price of the unsafe security, but have a correct understanding of the variance of the price. In particular, in period t, if the actual mean price of the unsafe security is μ_{t+1} in period $t+1$, then noise traders believe the mean price is $\mu_{t+1} + \rho_t$, where $\rho_t \sim N(\rho^*, \sigma_\rho^2)$, with ρ^* being a measure of the bullishness of the noise traders and $\sigma_\rho^2 > 0$ being the variance of the noise traders' misperceptions. The misperception ρ_t is common to all noise traders at time t, and iid across time.

All agents have constant absolute risk aversion utility function in second-period wealth (which is equivalent to consumption), i.e. $U(w) = -e^{-2\gamma w}$, with coefficient of absolute risk aversion 2γ. Maximizing expected utility here is equivalent to maximizing the certainty equivalent. A portfolio can be characterized by the quantity θ of the unsafe security purchased, with the remainder of the agent's first-period income invested in the safe security.[4] If a portfolio θ has market value $W_\theta \sim N(\mu_\theta, \sigma_\theta^2)$ in the next period $t+1$, then the agent's certainty equivalent is $\mu_\theta - \gamma\sigma_\theta^2$.

For the representative young sophisticated trader i at time t, the problem is one of choosing quantity θ_{it} to maximize

$$\mu_{\theta_{it}} - \gamma\sigma_{\theta_{it}}^2 = y(1+r) + \theta_{it}\left(r + \mathbb{E}_t[S_{t+1}] - S_t(1+r)\right) - \gamma\theta_{it}^2 \operatorname{var}_t(S_{t+1}),$$

where $\mathbb{E}_t[\cdot] = \mathbb{E}[\cdot \mid \mathcal{F}_t]$ and $\operatorname{var}_t(\cdot) = \operatorname{var}(\cdot \mid \mathcal{F}_t)$. Solving the first-order condition gives optimal demand

$$\theta_{it} = \frac{r + \mathbb{E}_t[S_{t+1}] - S_t(1+r)}{2\gamma \operatorname{var}_t(S_{t+1})}$$

for the unsafe security.

For the representative young noise trader n at time t, on the other hand, the problem is to choose quantity θ_{nt} to maximize

$$\mu_{\theta_{nt}} - \gamma\sigma_{\theta_{nt}}^2 = y(1+r) + \theta_{nt}\left(r + \rho_t + \mathbb{E}_t[S_{t+1}] - S_t(1+r)\right) - \gamma\theta_{nt}^2 \operatorname{var}_t(S_{t+1}).$$

That is, the noise trader's certainty equivalent includes the misperception ρ_t about the expected price next period. The first order condition gives optimal demand

$$\theta_{nt} = \frac{r + \mathbb{E}_t[S_{t+1}] - S_t(1+r)}{2\gamma \operatorname{var}_t(S_{t+1})} + \frac{\rho_t}{2\gamma \operatorname{var}_t(S_{t+1})}$$

for the unsafe security. As we might expect, if noise traders are bullish on average, then they will have greater demand than a sophisticated investor, so will take on more of the risk inherent in the unsafe security.

Since the supply of the unsafe security is fixed at quantity 1, we have market clearing condition,

$$(1-\mu)\theta_{it} + \mu\theta_{nt} = 1.$$

[4]This remainder can exceed the first-period income (if shorting the unsafe security) or be negative (if shorting the safe security).

This implies the price S_t of the unsafe asset at t must satisfy

$$S_t = \frac{1}{1+r}\left(r + \mathbb{E}_t[S_{t+1}] - 2\gamma\,\mathrm{var}_t(S_{t+1}) + \mu\rho_t\right).$$

Iterating forward,

$$\mathbb{E}_t[S_{t+1}] = \frac{1}{1+r}\left(r + \mathbb{E}_t[S_{t+2}] - 2\gamma\,\mathrm{var}_t(S_{t+2}) + \mu\rho^*\right).$$

This gives us a difference equation. DeLong et al. focus on steady-state equilibria here, so the unconditional distribution of S_{t+1} is identical to the unconditional distribution of S_t. Solving, we have

$$S_t = 1 + \frac{\mu(\rho_t - \rho^*)}{1+r} + \frac{\mu\rho^*}{r} - \frac{2\gamma}{r}\,\mathrm{var}_t(S_{t+1}).$$

On the right hand side, only ρ_t is a random variable, with $\mathrm{var}_t(S_{t+1}) = \mathrm{var}(S_{t+1}) = \left(\frac{\mu\sigma_\rho}{1+r}\right)^2$. Hence,

$$S_t = 1 + \frac{\mu(\rho_t - \rho^*)}{1+r} + \frac{\mu\rho^*}{r} - \frac{2\gamma\mu^2\sigma_\rho^2}{r(1+r)^2}.$$

The price of the unsafe security in equilibrium is thus distorted by the impact of noise traders, through the final three terms in this expression. The term $\frac{\mu(\rho_t-\rho^*)}{1+r}$ is the impact on price of variations in the noise traders' beliefs about their mean – the price is increasing in how bullish the noise traders are relative to their average beliefs. The term $\frac{\mu\rho^*}{r}$ is the impact on price of noise traders' average misperception – the more bullish they are on average, the higher the price of the unsafe security on average.

Finally, the term $-\frac{2\gamma\mu^2\sigma_\rho^2}{r(1+r)^2}$ captures the impact of the noise trader risk – for sophisticated investors to hold the unsafe security at t, they demand compensation for bearing the risk that the next generation of noise traders at $t+1$ will be more bearish than the current generation. Suppose $\rho^* > 0$, so noise traders are bullish on average. Then noise trader risk drives out the sophisticated arbitrageurs to some extent, driving down the average price of the unsafe security relative to its price if there was no noise-trader risk (i.e. if $\sigma_\rho^2 = 0$). Because noise traders are taking on more of this risk, they can potentially earn higher returns on average than sophisticated traders. The relative difference in returns at t is

$$\Delta R_{n-i,t+1} = (\theta_{nt} - \theta_{it})(r + S_{t+1} - S_t(1+r)).$$

Substituting for $\theta_{nt} - \theta_{it}$ and taking unconditional expectations, we get

$$\mathbb{E}[\Delta R_{n-i,t+1}] = \rho^* - \frac{(1+r)^2(\rho^*)^2 + (1+r)^2\sigma_\rho^2}{2\gamma\mu\sigma_\rho^2}.$$

This captures several effects. The first term ρ^* captures that the higher the noise traders' misperception, the more of the unsafe security they hold and thus the higher average reward they gain from exposure to noise trader risk. The two numerator terms capture two effects – 1. the more bullish noise traders are, the more they drive up the price, reducing the return to noise trader risk and 2. noise traders time the market poorly, buying high and selling low, reducing expected returns. The denominator captures that the greater the noise trader risk, the more the price of the unsafe security is driven down as sophisticated traders demand a greater risk premium. If the noise trader risk is large enough, the effects of bullishness and bad market timing are swamped by the extra reward the noise traders gain from taking on extra risk.

Thus we can have that noise traders earn greater profits in expectation than sophisticated traders. Note the results here do not depend on the OLG setting – it can be reformulated in a setting with infinitely-lived agents as long as the sophisticated arbitrageurs are myopic – and so we have a reason for why noise traders won't be driven out of the market (in fact, they might come to dominate it). The caveat here is this is a partial model. Since DeLong, Shleifer, Summers & Waldmann (1990), a large *market selection* literature has developed that seeks to understand under what conditions agents with false beliefs will be driven out of the market.[5]

Finally, a point on welfare. While noise traders can earn higher returns than sophisticated investors, this is only because their misperceptions lead them to take on excessive risk given their preferences. If they were to learn the actual return distribution of the unsafe security, they would not be willng to take on this risk. This is the reason why sophisticated investors do not simply imitate noise traders.

3.2.2 Shleifer & Vishny (1997): performance-based arbitrage

The overlapping generations model of DeLong et al. (1990) is a bit of an extreme case – arbitrageurs only care about consumption in the second period of their life, since they die after "retirement". As mentioned, the results of the model carry through to a setting with infinitely-lived traders provided those traders are myopic. The myopia is really key here – in DeLong et al's setting, if there are instead infinitely-lived traders and the unsafe security trades above the discounted value of its dividends, i.e. $S_t > 1$, then an arbitrageur could short a unit of the unsafe security indefinitely and purchase a unit of the safe security. Since the two securities have identical dividend payments, the arbitrageur can meets all future dividend payments on the short position, and realizes a risk-free profit $S_t - 1 > 0$ at time t.

In reality, the traders typically in the position to execute arbitrages are institutional, and are thus effectively permanent entities – banks, hedge funds, mutual funds, and so on – who manage other investors' money. Thus an important question if we think noise trader risk matters is: why would these traders be myopic? Shleifer & Vishny (1997) provide one reason that covers a large range of institutional traders – liquidity risk. As mentioned earlier, arbitrage typically requires putting up capital. For institutional traders, say a fund, this capital comes from investors who have invested in the fund, and who care about the performance of their investment. Even if an arbitrage trade

[5] I might cover this in another note because it seems interesting.

guarantees long run risk-free profits, short-run paper losses might force the fund to exit the arbitrage position at a realized loss if investors demand their investments back early, as they are want to do if poor short-run performance persuades them that the fund manager is low quality (or misguided).[6]

A nice example of this is Scion Capital (popularized in the book/film *The Big Short*). Scion purchased a large quantity of credit default swaps (CDS) on subprime mortgage-backed instruments in the years prior to the Great Financial Crisis. Until the underlying securities defaulted, the fund booked sizeable short-run losses due to CDS premium costs. The losses caused investors in the fund to attempt to pull their investments (due to the unusual design of the fund, many irate investors struggled to do this). When the underlying securities eventually defaulted, the fund's profits dwarfed the short-run losses and remaining investors made a substantial profit.

Shleifer & Vishny (1997) illustrate the point in a simple model. There are three types of agents: risk-neutral arbitrageurs (which are open-end funds), noise traders, and risk-neutral investors in the arbitrageurs. There is a market for a single security in unit supply, and the arbitrageurs are fully specialized in trading in this market alone, whereas investors are free to allocate their funds to outside options (such as to arbitrageurs in other markets). Time is discrete and the market lasts three periods. The price of the security at time t is S_t.

The fundamental value of the security is known to the arbitrageurs to be V. At time $t = 3$, the fundamental value V becomes common knowledge to all market participants, and so the price of the security becomes $S_3 = V$. Suppose noise traders are pessimistic. That is, in the first two periods, the representative noise trader experiences pessimism shocks $\epsilon_t \geq 0$, so that aggregate noise trader demand for the security at $t < 3$ is $Q_t^n = \frac{V - \epsilon_t}{S_t}$. To make the point, suppose $\epsilon_1 < 0$ is deterministic and

$$\epsilon_2 = \begin{cases} \bar{\epsilon}_2 & \text{with probability } q, \\ 0 & \text{with probability } 1 - q, \end{cases}$$

for $q \in (0, 1)$ and $\bar{\epsilon}_2 > \epsilon_1 > 0$. Total capital available to arbitrageurs at t is F_t. The initial available capital F_1 is exogenous, and the arbitrageurs know the size of the period 1 pessimism shock, investing D_1 in the security at $t = 1$, so arbitrageurs' aggregate demand is $Q_1^a = \frac{D_1}{S_1}$. Thus the price of the security in period 1 is $S_1 = V - \epsilon_1 + D_1$. At $t = 2$, if noise traders become more pessimistic (assume $\bar{\epsilon}_2 > F_2$, for simplicity), then arbitrageurs

[6] Whether investors can alter the capital available to a fund manager depends on the structure of the investment vehicle. Closed-end funds (such as closed-end mutual funds) raise capital through an initial equity issue (potentially, they may also borrow if they are leveraged). In this case, the investor is simply a shareholder in an investment corporation that has issued a fixed supply of shares, who can only exit her position by selling the shares to another investor. The capital available to the closed-end fund is thus not subject to outflows from investors pulling their capital. Open-end funds (such as open-end mutual funds or hedge funds) issue new shares whenever an investor invests in the fund, and investors have the right to have these shares redeemed directly by the fund, usually at any time. Such funds thus have the kind of structure Vishny & Shleifer (1997) study. Both open-end and closed-end funds can also increase available capital by borrowing. In this case, they may face liquidity risk if the debt covenant allows creditors to call the debt early or declare the fund insolvent under certain circumstances.

invest all available capital, so $D_2 = F_2$. In this case, we have $S_2 = V - \epsilon_2 + F_2$. If noise traders stop being pessimistic ($\epsilon_2 = 0$), then $S_2 = V$ and the arbitrageurs invest in cash.

