Online Appendix for "Sentiment and speculation in a market with heterogeneous beliefs"

Ian Martin Dimitris Papadimitriou[*]

March, 2022

1 Miscellaneous results

First of all, let us write down a version of the de Moivre–Laplace theorem; this theorem is essentially a special case of the Central limit theorem that first appeared in 1716 in de Moivre's *The Doctrine of Chances*. For a proof, see the textbook of Chung (2012).

Theorem 1. *Suppose* $0 < p_n < 1$, $p_n + q_n = 1$, $p_n \to p$ *and*

$$x_k = \frac{k - np_n}{\sqrt{np_nq_n}},\ 0 \le k \le n$$

Let A be an arbitrary, fixed positive number. Then in the range of k such that $|x_k| \le A$ we get

$$\binom{n}{k} p_n^k q_n^{n-k} \sim \frac{1}{\sqrt{2\pi np_nq_n}} e^{-\frac{x_k^2}{2}}$$

where the convergence is uniform and the notation $\sim$ means that the ratio of the right hand side to the left hand side tends to 1 as $n \to \infty$.[1] *Moreover if S_n has the Binomial(n, p_n) distribution then, for any two constants $a < b$ we have:*

$$\lim_{n\to\infty} P\left(a \le \frac{S_n - np_n}{\sqrt{np_nq_n}} \le b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx$$

We also use a stronger version of the de Moivre–Laplace theorem presented in Osius (1989) which implies that there is convergence of the moment generating function of a

[*]http://personal.lse.ac.uk/martiniw/ and https://sites.google.com/site/papadimitrioudea/.

[1] $f(x) \sim g(x)$ is equivalent to $f(x) = g(x)(1 + o(1))$.

standardized binomial to that of a standardized Normal (there is *convergence of infinite exponential order*).

We use a central limit theorem for beta-binomial random variables that appears[2] in Paul and Plackett (1978) in a slightly generalized form that allows α and β to have a term proportional to $\sqrt{N}$.

Theorem 2. *If $Y \sim BB(\overline{\lambda}N, \alpha, \beta)$, where $\overline{\lambda} > 0$, $\alpha = \theta N + \eta\sqrt{N}, \beta = \theta N - \eta\sqrt{N}$, and we let $N \to \infty$, then:*

$$\frac{Y - \frac{1}{2}\overline{\lambda}N - \frac{\eta}{2\theta}\overline{\lambda}\sqrt{N}}{\sqrt{\frac{(\overline{\lambda}+2\theta)}{8\theta}\overline{\lambda}N}} \to N(0,1)$$

Note that the convergence of Beta-Binomial distribution to Normal holds not only in distribution but also in moment generating functions. Indeed, by the Moment Continuity Theorem, convergence in distribution of subgaussian random variables implies convergence of moment generating functions.

1.1 Proof of Lemma 2

This section provides a proof of Lemma 2, which we restate in its general form:

Lemma. *If $Y_1 \sim BB(T, \overline{\alpha}, \lambda\overline{\alpha})$ and $Y_2 \sim BB(T, \underline{\alpha}, \lambda\underline{\alpha})$, for $\overline{\alpha} > \underline{\alpha}$ and $\lambda > 0$ then Y_1 second order stochastically dominates Y_2.*

Proof. First note that $\mathbb{E}[Y_1] = \mathbb{E}[Y_2] = \frac{T}{1+\lambda}$. Write $f_{\overline{\alpha}}(\cdot)$ and $f_{\underline{\alpha}}(\cdot)$ for the probability mass functions of Y_1 and Y_2, respectively. It is enough to show that the likelihood ratio $f_{\overline{\alpha}}(k)/f_{\underline{\alpha}}(k)$ is increasing for integers $k \in [0, T/(1+\lambda))$ and decreasing for integers $k \in (T/(1+\lambda), T]$. This implies that Y_1 second order stochastically dominates Y_2, by Theorem 2.2 of Ramos et al. (2000).[3]

We start by showing that

$$\frac{B(k+\overline{\alpha}, T-k+\lambda\