Arbitrageurs are perfectly competitive so set fees equal to marginal cost, which we assume is some constant. The investors who invest in the arbitrageurs are assumed not to have any information on the arbitrageurs' strategy, because in reality such a strategy is potentially very complex, and arbitrageurs are unwilling to share valuable information on their strategies that could lead to copying. Investors thus update their beliefs about the arbitrageurs' future returns purely on past performance. Their aggregate supply of funds to the arbitrageurs in this market in period 2 is increasing in the arbitrageurs' gross return between periods 1 and 2, given by a function $G(x) = ax + 1 - a$ for some constant $a \geq 1$ (more generally, we can take G to be concave with $G' \geq 1$ without affecting the results). Arbitrageurs' total available capital at $t = 2$ is therefore

$$\begin{aligned} F_2 &= F_1 G\left(\frac{D_1}{F_1}\frac{S_2}{S_1} + \frac{F_1 - D_1}{F_1}\right) \\ &= F_1 - aD_1\left(1 - \frac{S_2}{S_1}\right). \end{aligned}$$

This has the interpretation that arbitrageurs gain (lose) available capital if they outperform (underperform) some outside option benchmark return for investors, which is normalized to zero. In a very reduced-form way, this is capturing that investors face a signal extraction problem: poor performance could be due to increasing pessimism from noise traders or it could be due to poor ability on the part of the arbitrageurs; investors place higher probability on an arbitrageur being low-skilled in response to observing poor performance, leading to capital outflows.

Arbitrageurs maximize expected profits at $t = 3$, which under the competitiveness assumption means maximizing assets under management. Focus first on $t = 2$. If $\epsilon_2 = 0$, then the price of the security rebounds to V, so arbitrageurs can optimally liquidate their position and hold cash, and thus

$$F_3 = a\left(D_1\frac{V}{S_1} + F_1 - D_1\right) + (1-a)F_1.$$

If $\epsilon_2 = \bar{\epsilon}_2$, so noise trader sentiment has worsened, then

$$F_3 = \frac{V}{S_2}\left(a\left(\frac{D_1 S_2}{S_1} + F_1 - D_1\right) + (1-a)F_1\right).$$

The arbitrageurs' problem is therefore to maximize expected period 3 assets under management,

$$\begin{aligned} \max_{D_1 \in [0, F_1]} (1-q)&\left(a\left(D_1\frac{V}{S_1} + F_1 - D_1\right) + (1-a)F_1\right) \\ &+ q\frac{V}{S_2}\left(a\left(\frac{D_1 S_2}{S_1} + F_1 - D_1\right) + (1-a)F_1\right). \end{aligned}$$

We have first order condition,

$$(1-q)a\left(\frac{V}{S_1}-1\right)+qa\left(\frac{S_2}{S_1}-1\right)\frac{V}{S_2}\geq 0,$$

with strict inequality iff $D_1 = F_1$. If there is sufficient risk that the price deteriorates further in period 2 (that is, if q is above some threshold level q^*), then arbitrageurs will invest $D_1 < F_1$ in the security in period 1, holding back some cash to take advantage in case the price falls even further. If the risk of price deterioration is sufficiently low ($q \leq q^*$) then arbitrageurs will be fully invested in the security at $t = 1$. If the arbitrageurs are sufficiently invested at $t = 1$, then a deterioration in noise trader sentiment at $t = 2$ will drive down the price of the security, resulting in capital outflows from the arbitrageurs. For example, suppose the arbitrageurs are fully invested at $t = 1$ and noise trader sentiment worsens at $t = 2$. If $a > 1$, then $F_2 < D_1 = F_1$ and $\frac{F_2}{S_2} < \frac{D_1}{S_1}$. Here, investors pull funds from the arbitrageurs and demand that the arbitrageurs redeem them in cash. Because the arbitrageurs are fully invested, they can only raise cash by a fire sale, liquidating their position in the underpriced security at an extremely unfavourable price and thus locking in a paper loss. This is even though holding the position to $t = 3$ would guarantee a profit. Assuming $aF_1 < S_1$ (so arbitrageurs do not exit completely when there is a negative sentiment shock), the price at $t = 2$ in this case is

$$S_2 = \frac{V - \bar{\epsilon}_2 - aF_1 + F_1}{1 - a\frac{F_1}{S_1}}.$$

Differentiating, we see that

$$\frac{\partial S_2}{\partial \bar{\epsilon}_2} < -1 \quad \text{and} \quad \frac{\partial^2 S_2}{\partial a \partial S} < 0.$$

That is, prices fall more than one-for-one with the deepening noise trader sentiment shock, since withdrawals leave the arbitrageurs with less available capital to counter the shock. The more sensitive investors are to performance (i.e. the larger the magnitude of a), the more extreme this mispricing becomes. This is a nice example where arbitrageurs are the least effective when mispricings are the largest.

3.3 Synchronization risk

Abreu & Brunnermeier (2002, 2003) highlight another kind of risk that limits arbitrage activity. In some cases, a single arbitrageur might be large enough relative to the mispriced market to be able to correct mispricings alone. However, in many markets, arbitrageurs are small relative to the market, and so face a coordination problem in exploiting an arbitrage opportunity. Abreu & Brunnermeier term the risk of miscoordination "synchronization risk".

Abreu & Brunnermeier (2003) focus on the case of bubbles. Arbitrageurs in their setting understand the bubble will collapse at some point, but prior to that point can profit from "riding the bubble" while the security price continues to grow. An arbitrageur's problem is thus one of market timing.

Suppose there is a stock price index whose value initially coincides with its fundamental value and grows at the risk-free interest rate r. Arbitrageurs (who are all a priori identical) are fully invested in the stock market. At a normalized time $t = 0$, the index price S_0 is normalized to 1, and the price starts to grow at a rate $\gamma > r$, so $S_t = e^{\gamma t}$ for $t > 0$. This is initially justified by improving fundamentals (for example, a series of positive shocks) but at a random time $t_0 > 0$, the growth rate γ becomes uncoupled from improving fundamentals and we have a bubble. This random time t_0 is exponentially distributed with intensity λ, so we have cdf $F(t_0) = 1 - e^{-\lambda t_0}$. Beyond t_0, some fraction $\beta(t - t_0)$ of the price is a bubble component whereas the remaining fraction $1 - \beta(t - t_0)$ reflects fundamentals, with β being a strictly increasing continuous function. Irrationally exuberant behavioural traders keep the bubble going (due to mistaken beliefs about the fundamentals, e.g. believing internet stocks will continue to grow because e-commerce will revolutionize the economy).

The behavioural traders have an absorption capacity κ. If the selling pressure of arbitrageurs exceeds κ at time $t_0 + \tau$, the price drops by fraction $\beta(\tau - t_0)$ to the fundamental value, and grows at rate r thereafter. Moreover, regardless of selling pressure, the bubble will exogenously burst at a date $t_0 + \bar{\tau}$, once it reaches a maximum size $\bar{\beta}$ as a fraction of price.

From the moment t_0 that the fundamental value of the security and its price begin to diverge, arbitrageurs start to realize. However, they do so sequentially. For any arbitrageur, awareness "arrives" at a rate $\frac{1}{\eta}$, iid across arbitrageurs, and thus by the exact law of large numbers, a fraction $\frac{t-t_0}{\eta}$ of arbitrageurs are aware the security is overpriced at $t \in [t_0, t_0 + \eta]$, with all arbitrageurs aware of the bubble by $t_0 + \eta$. Let arbitrageur i be an arbitrageur who becomes aware of the bubble at time t_i. At t_i, she does not know what fraction of arbitrageurs have already learned that a bubble has started to form, and thus she cannot infer t_0. Her posterior belief about t_0 is a distribution with support on $[t_i - \eta, t_i]$ (she can infer that $t_0 \geq t_i - \eta$, since even if she became informed last, not more than η units of time would have passed). In particular, her posterior has cdf,

$$F(t_0 \mid t_i) = \frac{e^{\lambda\eta} - e^{\lambda(t_i - t_0)}}{e^{\lambda\eta} - 1} \qquad \text{for } t_0 \in [t_i - \eta, t_i].$$

To ensure arbitrageurs do not prefer to close out their position before becoming aware of the bubble, assume

$$\frac{\lambda}{1 - e^{-\lambda\eta\kappa}} < \frac{\gamma - r}{\beta(\eta\kappa)}.$$

The interesting case will be when a bubble is *persistent*, that is, when the bubble survives beyond the time $t_0 + \eta\kappa$ at which a large enough mass of arbitrageurs are aware of the existence of the bubble to be able correct the mispricing.

To avoid arbitrageurs taking infinite positions, suppose there are bounds on the magnitude of long and short positions each arbitrageur can take. Without loss, an arbitrageur i's position at t can be taken to be $\theta_{it} \in [0,1]$, where 0 is the maximum long position and 1 is the maximum short position. The process θ_i summarizes arbitrageur t_i's position across time, and is given by $\theta_{it} = \sigma(t, t_i)$ where $\sigma : [0,\infty) \times [0,\infty) \to [0,1]$ is arbitrageur i's strategy, a measurable function of time t and the instant t_i at which i becomes aware of the bubble. Aggregate selling pressure at time $t \geq t_0$ is $s(t, t_0) = \int_{t_0}^{\min\{t, t_0+\eta\}} \sigma(t, t_i)\, \mathrm{d}t_i$. The "bursting time" of the bubble is a stopping time T^* defined by

$$T^*(t_0) = \inf\left(\{t \mid s(t, t_0) \geq \kappa\} \cup \{t_0 + \bar{\tau}\}\right).$$

Arbitrageur i's posterior belief about the bursting time has cdf

$$G(t \mid t_i) = \int_0^t 1_{T^*(t_0)<t} F(\mathrm{d}t_0 \mid t_i).$$

If an arbitrageur trades exactly at the instant when the bubble bursts, she trades at the pre-crash price $S_{T^*} = e^{gT^*}$ if $s(T^*, t_0) \leq \kappa$, whereas if $s(T^*, t_0) > \kappa$, then only a randomly chosen fraction of orders are fulfilled at the precrash price, with the remainder being fulfilled at the post-crash price $(1 - \beta(T^* - t_0))S_{T^*}$. The expected execution price is thus $[1 - \alpha + \alpha(1 - \beta(T^* - t))]S_{T^*}$ at the point where the bubble bursts, with $\alpha = 0$ if $s(t, T^*) \leq \kappa$ and $\alpha > 0$ otherwise. When trading at a time t, arbitrageurs face a transaction cost ce^{rt} for a constant $c > 0$. This means the present value of the cost of trading at t is c, and arbitrageurs will only modify their position finitely many times since trading infinitely many times incurs infinite cumulative costs.

Equilibrium in this setting is a perfect Bayesian equilibrium in which any arbitrageur who is not fully invested correctly believes all other arbitrageurs who became aware of the bubble before her also are not fully invested. The unique equilibrium strategy is a "trigger strategy" in which each arbitrageur i remains fully invested up to some point at which she sells out completely (i.e. takes a maximum short position), and remains fully shorted thereafter until the bubble bursts.[7] We will not prove uniqueness but the proof is in section 4 of Abreu & Brunnermeier (2003). Restricting to these threshold strategies, let $T(t_i)$ be the threshold time at which an arbitrageur becoming aware of the bubble at time t_i switches to shorting. By definition of equilibrium, $T(t_j) \leq T(t_i)$ whenever $t_j < t_i$, so all arbitrageurs who already knew about the bubble have divested by the time arbitrageur i divests. Likewise, if i has not yet divested, then nor have those arbitrageurs who learned about the bubble after her. It follows that unless the bubble bursts exogenously, the bubble will burst at the threshold time of the arbitrageurs who are in the κth quantile in terms of learning times. Hence the bursting time is

$$T^* = \min\{T(t_0 + \eta\kappa), t_0 + \bar{\tau}\}.$$

[7] Models of speculative currency attacks as global games (e.g. Morris & Shin, 1998) typically restrict to binary actions (attack or refrain from attacking). One might complain the action space should be much richer in these models – Abreu & Brunnermeier (2003) is nice because it justifies why restricting to binary actions can be without loss, although note that the setting of Abreu & Brunnermeier (2003) is not a global game, so the comparison is not direct.

Now at arbitrageur i's threshold time $T(t_i)$, i cannot believe that more than a fraction κ of arbitrageurs learned about the bubble before her, because then the bubble would have already burst and she would obtain the post-crash price. One can prove from this that T^* is strictly increasing and continuous in t_0 (and thus so is its inverse), and that T is continuous. It also turns out that arbitrageur i will believe the bubble will burst at $T(t_i)$ with probability zero. Abreu & Brunnermeier restrict to absolutely continuous inverse burst times T^{*-1} such that the unconditional cdf for the burst time, $G(t) = F(T^{*-1}(t))$, is absolutely continuous. Recalling the conditional cdf for the burst time is $G(t \mid t_i)$, let $g(t \mid t_i)$ denote the corresponding conditional density. Denote the conditional hazard rate that the bubble bursts at t by $h(t \mid t_i) = \frac{g(t|t_i)}{1-G(t|t_i)}$.

Now, the expected payoff for arbitrageur i (who learns there is a bubble at time t_i) under the trigger strategy selling at t is

$$\int_{t_i}^{t} e^{-ru}(1 - \beta(u - T^{*-1}(u)))S(u)g(u \mid t_i)\,\mathrm{d}u + e^{-rt}S_t(1 - G(t \mid t_i)) - c.$$

Differentiating with respect to t, we have from the first order condition that the threshold $T(t_i)$ solves

$$h(T(t_i) \mid t_i) = \frac{\gamma - r}{\beta(T(t_i) - T^{*-1}(T(t_i)))}.$$

The right hand side is a "cost-benefit ratio." Once arbitrageur i's hazard rate hits that ratio, she attacks, switching from her maximum long position to a maximum short position.

Bubbles in Abreu & Brunnermeier's model are persistent but will always burst (if not endogenously, then because they burst at $t = t_0 + \bar{\tau}$). For any arbitrageur, the first-best is to attack just before the bubble bursts, when the security price is at its zenith. Consider arbitrageur i and suppose the other arbitrageur's never attack. Then arbitrageur i's best response is to attack at time $t_i + \tau^1$. If all other arbitrageurs adopt this best response, then i's best response is to attack at time $t_i + \tau^2$. Given τ^k, we can thus generate a new best response waiting time τ^{k+1}. This process removes symmetric trigger strategies that are not rationalizable. Now, suppose $\frac{\lambda}{1-e^{-\lambda\eta\kappa}} \leq \frac{\gamma-r}{\bar{\beta}}$. In this case,

$$\lim_{k\to\infty} \tau^k = \tau^1 = \bar{\tau} - \frac{1}{\lambda}\ln\left(\frac{\gamma - r}{\gamma - r - \lambda\bar{\beta}}\right) < \bar{\tau},$$

and the bubble bursts exogenously at time $t_0 + \bar{\tau}$. On the other hand, if $\frac{\lambda}{1-e^{-\lambda\eta\kappa}} > \frac{\gamma-r}{\bar{\beta}}$, then in equilibrium, arbitrageur selling pressure will exceed κ before the bubble exogenously bursts. In this case, if an arbitrageur i learns about the bubble at $t_i \geq \eta\kappa$, she attacks at $t_i + \tau^*$ where

$$\tau^* = \beta^{-1}\left(\frac{\gamma - r}{\lambda/(1 - e^{-\lambda\eta\kappa})}\right) - \eta\kappa.$$

If $t_i < \eta\kappa$, then the arbitrageur attacks at $\eta\kappa + \tau^*$ (her posterior beliefs differ here because we have the truncation $t_0 \geq 0$). In any case, the bubble bursts as soon as the bubble fraction of the price reaches $\beta^* = \frac{\gamma-r}{\lambda/(1-e^{-\lambda\eta\kappa})}$.

Miscoordination thus allows bubbles to persist. An interesting possibility is that the arrival of (even uninformative) news can potentially serve as a coordination device. If all arbitrageurs believe that other arbitrageurs will attack on the arrival of some news event, then it is a best response for them to do the same – coordinating on uninformative news can thus improve price discovery by potentially bursting the bubble earlier. Finally, Abreu & Brunnermeier (2002) extend the analysis to more general mispricing situations beyond bubbles that grow at an exponential rate.

3.4 Margin requirements and liquidity spirals

Arbitrageurs often rely on borrowing to fund their trading activities. When buying a risky security, the security can be used as collateral for borrowing, but because the security is risky, its collateral value is less than its current price. The security's *margin* for the trader is the difference between its price and its collateral value. When shorting a risky security, the trader must hold margin capital against the borrowed security so that the short position can be closed at a later date even if the price of the shorted security increases significantly. In both cases, the margin capital comes out of the trader's own capital. Less liquid markets feature larger liquidity risk, and thus greater margin requirements. However, larger margin requirements mean traders have less funding available to take positions in the market, worsening the liquidity of the market. Brunnermeier & Pedersen (2008) show this vicious cycle can create *liquidity spirals*, where prices move further and further away from fundamentals and so margins keep increasing, forcing traders to incur increasing losses.[8]

Fix a probability space $(\Omega, \mathcal{F}, P)$. There are J securities traded, all in zero net supply. There are four periods $t = 0, 1, 2, 3$, and the risk-free rate is 0. At time $t = 3$, each asset j has a random payoff X^j. Let $\mathbb{F} = \{\mathcal{F}_t\}_{t=0}^3$ be a filtration such that the fundamental value v_t^j evolves according to a autoregressive conditional heteroskedasticity process,

$$v_{t+1}^j = v_t^j + \sigma_{t+1}^j \epsilon_{t+1}^j,$$

where ϵ_t^j are iid shocks drawn from a standard normal distribution, and volatility σ_t^j observes

$$\sigma_{t+1}^j = \underline{\sigma}^j + \theta^j \left| v_t^j - v_{t-1}^j \right|$$

with $\underline{\sigma}^j, \theta^j \geq 0$. In the absence of any aggregate risk, the fundamental value of asset j at time $t < 3$ is $v_t^j = \mathbb{E}[X^j \mid \mathcal{F}_t]$. These securities have prices $S_t = (S_t^1, \ldots, S_t^J)$ at time t.

There are three kinds of agent. First, there are 3 risk-averse customers, who each have constant absolute risk aversion and maximize their utility $u(W_3^c) = e^{-\gamma W_3^c}$ over final period wealth W_3^c. At time $t = 0$, each customer $c \in \{0, 1, 2\}$ is endowed with cash W_0^c and no risky securities. However, each customer knows they will face an endowment shock $z^c = (z^{c,1}, \ldots, z^{c,J})$ in the securities holdings at $t = 3$, with $\sum_{c=0}^2 z^c = 0$. With

[8] Brunnermeier & Pedersen (2009) has turned out to be influential in policymaking – Geoffrey & Lee (2020) find it is the most cited paper in SEC rule proposals made since 2007.

probability $1-p$, all customers arrive simultaneously at the start of trading, and can trade in periods $t = 0, 1, 2$. With probability p, customers arrive sequentially, with customer c arriving at time $t = c$. In this case, denote the total demand shock for customers who have already arrived in period t by $Z_t = \sum_{c=0}^{t} z^c$. Customer c's demand is obviously $y_t^c = 0$ if the customer has not yet arrived at the start of period t, while if the customer has arrived, she chooses her period t position $y_t^c = (y_t^{c,1}, \ldots, y_t^{c,J})$ to maximize expected utility over final period wealth. Customer c's wealth evolves according to $W_{t+1}^c = W_t^c + (S_{t+1} - S_t)'(y_t^c + z^c)$.

Next, there are speculators, who act as market makers for the customers, and financiers, who provide funding to support the speculators' positions. Speculators are risk-neutral traders who seek to maximize their final period wealth. Speculators' position $x_t = (x_t^1, \ldots, x_t^j)$ is subject to a total margin requirement,

$$\sum_{j=1}^{J} x_t^{j+}(m_t^{j+} + x_t^{j-} m_t^{j-}) \leq W_t,$$

where $x_t^{j+} = x_t^j 1_{x_t^j \geq 0}$ and $x_t^{j-} = -x_t^j 1_{x_t^j \leq 0}$ are the positive and negative parts of speculators' position on security j, and $m_t^{j+}, m_t^{j-} \geq 0$ are the margin on long and short positions in security j. At the start of period $t = 0$, the speculators have an initial cash holding W_0 and no position in risky securities. Their wealth then evolves according to independent wealth shocks η_t (due to other non-trading activities) and the change in value of their position in the risky securities, so

$$W_{t+1} = W_t + (S_t - S_{t-1})' x_{t-1} + \eta_t.$$

If speculators' wealth is $W_t \leq 0$, then they have lost all their capital and are unable to borrow further because of the margin constraints, so $x_t = 0$. Speculators' utility is then $\varphi_t W_t$, where $\varphi_t \geq 0$. If there is costless limited liability, then $\varphi_t = 0$, while there are bankruptcy costs if $\varphi_t > 0$. The choice of φ_t does not change the qualitative results, so fix $\varphi_2 = 1$.

Financiers lend to the speculators, and face credit counterparty risk because the speculators might default. Let $\mathbb{G} = \{\mathcal{G}_t\}_{t=0}^{3}$ be the filtration describing the financiers' information. To limit exposure to counterparty risk, the financiers set the margins m_t^{j+} and m_t^{j-} to cover the π-value-at-risk of each speculators' position, i.e.

$$\pi = P(-(p_{t+1}^j - p_t^j) > m_t^{j+} \mid \mathcal{G}_t) = P(p_{t+1}^j - p_t^j > m_t^{j-} \mid \mathcal{G}_t),$$

for a target probability π (e.g. $\pi = 0.01$ means the speculator sets the margin so that they believe there is a 1% probability that an adverse price change is not covered by the margin, for both long and short positions.)

Brunnermeier & Pedersen link two kinds of liquidity in this model. The first is *market liquidity* – how easily an asset is traded – and the second is *funding liquidity* – how easily speculators can obtain funding. To keep track of market liquidity, let $\Lambda_t^j = p_t^j - v_t^j$ be the deviation in security j's price from its fundamental value. Market illiquidity is then measured by $\left|\Lambda_t^j\right|$. Speculators' funding illiquidity, on the other hand, is measured by the marginal value ϕ_t to speculators of additional funding.

In competitive equilibrium, prices S_t are such that speculators' position x_t maximizes expected profits subject to the margin constraint each period with the margins set as above, each customer c's position y_t^c maximizes her expected utility once she arrives in the market (and is 0 before arrival), and markets clear, i.e. $x_t + \sum_{c=0}^{2} y_t^c = 0$. The equilibrium can be solved by backward induction.

Let U_t denote a customer's value function at t and let V_t denote speculators' value function t. In period 2, customer c solves

$$U_2(W_2^c, S_2, v_2) = \max_{y_2^c} \left\{ -e^{-\gamma \mathbb{E}[W_3^c|\mathcal{F}_t] - \frac{\gamma^2}{2} \operatorname{var}(W_3^c|\mathcal{F}_t)} \right\},$$

which gives $y_2^{c,j} = \frac{v_2^j - S_2^j}{\gamma(\sigma_3^j)^2} - z^{c,j}$. Since all customers have arrived by period 2, $S_2 = v_2$ and thus speculator demand is 0, with customers trading their endowment of risky assets among themselves. We thus have $U_2(W_2^c, v_2, v_2) = -e^{-\gamma W_2^c}$ and $V_2(W_2, v_2, v_2) = W_2$. Now in periods $t \le 1$, if all customers arrived simultaneously in period 0 then the same argument gives $S_t = v_t$. Suppose then that customers arrive sequentially. In period 1, customers $c = 0, 1$ have demand $y_1^{c,j} = \frac{v_1^j - S_1^j}{\gamma(\sigma_2^j)^2} - z^{c,j}$, as in period 2, and $U_1(W_1^c, S_1, v_1) = -e^{-\gamma W_1^c - \frac{1}{2}\sum_j (v_1^j - S_1^j)^2/(\sigma_1^j)^2}$. If speculators take a position in security j, their profit per dollar invested is $\frac{v_1^j - S_1^j}{m_1^{j+}}$ on a long position and $-\frac{v_1^j - S_1^j}{m_1^{j-}}$ on a short position. The speculators take the profit-maximizing position across securities. Thus, provided they are solvent, speculators' shadow cost of capital is

$$\phi_1 = 1 + \max_j \left\{ \max \left\{ \frac{v_1^j - S_1^j}{m_1^{j+}}, -\frac{v_1^j - S_1^j}{m_1^{j-}} \right\} \right\}.$$

If speculators are insolvent (i.e. $W_1 < 0$) then $\phi_1 = \varphi_1$. Speculators' value function is $V_1(W_1, S_1, v_1) = \phi_1 W_1$ at $t = 1$.

In equilibrium, the market illiquidity and funding illiquidity at $t = 1$ are linked by the margin requirements on the positions taken by speculators. If speculators are long on security j (i.e. $x_1^j > 0$), then $\left|\Lambda_1^j\right| = m_1^{j+}(\phi_1 - 1)$, whereas if they are short ($x_1^j < 0$), then $\left|\Lambda_1^j\right| = m_1^{j-}(\phi_1 - 1)$. If $x_1^j = 0$, then $\left|\Lambda_1^j\right| \le \min\{m_1^{j+}, m_1^{j-}\}(\phi_1 - 1)$.

If the financiers who set the margin requirements are as informed as the speculators (i.e. $\mathcal{G}_t = \mathcal{F}_t$), then the financiers know the deviation in price from fundamentals Λ_t at t, and they know $\Lambda_2 = 0$. Thus lending to speculators is relatively safer the larger the deviation in price from fundamentals, so we would expect margins to be lower when mispricing is higher. Indeed, informed speculators set the margins so that

$$\pi = 1 - \Phi\left(\frac{m_1^{j+} - \Lambda_1^j}{\sigma_2^j}\right) = 1 - \Phi\left(\frac{m_1^{j-} + \Lambda_1^j}{\sigma_2^j}\right).$$

This implies $m_1^{j+} = \max\{\bar{\sigma}^j + \bar{\theta}^j \left| v_1^j - v_0^j \right| + \Lambda_1^j, 0\}$ and $m_1^{j-} = \max\{\bar{\sigma}^j + \bar{\theta}^j \left| v_1^j - v_0^j \right| - \Lambda_1^j, 0\}$, where $\bar{\sigma}^j = \underline{\sigma}^j \Phi^{-1}(1-\pi)$ and $\bar{\theta}^j = \theta^j \Phi^{-1}(1-\pi)$. The more market illiquidity there is, then, the more funding *liquidity.* Thus margin requirements move to stabilize markets.

If the financiers are uninformed, observing only prices ($\mathcal{G}_t = \sigma(\{S_u \mid u \le t\})$), then the opposite can be true. In particular, as the probability p that customers arrive sequential tends to 0, both long and short margin requirements tend to $m_1^j = \bar{\sigma}^j + \bar{\theta}^j \left| (v_1^j - v_0^j) + (\Lambda_1^j - \Lambda_0^j) \right|$. Margins here are increasing in price volatility. Moreover, when an illiquidity shock $\Lambda_1^j - \Lambda_0^j$ moves in the same direction as a fundamental shock $v_1^j - v_0^j$, it increases the margins. Intuitively, if uninformed financiers believe that prices reflect fundamentals with high probability, then they take price volatility as a sign of fundamental risk that they must protect themselves against. If the illiquidity shocks tend to comove with fundamentals (e.g. if the market overcorrects on bad news) then margin requirements can have a destabilizing effect.

Margin requirements can make the excess demand $x_t + \sum_c y_t^c$ non-monotonic in price, and thus the equilibrium price may not be continuous in the endowment shocks, an effect Brunnermeier & Pedersen label "fragility". For example suppose there is just one risky stock and speculators enter period 1 with a short position on that security. As the period 1 price of the stock increases, speculators take heavier losses. This may cause the margin requirements to bind – particularly if financiers are uninformed, so margins are destabilizing – which may force the speculators to close the short position early by purchasing the stock, driving up demand even as price increases.

Call an equilibrium with price S stable if in a neighbourhood of S excess demand is positive for $S' < S$ and negative for $S' > S$. Even small negative shocks to speculators at a stable equilibrium can be destabilizing because of two amplification channels. First, the "loss spiral" channel as outlined above, where speculator losses can force them to close their position at a loss to meet binding margin requirements. Second, if financiers are uninformed and place a sufficiently low prior on sequential arrival of customers, then the ARCH structure of fundamental values leads the financiers to conclude that adverse movements in price reflect greater fundamental volatility, pushing them to increase margins – this is the "margin spiral" channel. Both channels worsen speculators' funding position.[9]

Returning to the setting with J risky securities, if both the probability p of sequential arrivals and the sensitivities $\theta^j > 0$ of fundamental volatility for all securities j to fundamental shocks are all sufficiently low, then with either informed or uninformed financiers, the model exhibits some attractive features. Firstly, for risky assets i, j, we have $\operatorname{cov}(\left|\Lambda_1^i\right|, \left|\Lambda_1^j\right|) > 0$, so illiquidity tends to comove in the same direction across risky assets – thus market liquidity problems in one risky asset class can appear to "spill over" to other risky assets. Second, market liquidity is worse for assets with higher fundamental volatility, and this liquidity difference grows larger when speculators' margin constraints

[9]Intuitively, a stable equilibrium is one with a basin of attraction for a tâtonnement process.

are tighter – a "flight to safety" effect. Intuitively, this is because investing in less volatile assets places lower funding requirements on speculators, so increasing exposure to safer assets and reducing exposure to more volatile assets reduces speculators' funding problems. Escaping high margin requirements thus provides a rationale for flights to safety during liquidity crises.

The risk posed by funding liquidity constraints also affects speculator behaviour in period 0. At this stage, speculators maximize $\mathbb{E}[W_1\phi_1 \mid \mathcal{F}_0]$ and customer 0 maximizes $\mathbb{E}[U_1(W_1^0, S_1, v_1) \mid \mathcal{F}_0]$. If $\varphi_1 = 0$ so speculators have limited liability without bankruptcy costs, then the analysis is as for period 1. Assume instead that $\varphi_1 > 0$, and thus if $W_1 < 0$, then $W_1\phi_1 = W_1\varphi_1 < 0$. This would capture that the speculator might have other capital not pledged to trading (for example, banks engage in many other activities beyond trading that generate operational cash) and this capital will need to absorb trading losses when all capital pledged to trading is wiped out. If the speculator is not margin-constrained in period 0, we have first order condition $\mathbb{E}[\phi_1(S_1^j - S_0^j) \mid \mathcal{F}_0] = 0$. Expectations about funding illiquidity ϕ_1 in period 1 thus pin down the stochastic discount factor, since

$$S_0^j = \mathbb{E}\left[\frac{\phi_1}{\mathbb{E}[\phi_1 \mid \mathcal{F}_0]} S_1^j \mid \mathcal{F}_0\right] = \mathbb{E}[S_1^j \mid \mathcal{F}_0] + \frac{\operatorname{cov}(\phi_1, S_1^j \mid \mathcal{F}_0)}{\mathbb{E}[\phi_1 \mid \mathcal{F}_0]},$$

so the stochastic discount factor is $\frac{\phi_1}{\mathbb{E}[\phi_1 | \mathcal{F}_0]}$. The price of security j is depressed in period 0 when $\operatorname{cov}(\phi_1, S_1^j \mid \mathcal{F}_0) < 0$ so states of future funding illiquidity are associated with a lower price. An asset that has higher payoffs in liquidity crises would trade above its expected price in period 0, because there is an insurance motive to hold such assets.[10]

3.5 Slow moving capital

Frictions in the mobility of capital into or out of segmented securities markets can lead to persistent mispricings. There is good evidence of slow movement of capital in a range of markets:

- Mitchell, Pedersen & Pulvino (2007) catalogue slow moving capital effects in the convertible bond market. Circa 2005, many hedge funds were forced to sell bonds in order to redeem investors who were demanding redemptions. This depressed the price of convertible bonds below fundamental values. Nevertheless, roughly half of the multistrategy hedge funds who *were not* capital constrained were net sellers. They interpret this as evidence that within-firm information barriers led to internal capital constraints for the trading desks at these hedge funds. They similarly find hedge funds were slow to enter the depressed convertible bond market following the liquidation of LTCM in 1998.

- Duffie (2010) gives other examples. The most extreme is in the market for catastrophe risk insurance. Premiums in this market are very volatile, and can vary by

[10] For example, US Treasury securities and gold are assets typically associated with a flight to safety.

up to 50% over multiyear periods (Enz, 2001), despite the probability of, say, a hurricane, being roughly constant over time. Capital replenishment at insurance and reinsurance firms can take years.

- Fleckenstein, Longstaff & Lustig (2014) interpret the TIPS-Treasury bond puzzle as evidence of slow moving capital.

3.5.1 Duffie & Strulovici (2012): intermediation frictions

Duffie & Strulovici (2012) model capital movement between a pair of partially segmented markets – the catastrophe insurance market is a motivating example. Fix a probability space $(\Omega, \mathcal{F}, P)$ and a filtration $\mathbb{F} = \{\mathcal{F}_t \mid t \in [0, \infty)\}$ satisfying the usual conditions. There are two markets, a and b. There are three kinds of agent. First, each market has a continuum of "local hedgers", who are risk averse agents who own short-lived risky assets. The local hedgers cannot trade across markets. Second, there is a continuum of infinitely-lived, risk-neutral investors who are willing to supply capital by purchasing the risky assets from the local hedgers whenever the risk premium is positive. The continuum of investors is a non-atomic measure space $(A, \mathcal{A}, \gamma)$ and we assume $(\Omega, \mathcal{F}, P)$ and $(A, \mathcal{A}, \gamma)$ have a Fubini extension $(A \times \Omega, \mathcal{W}, \xi)$.[11] Once an investor has supplied capital to one sector, she can only move the capital to the other sector via an intermediary. Investors maximize the discounted present value of their wealth net of intermediation fees. Intermediaries are the third type of agent – they charge investors a fee to move capital between the two markets.

Let X_{it} be the level of capital in market $i = a, b$ at time t. The investors in each market can roll over their investments in the short-run risky assets continuously, receiving an equilibrium dividend rate $\pi(X_{it}) = k_0 + \frac{k}{X_{it}}$.[12] This is paid in cash and is not reinvested automatically. The short-lived assets are risky because there are loss events that arrive according to a Poisson process. All of the capital invested in market i's risky assets is lost when a loss event occurs. Loss events arrive at mean arrival rate η.[13] As long as $\pi(X_{it}) - \eta > r$, investors in market i are willing to supply all their capital to the local hedgers, since the investors' mean rate of return exceeds their rate of time preference. Since a unit of capital in both markets faces identical risk of losses, the difference in mean rates of return across markets is just the difference in dividend rates.

Absent capital movement frictions (i.e. if intermediaries charged zero fees), capital would move to the market with the higher return, so in equilibrium, we would have $\pi(X_{at}) = \pi(X_{bt})$, implying $X_{at} = X_{bt}$. Capital frictions can result in unequal capital levels and thus unequal dividend rates.

To be able to move capital between the two markets, an investor needs to be in contact with an intermediary. Investors are contacted with intensity $\lambda_t \in [0, \bar{\lambda}]$ where $\bar{\lambda}$ is the

[11] This is necessary to apply the exact law of large numbers. See Sun (2006).

[12] This is determined through a double auction, and will take this form in equilibrium under appropriate assumptions about how local hedgers' risk aversion is distributed.

[13] They extend this to a much more general setting where there a random proportion of capital is lost. The total capital loss assumption is a simplification.

maximal feasible intensity, and are contacted pairwise independently conditional on the realized path of the intensity process λ. The intermediary incurs a flow cost $c\lambda_t$ with $c \geq 0$. The intermediation fee they charge is a fraction q of the investor's gain in present value from redeploying capital from one market to the other (q is the bargaining power of the intermediary). The intermediary's profit maximization problem is to choose the intensity process λ to maximize the expected present value of intermediation fee revenues net of contacting costs. Since the mean rate of return is higher in the market with less capital, the only capital movements in will be from the market with more capital to the market with less.

Suppose intermediation is monopolistic, and restrict consideration to equilibria where contacting intensity takes the form $\lambda_t = \Lambda(X_t, Y_t)$ for $X_t = \max\{X_{at}, X_{bt}\}$ and $Y_t = \min\{X_{at}, X_{bt}\}$. Let $W_{ij}(t)$ be the capital invested in market i by investor j at t. By the exact law of large numbers, the total flow rate of capital from market a to market b is

$$\int_A \lambda_t 1_{X_{at} > X_{bt}} W_{aj}(t)\,\mathrm{d}\gamma(j) = \lambda_t 1_{X_{at} > X_{bt}} \int_A W_{aj}(t)\,\mathrm{d}\gamma(j) = \lambda_t 1_{X_{at} > X_{bt}} X_{at},$$

almost surely, and the total flow rate of capital from market b to a is $\lambda_t 1_{X_{bt} > X_{at}} X_{bt}$ a.s.

Given (X_t, Y_t) and fixing intermediation policy Λ, we can represent the continuation value to an investor of holding a unit of capital in the overcapitalized market by a function $G^\Lambda : \mathbb{R}_+^2 \to [0, \infty)$ and the present value to an investor of a unit of capital in the undercapitalized market by $H^\Lambda : \mathbb{R}_+^2 \to [0, \infty)$. If these are differentiable, then Itô's lemma implies they satisfy coupled equations,

$$\begin{aligned} rG^\Lambda(x,y) = \pi(x) + [G_y^\Lambda(x,y) - G_x^\Lambda(x,y)]x\Lambda(x,y) + (1-q)\Lambda(x,y)[H^\Lambda(x,y) - G^\Lambda(x,y)] \\ - \eta G^\Lambda(x,y) + \eta(G^\Lambda(x,0) - G^\Lambda(x,y)), \\ rH^\Lambda(x,y) = \pi(\underline{x}) + [H_y^\Lambda(x,y) - H_x^\Lambda(x,y)]x\Lambda(x,y) + \eta[G^\Lambda(y,0) - H^\Lambda(x,y)] - \eta H^\Lambda(x,y). \end{aligned}$$

These can be interpreted as the loss in value due to time preference equals the expected net profits from investment in each market. Now supposing investors conjecture the intermediary's intermediation policy is Γ (and so the intermediary's fees are consistent with Γ), the intermediary's value function V given maximal capital level x and minimal capital level y across the two markets is given by

$$V(x,y) = \sup_\Lambda \mathbb{E}\left[\int_0^\infty e^{-rt}\Lambda(X_t, Y_t)\left(X_t q\left(H^\Gamma(X_t,Y_t) - G^\Gamma(X_t,Y_t)\right)\right)\mathrm{d}t\right].$$

Intermediation policy Λ is an equilibrium policy if it solves this problem when investors correctly conjecture the policy, $\Gamma = \Lambda$. Suppose that V is finite and differentiable. The value function solves the Hamilton-Jacobi-Bellman equation,

$$0 = \sup_{\hat\lambda \in [0,\bar\lambda]} \left[-rV(x,y) + \mathcal{D}V(x,y,\hat\lambda)\right]$$

where

$$\mathcal{D}V(x,y,\hat{\lambda}) = -V_x(x,y)\hat{\lambda}x + V_y(x,y)\hat{\lambda}x + \eta[V(y,0)+V(x,0)-2V(x,y)] \\ + \hat{\lambda}\left[xq\left(H^{\Gamma}(x,y)-G^{\Gamma}(x,y)\right)-c\right].$$

By a martingale verification argument, one can establish that for any given intermediation policy Γ, any bounded differentiable function $\hat{V}$ satisfying the HJB equation is indeed the value function. Since the contact intensity λ enters linearly in the objective function in the HJB equation, the optimal contact intensity for the intermediary at any instant is either $\bar{\lambda}$ or 0. The intermediary optimally contacts with intensity $\bar{\lambda}$ at time t if

$$q\left[H^{\Lambda}(X_t,Y_t)-G^{\Lambda}(X_t,Y_t)\right]-\left[V_x(X_t,Y_t)-V_y(X_y,Y_t)\right] > \frac{c}{X_t}.$$

For investors in market i, if there is no opportunity to move capital to the other market then it is optimal to reinvest capital in market i as long as the mean rate of return exceeds the discount rate, i.e. $\pi(x)-\eta-r = k_0-\frac{k}{x}-\eta-r \geq 0$. This implies $k_0 \geq \eta+r$. However, since k_0 is common to the rate of return in both markets, it can for the sake of analysing equilibrium cross-market capital movements be taken as zero without loss of generality, which makes π homogeneous of degree -1.

Suppose in equilibrium, the contact intensity is a function of the cross-market capital ratio $Z_t = X_t/Y_t$. Since loss events destroy all capital in one of the markets, this ratio becomes infinite when a loss event occurs. By homogeneity of dividend rate π and intermediation policy Λ, H^{Λ} and G^{Λ} are homogeneous of degree -1, implying $G^{\Lambda}(z,0) = g_0kz^{-1}$ for some constant $g_0>0$. Put $f(z)=[H^{\Lambda}(z,1)-G^{\Lambda}(z,1)]/k$ and $L(z)=\Lambda(z,1)$. The coupled equations for G^{Λ} and H^{Λ} then imply f solves ODE,

$$0 = -rf(z)+1-z^{-1}-zL(z)f'(z)-(f(z)+zf'(z))L(z)z-(1-q)f(z)L(z) \\ +\eta[g_0(1-z^{-1})-2f(z)],$$

with boundary condition $f(1)=0$. For any intermediation policy Λ, $f(z)>0$ if $z>1$, so any investor in the overcapitalized market optimally shifts all their capital to the undercapitalized market when contacted by the intermediary. Given G^{Λ} and H^{Λ} (and thus given f), the intermediary's value function V is homogeneous of degree 0, and thus so is the intermediation policy, verifying $\Lambda(x,y)=L(x/y)$ for some function L. Since f depends on L and L depends on f, finding the equilibrium is a fixed point problem. As noted before, equilibrium $\Lambda(x,y)$ takes only values $\bar{\lambda}$ or 0 – one can show there is some threshold ratio $T\geq 1$ in equilibrium so that $\Lambda(x,y)=1_{x/y\geq T}\bar{\lambda}$. Finding this threshold T is all that remains to characterize the equilibrium. Given the form of Λ in equilibrium, the ODE for f gives us $f(z)=\frac{1+\eta g_0}{r+2\eta}(1-1/z)$ for $z\in[1,T]$. For $z\geq T$, the ODE becomes

$$(r+2\eta+\bar{\lambda}((1-q)+z))f(z)+\bar{\lambda}(1+z)zf'(z)=(1+\eta g_0)(1-1/z).$$

Putting $v(z):=V(z,1)/k$, we can write the HJB equation as

$$0=\sup_{\hat{\lambda}\in[0,\bar{\lambda}]}\left[-rv(z)-\hat{\lambda}(z+z^2)v'(z)+2\eta(v_0-v(z))+(qzf(z)-c/k)\hat{\lambda}\right],$$

where $v_0 := V(y, 0)/k = V(x, 0)/k$. We get $v(z) = \frac{2\eta}{r+2\eta} v_0$ for $z \in [1, T]$, and

$$\frac{r + 2\eta}{\bar{\lambda}} v(z) + v'(z) z(1 + z) = \frac{2\eta v_0}{\bar{\lambda}} - \frac{c}{k} + qzf(z) \quad \text{for } z \geq T.$$

It turns out that even though f is not generally monotone, the aggregate trade surplus from intermediation, $zf(z)$, is strictly increasing and thus the normalized homogenized value function $v(z)$ is strictly increasing. Moreover, v is bounded. Now, from the smooth pasting condition $v'(T) = 0$, we have $qTf(T) = \frac{c}{k}$, giving $T = 1 + \frac{c(r+2\eta)}{(1+\eta g_0)qk}$. Finally, the aggregate of the investors' value functions plus the intermediary's value function must equal the present value of the cash dividend payments net of the intermediary's search costs. This can be used to pin down $g_0 = \frac{2}{kr} - \frac{c\bar{\lambda}}{kr}(1 - e^{-(2\eta+r)\ln(1+1/T)/\bar{\lambda}}) - v_0$. This threshold strategy equilibrium exists and is the unique equilibrium.

In equilibrium, if there are no intermediation costs then $T = 1$, so the intermediary always contacts investors at the maximal intensity. As these costs increase, the threshold T increases. Increasing the dividend rate coefficient k decreases the threshold, because it makes the dividend in the two markets more sensitive to capital levels. Increasing the discount rate r directly reduces investors' expected present value gain from moving to the undercapitalized market, and indirectly reduces the fees investors pay due to this lower gain – the first effect dominates the second, so increasing r raises the threshold T. Holding the dividend rate function π fixed, increasing the intensity of loss events η increases the threshold T, although raising η would also change the equilibrium π (this leads to an ambiguous effect on T). Likewise, while raising the intermediary's bargaining power q increases the intermediary's incentives to search holding fixed investors' present value from moving capital, raising q also lowers investors' present value from moving capital in the future and thus the gains to be shared with the intermediary. If $q < \frac{1}{2}$, the former effect dominates the latter.

Duffie & Strulovici (2012) extend the analysis to oligopolistic and perfectly competitive intermediaries. Competition provides investors with outside options so limits the bargaining power of intermediaries, reducing the profitability of intermediation. When loss events destroy all capital, this is the only effect of competition, and so competition raises the threshold T and reduces capital mobility. When loss events only partially destroy capital, there is another channel – each intermediary does not internalize the effect of its intermediation activities on reducing the profits of other intermediaries. By a standard Cournot-model-type intuition, this force acts to increase intermediation as we increase the number of intermediaries, if we hold bargaining power q fixed. The channel is shut down when loss events destroy all capital because then pre-shock heterogeneity in capital levels and post-shock heterogeneity are completely decoupled.

3.5.2 Dow, Han & Sangiorgi (2021): liquidity hysteresis

Dow, Han & Sangiorgi (2021) consider a model where arbitrageur's capital can flow *out* of a market experiencing a liquidity shock, exacerbating mispricing. They term this flight-to-liquidity phenomenon "liquidity hysteresis". This provides a more worked-out story for the hedge fund behaviour documented by Mitchell, Pedersen & Pulvino (2007).

There is a continuum of infinitely-lived agents in a discrete time setting. Agents are risk-neutral, and have discount factor β. There are two markets for risky assets – a market in long-term securities, and a market in short-term securities. Both of these markets consist of a continuum of securities. A short-term security s that is available at time t will be liquidated at time $t+1$, and has a random liquidation value $v_s \in \{V_s^L, V_s^H\}$ where $V_s^L < V_s^H$ and the value is drawn uniformly at random. A long-term security ℓ has a random maturity, maturing with probability $q > 0$ each period, and also has a random liquidation value $v_\ell \in \{V_\ell^L, V_\ell^H\}$, where $V_\ell^L < V_\ell^H$ and the value is drawn uniformly at random. That is, in both markets, there are high and low quality assets, and quality is not known ex ante to agents. Liquidation values are iid across securities and across time, and satisfy

$$\frac{\beta q}{1-\beta(1-q)} V_\ell^L = \beta V_s^L \quad \text{and} \quad \frac{\beta q}{1-\beta(1-q)} V_\ell^H = \beta V_s^H.$$

This implies the present value of the securities in the two markets are identical. Finally, there is riskless short-term borrowing, with a short rate $r = \beta^{-1} - 1$. The price of an asset in equilibrium may fully reveal the quality of the asset. The mass of "unrevealed" assets is always normalized to 1 in each market, for simplicity.

There are three types of agent. First, there is a continuum (non-atomic measure space) $(A, \mathcal{A}, \gamma)$ of arbitrageurs, with the normalization $\gamma(A) = 1$. Each arbitrageur can costlessly generate a private signal that perfectly reveals the liquidation value of one security each period. However, arbitrageurs are capital-constrained, so can each only hold a unit position (long or short) in a single risky asset at any point in time. Arbitrageur a submits an order $x_i^a(t) \in \{-1, 0, 1\}$ for security i at time t (noting the capital constraints imply $x_i^a(t) = 0$ for all but one asset i). An arbitrageur's position is open unless either (i) the asset matures, (ii) the price fully reveals the asset's quality, or (iii) the arbitrageur decides to close out the position early, before realizing liquidation profits.

Second, there is a continuum of noise traders. The noise traders submit a random order flow ζ_{it} for each asset i. For each market h, ζ_{it} is an iid draw from a uniform distribution over an interval $[-z_{ht}, z_{ht}]$. For the short-term asset, the endpoint $z_{st} := z_s$ is fixed. However, for the long-term asset, $z_{\ell t}$ varies over time according to an N-state Markov chain. The realization of $z_{\ell t}$ is commonly known to all agents each period. Finally, there are competitive market makers. The other traders submit orders to the market makers, who set trading prices to clear the market.

Each period t, a fraction ξ_t of arbitrageurs are *active* and the remaining fraction $1-\xi_t$ of arbitrageurs are inactive (i.e. their capital is currently "locked-in" in existing trade positions). Active arbitrage capital is available for an arbitrageur to invest, whereas inactive arbitrage capital is capital that is currently committed to some investment position in an unrevealed asset. Of the active arbitrageurs, a fraction δ_t choose to invest in the long-term market, with the remainder investing in short-term securities. Once choosing a market, an active arbitrageur collects information about a single security in that market. Of the inactive arbitrageurs, a fraction η_t choose to liquidate their investment early (i.e. before maturity and before the price fully reveals the liquidation payoff).

As in a Kyle model, both the noise traders and the informed arbitrageurs submit orders to risk neutral market makers. Because the market makers are competitive and risk neutral, they set the price S_{it} of asset i to break-even in expectation given the observed order flows, i.e.

$$S_{it} = \mathbb{E}\left[\beta^{\tau_i} v_i \mid \xi_t, z_{\ell t}, X_{it}\right],$$

where $X_{it} = X_{it}^a + \zeta_{it}$ is the aggregate order flow for asset i at time t, $X_{it}^a = \int_A x_i^a(t)\,\mathrm{d}\gamma$ is the aggregate order from by arbitrageurs, ζ_{it} is noise traders' order flow and τ_i is the liquidation stopping time for asset i. Now, if a mass μ_{it} of arbitrageurs invest in asset i at time t in market h, they either all buy the asset ($X_{it}^a = \mu_{it}$) if the liquidation value is high (V_h^H) or all short the asset ($X_{it} = -\mu_{it}$) if the liquidation value is low (V_h^L). Since ζ_{it} is a uniform distribution on $[-z_{it}, z_{it}]$, X_{it} is uniform on $[-z_{it} \pm \mu_{it}, z_{it} \pm \mu_{it}]$. Bayes' rule then implies that if the market makers' prior belief that $v_i = V_h^H$ is $p = \frac{1}{2}$ then their posterior belief $\hat{p}_{it}$ is

$$\hat{p}_{it} = \begin{cases} 0 & \text{if } -z_{it} - \mu_{it} \leq X_{it} < -z_{it} + \mu_{it}, \\ p & \text{if } -z_{it} + \mu_{it} \leq X_{it} \leq z_{it} - \mu_{it}, \\ 1 & \text{if } z_{it} - \mu_{it} < X_{it} \leq z_{it} + \mu_{it}. \end{cases}$$

Thus prices are either fully revealing of asset i's liquidation value ($S_i = S^H := \beta V_s^H$ or $S_i = S^L := \beta V_s^L$) or completely non-revealing ($S_i = S^0 := \beta(V_s^L + V_s^H)/2$) [recall present value for the two asset classes is equal]. The probability asset i's price is fully revealing at t is $\lambda_{it} = P(\{\hat{p}_i \in \{0,1\}\}) = \frac{\mu_{it}}{z_{it}}$. This probability will turn out to reflect both price effiency and liquidity.

Consider interior, stationary, symmetric rational expectations equilibria. Then within each market h, $\lambda_i = \lambda^h$ for all unrevealed assets i. The process giving the mass of active arbitrageur capital ξ_t evolves according to

$$\xi_{t+1} = (1-\delta_t)\xi_t + (\delta_t \xi_t + 1 - \xi_t)(q + (1-q)\lambda_{\ell t}) + \eta_t(1-\xi_t)(1-q)(1-\lambda_{\ell t}),$$

recalling δ_t is the fraction of active arbitrageurs trading in market L, q is the probability that a given long term security matures each period, and η_t is the mass of inactive arbitrageurs liquidating early.

One can show inactive arbitrageurs will never close a position early in equilibrium, so $\eta_t = 0$. Given the state $(\xi_t, z_{\ell t})$, the value function for an inactive arbitrageur is then $J_I(\xi_t, z_{\ell t}) = \lambda_{\ell t} S^H + (1-\lambda_{\ell t})S^0 + \mathbb{E}[J_\ell(\xi_{t+1}, z_{\ell t+1}) \mid \xi_t, z_{\ell t}]$, and the value function for an active arbitrageur is $J_f(\xi_t, z_{\ell t}) = \max\{J_s(\xi_t, z_{\ell t}), J_\ell(\xi_t, z_{\ell t})\}$, where

$$\begin{aligned} J_s(\xi_t, z_{\ell t}) &= -(\lambda_{st} S^H + (1-\lambda_{st})S^0) + \beta\left(V_s^H + \mathbb{E}[J_f(\xi_{t+1}, z_{\ell t+1}) \mid \xi_t, z_{\ell t}]\right), \\ J_\ell(\xi_t, z_{\ell t}) &= -(\lambda_{\ell t} S^H + (1-\lambda_{\ell t})S^0) + \beta\big[q V_\ell^H + (1-q)\lambda_{\ell t} S^H \\ &\qquad + (1-(1-\lambda_{\ell t})(1-q))\mathbb{E}[J_f(\xi_{t+1}, z_{\ell t+1}) \mid \xi_t, z_{\ell t}] \\ &\qquad + (1-\lambda_{\ell t})(1-q)\mathbb{E}[J_I(\xi_{t+1}, z_{\ell t+1}) \mid \xi_t, z_{\ell t}]\big]. \end{aligned}$$

In equilibrium, the active arbitrageur capital flows to the market with the higher value, so $\delta_t = 0$ if $J_\ell(\xi_t, z_{\ell t}) < J_s(\xi_t, z_{\ell t})$, $\delta_t = 1$ if $J_\ell(\xi_t, z_{\ell_t}) > J_s(\xi_t, z_{\ell t})$, and $\delta_t \in [0,1]$ otherwise. Together, $(J_f, J_I, J_\ell, J_s, \delta, \xi, \lambda_\ell, \lambda_s)$ characterizes the stationary symmetric equilibrium. In this equilibrium, price efficiency satisfies $\lambda_{\ell t} = \frac{\delta_t \xi_t}{z_{\ell t}}$ and $\lambda_{st} = \frac{(1-\delta_t)\xi_t}{z_{st}}$. Fixing the relative allocation δ_t of active arbitrage capital across the two markets, the price efficiency in market h improves when there is more active capital ξ_t available and worsens as the intensity of noise trading, z_{ht} becomes more extreme. Under plausible but complicated technical conditions, there is a unique stationary equilibrium in which the price efficiency in the long term asset market is increasing in ξ_t and decreasing in $z_{\ell t}$ (no longer fixing δ).

The interesting feature of this model is that there can be many liquidity "regimes" in equilibrium. A temporary liquidity shock can lead to a switch in regime, so have persistent effects. Fix a level of $z_{\ell t}$ at $\bar{z}_\ell$ and suppose $z_{\ell t}$ stays constant at this level over time (similarly, denote the constant level of z_{st} by $\bar{z}_s$). We can then focus on a steady state equilibrium. Steady state variables x are denoted as x^*, so ξ^* is the steady state fraction of active arbitrageurs, for example. One can show price efficiency in equilibrium across the two markets satisfies an indifference condition $\lambda_{st} - \lambda_{\ell t} = \beta(1-q)(1-\lambda_{\ell t})(1 - \mathbb{E}[\lambda_{\ell,t+1} \mid \xi_t, z_{\ell t}])$, which gives us $\lambda_s^* - \lambda_\ell^* = \beta(1-q)(1-\lambda_\ell^*)^2$ in steady state. Substituting for $\lambda_\ell^* = \frac{\delta^*\xi^*}{\bar{z}_\ell}$ and $\lambda_s^* = \frac{(1-\delta^*)\xi^*}{\bar{z}_s}$ gives

$$\frac{\bar{z}_s - (1-\delta^*)\xi^*}{\bar{z}_s} = \left(\frac{\bar{z}_\ell - \delta^*\xi^*}{\bar{z}_\ell}\right)\left(1 - \beta(1-q)\frac{\bar{z}_\ell - \delta^*\xi^*}{\bar{z}_\ell}\right).$$

Decreases in the fraction of arbitrageurs increase speculative profits in the short-term market s relatively more than they increase speculative profits in the long-term market ℓ, so the fraction of active arbitrageurs in the long-term market δ^* must fall to restore arbitrageur indifference between the two markets.

Next, substituting $\lambda_\ell^* = \frac{\delta^*\xi^*}{\bar{z}_\ell}$ into the law of motion for ξ_t gives

$$\xi^* = (1-\delta^*)\xi^* + (\delta^*\xi^* + 1 - \xi^*)\left(q + (1-q)\frac{\delta^*\xi^*}{\bar{z}_\ell}\right).$$

As the fraction of active arbitrageurs in the long-term market δ^* increases, more of the arbitrageurs remain locked into illiquid investments, putting downward pressure on the proportion of active arbitrageurs ξ^*. Yet at the same time, higher active arbitrageur participation in the long-term market improves price efficiency λ_ℓ^*, which makes the market less illiquid and thus freeing up arbitrage capital, an upward force on ξ^*. Which effect dominates depends on the parametrization. A steady state always exists, but there can be up to two stable steady states for some parameter values. Indeed, for some thresholds $0 < \underline{q} < \bar{q} < 1$ and $0 < \underline{\beta} < \bar{\beta} < 1$, there is a unique steady state if $q > \bar{q}$ and $\beta < \underline{\beta}$, and multiple steady states if $q < \underline{q}$, $\beta > \bar{\beta}$, and $\bar{z}_\ell < 1 - \frac{3}{4}\bar{z}_s$.

Returning to the case where $z_{\ell t}$ is subject to shocks, suppose now that $z_{\ell t}$ has a "normal level" $\bar{z}_\ell$ but with small probability undergoes transient shocks where $z_{\ell t}$ deviates from $\bar{z}_\ell$. Suppose we have two steady states at the normal level, and that we are initially at the steady state with the higher level of active capital ξ_1^*. Recall we had indifference condition $1-\lambda_{st} = (1-\lambda_{\ell t})(1-\beta(1-q)(1-\mathbb{E}[\lambda_{\ell,t+1} \mid \xi_t, z_{\ell t}]))$. A temporary shock that increases $z_{\ell t}$ from $\bar{z}_\ell$ would, if the fraction of active arbitrageurs in the long-term market remains fixed, reduce price efficiency $\lambda_{\ell t}$ in the long term market. As above, we have two effects in opposite directions: reduced price efficiency increases speculative profits but increases the expected capital lock-in time and thus opportunity cost. If the opportunity cost dominates, then the proportion δ_t of active arbitrageurs in the long-term market falls to maintain active arbitrageur indifference between the two markets. This worsens the price efficiency of the long-term market. Arbitrageurs already locked-in in that market thus expect to be locked-in for longer, and so ξ_{t+1} decreases. If the temporary shock lasts sufficiently many periods, this spiral can lead to the process ξ crossing a threshold so that the economy converges to the steady state with the lower active capital level $\xi_2^* < \xi_1^*$. Thus wa short term liquidity shocks can result in a switch to a lower liquidity regime.

4 Heterogeneous beliefs

A strand of the literature going back to Harrison & Kreps (1978) relaxes the common priors assumption. Per the no trade theorem (Aumann, 1976; Kreps, 1977; Milgrom & Stokey, 1982), if we have a market of strictly risk averse traders who share a common prior and it is common knowledge that (i) traders are rational Bayesian agents who share a common prior, and (ii) that the initial allocation is ex-ante efficient, then in any equilibrium, there will be no trade. Naturally, this rules out bubbles (as discussed earlier for Tirole (1982)). If traders have heterogeneous beliefs, then a trader might conclude their counterparty in a trade simply disagrees with them without having superior information, so trade becomes possible. There is a (somewhat tedious) debate about whether heterogeneous priors is a reasonable assumption. Regardless of one's stance in that debate, models of bubbles with heterogeneous beliefs and short-selling restrictions can fit some stylized facts quite well.

Recall that bubbles are often accompanied by abnormally high trading volumes and high trading volumes. Scheinkman & Xiong (2003) develop a model of bubbles that has these features by taking Harrison & Kreps' (1978) model to a continuous time setting.

Fix a probability space $(\Omega, \mathcal{F}, P)$ and let $\mathbb{F} = \{\mathcal{F}_t \mid t \in [0, \infty)\}$ be a filtration satisfying the usual conditions. There is a single risky security that pays dividends and is in fixed, finite supply. The cumulative dividend process D has drift given by a mean-reverting fundamental variable μ_t. The processes D and μ evolve according to

$$\begin{aligned} \mathrm{d}D_t &= \mu_t \,\mathrm{d}t + \sigma \,\mathrm{d}B_t^1, \\ \mathrm{d}\mu_t &= -\lambda(\mu_t - \bar{\mu}) \,\mathrm{d}t + v \,\mathrm{d}B_t^2, \end{aligned}$$

where $\bar{\mu}$ is the mean of μ, $\lambda \geq 0$ is a mean-reversion parameter, and $\sigma, v > 0$ are volatilities. As well as the dividend process, there are two signals s^1 and s^2 that evolve according to

$$\mathrm{d}s_t^a = \mu_t \,\mathrm{d}t + \hat{\sigma}\,\mathrm{d}B_t^a,$$
$$\mathrm{d}s_t^b = \mu_t \,\mathrm{d}t + \hat{\sigma}\,\mathrm{d}B_t^b,$$

where $\hat{\sigma} > 0$ is a common volatility. Assume B^1, B^2, B^a, B^b are mutually independent standard Brownian motions.

There is a continuum of risk neutral, infinitely-lived agents. Each agent has a type $i \in \{a, b\}$, and the mass of both types is equal. All agents observe both signals, but a type i agent believes signal s^i is more informative than it truly is. This is because type i falsely believes that

$$\mathrm{d}s_t^i = \mu_t \,\mathrm{d}t + \hat{\sigma}\varphi\,\mathrm{d}B_t^1 + \hat{\sigma}\sqrt{1-\varphi^2}\,\mathrm{d}B_t^i,$$

i.e. they falsely believe B^1 and B^i are correlated. This leads them to overreact to signal s^i. Type i has correct beliefs about the other signal s^{-i}.

Given their faulty beliefs about the structure of the problem, agents face an optimal nonlinear filtering problem in updating their beliefs about μ_t given information $\mathcal{G}_t = \sigma(\{D_u, s_u^a, s_u^b \mid u \le t\} \cup \mathcal{N})$.[14] Since all the variables are Gaussian, if the initial conditions D_0, s_0^a, s_0^b are Gaussian (or constant), then type i's beliefs about μ will be a Gaussian process, with conditional mean $\hat{\mu}^i$ and variance γ^i. For the stationary solution, the variance for both types' beliefs is the same, given by

$$\gamma = \frac{\sqrt{(\lambda + \varphi v/\hat{\sigma})^2 + (1-\varphi^2)(2v^2/\hat{\sigma}^2 + v^2/\sigma^2)} - (\lambda + \varphi v/\hat{\sigma})}{1/\sigma^2 + 2/\hat{\sigma}^2}.$$

This is decreasing in φ, so φ can be thought of as capturing overconfidence. The conditional mean $\hat{\mu}^i$ evolves according to

$$\mathrm{d}\hat{\mu}_t^i = -\lambda(\hat{\mu}_t^i - \bar{\mu})\,\mathrm{d}t + \frac{\varphi v\hat{\sigma} + \gamma}{\hat{\sigma}}\,\mathrm{d}\bar{B}_t^{i,i} + \frac{\gamma}{\hat{\sigma}}\bar{B}_t^{-i,i} + \frac{\gamma}{\sigma}\bar{B}_t^{1,i},$$

where (from the perspective of type i only) $\bar{B}^{i,i}$, $\bar{B}^{-i,i}$ and $\bar{B}^{1,i}$ are mutually independent standard Brownian motions, defined by

$$\mathrm{d}\bar{B}_t^{i,i} = \frac{1}{\hat{\sigma}}\,\mathrm{d}s_t^i - \frac{\hat{\mu}_t^i}{\hat{\sigma}}\,\mathrm{d}t,$$
$$\mathrm{d}\bar{B}_t^{-i,i} = \frac{1}{\hat{\sigma}}\,\mathrm{d}s_t^{-i} - \frac{\hat{\mu}_t^i}{\hat{\sigma}}\,\mathrm{d}t,$$
$$\mathrm{d}\bar{B}_t^{1,i} = \frac{1}{\sigma}\,\mathrm{d}D_t - \frac{\hat{\mu}_t^i}{\sigma}\,\mathrm{d}t.$$

The conditional mean process $\hat{\mu}^i$ of type i's beliefs is mean-reverting. While $\bar{B}^i = (\bar{B}^{i,i}, \bar{B}^{-i,i}, \bar{B}^{1,i})$ is a standard Brownian motion from the perspective of type i, it is not a standard Brownian motion from the perspective of the other type $-i$, because $-i$ has a different (but also incorrect) understanding of the structure of the model. Given

[14] Here, $\mathcal{N}$ denotes the set of all null sets.

constant variance in beliefs for both types, label the conditional mean process $\hat{\mu}^i$ as i's beliefs. The difference in i's beliefs relative to $-i$ is given by $\Delta^i_t = \hat{\mu}^{-i}_t - \hat{\mu}^i_t$. This "difference in beliefs" process evolves according to

$$\mathrm{d}\Delta^i_t = -\rho\Delta^i_t\,\mathrm{d}t + \sqrt{2}\varphi v\,\mathrm{d}W^i_t,$$

where W^i is a standard Brownian motion for type i that is independent of $\bar{B}^i$, and $\rho = \sqrt{(\lambda + \varphi v/\hat{\sigma})^2 + (1-\varphi^2)(2/\hat{\sigma}^2 + 1/\sigma^2)}$. This is a mean-reverting process. Moreover, if there is no overconfidence (i.e. $\varphi = 0$), then there is no volatility in the difference in beliefs between the types. With overconfidence, the difference in beliefs fluctuates. This generates trading activity.

Trading takes place in a highly organized market (i.e. an exchange), where short-selling is forbidden.[15] When selling the security, a trader incurs a transaction cost $c \geq 0$ per unit of security sold. At time t, suppose type i does not own the asset and type $j \neq i$ currently owns the asset. Type i's willingness to pay for the security is S^i_t. Because we have a continuum of agents and short-sale constraints, this is the price a type i agent will pay if they acquire the security at time t. The current owners, type j, choose a stopping time τ at which to sell to type i. They choose the stopping time to maximize their own expected profits. Let the expectation operator $\mathbb{E}^j_t[\cdot]$ denote type j's expectation conditional on information $\mathcal{G}_t$ and given type j's (faulty) structural assumptions. Then the price type j is demands at t is

$$S^j_t = \sup_{\tau\geq 0}\mathbb{E}^j_t\left[\int_t^\tau e^{-r(s-t)}\,\mathrm{d}D_s + e^{-r(\tau-t)}(S^i_\tau - c)\,\mathrm{d}s\right].$$

Type j's problem here is like the problem of choosing the optimal stopping time for exercising an American call option that also pays dividends, where the price process of the underlying security for the American option is type i's willingness to pay, S^i and the strike price is c.

Noting type j's conditional mean beliefs are a martingale and $\mathrm{d}D_t = \hat{\mu}^j_t\,\mathrm{d}t + \sigma\,\mathrm{d}\bar{B}^{1,j}_t$, we have from the evolution of conditional mean beliefs that

$$\int_t^\tau e^{-r(s-t)}\,\mathrm{d}D_s = \int_t^\tau e^{-r(s-t)}\left(\bar{\mu} + e^{-\lambda(s-t)}(\hat{\mu}^j_s - \bar{\mu})\right)\mathrm{d}s + M^j_\tau - M^j_t$$

where M^j is some martingale for type j so $\mathbb{E}^j_t[M^j_\tau - M^j_t] = 0$. Substituting this into the expression for S^j_t gives

$$S^j_t = \sup_{\tau\geq 0}\mathbb{E}^j_t\left[\int_t^\tau e^{-r(s-t)}\left(\bar{\mu} + e^{-\lambda(s-t)}(\hat{\mu}^j_s - \bar{\mu})\right)\mathrm{d}s + e^{-r(\tau-t)}(S^i_t - c)\,\mathrm{d}s\right].$$

Suppose equilibrium price type j demands at t takes the form

$$S^j_t = f(\hat{\mu}^j_t, \Delta^j_t) = g(\mu^j_t) + h(\Delta^j_t),$$

[15] More generally, one can partially relax this to allow some short-selling, provided the opportunity for each type i to earn speculative profits from the other type remains.

where the fundamental valuation component $g(\mu_t^j) = \frac{\bar{\mu}}{r} + \frac{\hat{\mu}_t^j - \bar{\mu}}{r+\lambda}$ is the expected present value of future dividends under j's beliefs, and the resale option valuation component $h : \mathbb{R} \to \mathbb{R}$ is some positive, strictly increasing continuous function that depends on the gap between j's beliefs and the beliefs of the other type. From j's perspective, this resale option component is the bubble component.

Conjecturing equilibrium takes this form and substituting into the expression for S_t^j gives

$$S_t^j = \frac{\bar{\mu}}{r} + \frac{\hat{\mu}_t^j - \bar{\mu}}{r+\lambda} + \sup_{\tau \geq 0} \mathbb{E}_t^j \left[e^{-r(\tau - t)} \left(\frac{\Delta_\tau^j}{r+\lambda} + h(\Delta_\tau^i) - c \right) \right].$$

The resale option valuation h thus satisfies

$$h(\Delta_t^j) = \sup_{\tau \geq 0} \mathbb{E}_t^j \left[e^{-r(\tau - t)} \left(\frac{\Delta_\tau^j}{r+\lambda} + h(\Delta_\tau^i) - c \right) \right].$$

Since type j can sell immediately to type i, the value of the option must be at least S_t^i: the region where S_t^i weakly exceeds the value of the resale option is the stopping region. From type j's perspective, the deflated value process $e^{-rt} h(\Delta_t^j)$ for the resale option is a martingale in the continuation region ($t < \tau$) and a supermartingale in the stopping region ($t \geq \tau$). From Itô's lemma and the evolution of Δ^j, we can write these as

$$\begin{aligned} h(x) &\geq \frac{x}{r+\lambda} + h(-x) - c, \\ 0 &\geq \frac{1}{2}\sqrt{2}\varphi v h''(x) - \rho x h'(x) - r h(x), \end{aligned}$$

where the second inequality holds strictly when the first holds with equality. The smooth pasting implies that h is continuously differentiable. Moreover, conjecture the continuation region is some lower interval $(-\infty, \bar{\Delta})$ for some threshold difference in beliefs $\bar{\Delta} > 0$. In the continuation region, the value function h solves the second order ODE,

$$\varphi^2 v^2 u''(x) - \rho x u'(x) - r u(x) = 0.$$

Now note if u satisfies this ODE then it can be written as $u(x) = w\left(\frac{\rho}{2\varphi^2 v^2} x^2\right)$, where w solves the second order ODE

$$y w''(y) + \left(\frac{1}{2} - y\right) w'(y) - \frac{r}{2\rho} w(y) = 0.$$

This second order ODE for w is a very well-studied form of ODE known as *Kummer's equation.* Solutions to Kummer's equation take the form

$$w(y) = \alpha M\left(\frac{r}{2\rho}, \frac{1}{2}, y\right) + \beta U\left(\frac{r}{2\rho}, \frac{1}{2}, y\right)$$

for arbitrary constants α, β and confluent hypergeometric functions M and U defined by

$$M(a,b,y) = \sum_{n=0}^{\infty} \frac{a^{(n)} y^n}{b^{(n)} n!}, \qquad \text{where } a^{(n)} := a(a+1)\cdots(a+n-1) \text{ and } a^{(0)} := 1,$$

and

$$U(a,b,y) = \frac{\pi}{\sin(\pi b)} \left(\frac{M(a,b,y)}{\Gamma(1+a-b)\Gamma(b)} - y^{1-b} \frac{M(1+a-b, 2-b, y)}{\Gamma(a)\Gamma(2-b)} \right),$$

with Γ being the gamma function. From here, one can show that

$$\hat{u}(x) = \begin{cases} U\left(\frac{r}{2\rho}, \frac{1}{2}, \frac{\rho}{2\varphi^2 v^2}\right) & \text{if } x \leq 0, \\ \alpha M\left(\frac{r}{2\rho}, \frac{1}{2}, \frac{\rho}{2\varphi^2 v^2}\right) - U\left(\frac{r}{2\rho}, \frac{1}{2}, \frac{\rho}{2\varphi^2 v^2}\right) & \text{if } x > 0, \end{cases}$$

where $\alpha = \frac{2\pi}{\Gamma(1/2 + r/(2\rho))\Gamma(1/2)}$ and $\hat{u}(0) = \frac{\alpha}{2}$, is a solution to the second order ODE in u. For $x \in (-\infty, 0)$, it is also strictly positive and increasing, and any other solution that is strictly positive and increasing on $(-\infty, 0)$ must take the form $u(x) = \beta_1 \hat{u}(x)$ for some constant $\beta_1 > 0$. In fact, $\hat{u} > 0$ $\hat{u}' > 0$, $\hat{u}'' > 0$ and $\hat{u}''' > 0$ everywhere, and $\lim_{x \to -\infty} \hat{u}(x) = 0$ and $\lim_{x \to -\infty} \hat{u}'(x) = 0$.

Given h is strictly increasing and positive on $(-\infty, \bar{\Delta})$, we have

$$h(x) = \begin{cases} \beta_1 \hat{u}(x) & \text{for } x < \bar{\Delta}, \\ \frac{x}{r+\lambda} + \beta_1 \hat{u}(-x) - c & \text{for } x \geq \bar{\Delta}. \end{cases}$$

Moreover, since it is continuous and, by smooth pasting, continuously differentiable at the threshold $\bar{\Delta}$, we have

$$\begin{aligned} \beta_1 \hat{u}(\bar{\Delta}) - \frac{\bar{\Delta}}{r+\lambda} - \beta_1 \hat{u}(-\bar{\Delta}) + c &= 0 && \text{(by continuity)}, \\ \beta_1 \hat{u}'(\bar{\Delta}) + \beta_1 \hat{u}'(-\bar{\Delta}) - \frac{1}{r+\lambda} &= 0 && \text{(differentiating both sides)}, \end{aligned}$$

implying $\beta_1 = \frac{1}{(\hat{u}'(\bar{\Delta}) + \hat{u}'(-\bar{\Delta}))(r+\lambda)}$ and that $\bar{\Delta}$ solves

$$(\bar{\Delta} - c(r+\lambda))(\hat{u}'(\bar{\Delta}) + \hat{u}'(-\bar{\Delta})) - \hat{u}(\bar{\Delta}) + \hat{u}(\bar{\Delta}) = 0.$$

There is a unique solution $\bar{\Delta}$ to this equation for any nonnegative cost c (in particular, $c = 0$ gives $\bar{\Delta} = 0$). Having obtained the threshold $\bar{\Delta}$, we can fully characterize the resale option value h as

$$h(x) = \begin{cases} \frac{b}{\hat{u}(-\bar{\Delta})} \hat{u}(x) & \text{if } x < \bar{\Delta}, \\ \frac{x}{r+\lambda} + \frac{b}{\hat{u}(-\bar{\Delta})} \hat{u}(-x) + c & \text{if } x \geq \bar{\Delta}, \end{cases}$$

where $b := \hat{u}(-\bar{\Delta}) = \frac{1}{r+\lambda}\frac{\hat{u}(-\bar{\Delta})}{\hat{u}'(\bar{\Delta})+\hat{u}'(-\bar{\Delta})}$. This now completely characterizes the equilibrium.

Note if type j are the current owners at time t, then the optimal stopping time τ^t at which j sells the asset takes the form $\tau^t(\omega) = \inf\{s \geq t \mid \Delta^j_s(\omega) \geq \bar{\Delta}\}$, so is a function only of the future path of Δ^j. If type j has just acquired the security by purchasing it from type i at time t, then type i must have "exercised the resale option", which is optimal at $\Delta^i_t = \bar{\Delta}$. Since $\Delta^i_t = -\Delta^j_t$, we have $\Delta^j_t = -\Delta$. The expected wait at t for type j until selling to type i is thus $\mathbb{E}^j[\tau - t \mid \Delta^j_t = -\bar{\Delta}]$. Now, suppose costs c are zero. Then $\bar{\Delta} = 0$. Since the difference in belief process Δ^j_t is an Itô process with strictly positive volatility, if it hits 0 at t then with probability 1 it crosses zero infinitely many times in any interval $(t, t+\epsilon)$ for $\epsilon > 0$, and the expected time between trades if a trade has just occurred is zero. Trade volume is thus infinite in the vicinity of any trade. Of course, in any time interval where the path of Δ^j does not cross zero, no trade occurs and the expected time until a trade is nonzero. The expected time until the next trade is a continuous function of c, so if the trading cost $c > 0$ is very small, we would expect large trading volumes in the vicinity of any trade.

Even if trading costs are zero so trade occurs whenever beliefs cross, there is still a bubble – i.e. the resale option has a positive price, which when exercised at $\bar{\Delta} = 0$ is $b = \frac{1}{r+\lambda}\frac{\hat{u}(0)}{2\hat{u}'(0)} > 0$. This is even though the security swaps hands as soon as one type becomes more bullish than the other. It is easier to see where the bubble comes from by considering some small cost $c > 0$ so that $\bar{\Delta} = \epsilon > 0$ for some small ϵ. Letting b^j be the value of the resale option to type j when selling when $\Delta^j_t = \epsilon$ and letting b^i be the value of the resale option to type i when selling when $\Delta^j_t = -\epsilon$, we have $b^j = \left(\frac{\epsilon}{r+\lambda} + b^i\right)\frac{\hat{u}(-\epsilon)}{\hat{u}(\epsilon)}$.[16] By symmetry, $b^j = b^i$ so $b^j = b^i = \frac{\epsilon}{r+\lambda}\frac{\hat{u}(-\epsilon)}{\hat{u}(\epsilon)-\hat{u}(-\epsilon)}$. This converges to b as $\epsilon \to 0$. The bubble exists in the limit because as we approach the limit, the volume of trade approaches infinity, and the vanishing gains traders make per trade nevertheless aggregate to a positive bubble. The size of the bubble is increasing in the variance $\sqrt{2}\varphi v$ and mean reversion coefficient ρ the difference in belief process Δ^j.

As with real-world bubbles, the bubbles in Scheinkman & Xiong (2003) are accompanied by high trading volumes. The bubble is sustained by traders' hopes to resell to other traders, fitting the typical definition of a "bubble" (as described by, say, Keynes, 1931) perfectly.

[16] $\frac{\hat{u}(-\epsilon)}{\hat{u}(\epsilon)}$ is the rate at which type j discounts the cashflow from the sale of the security.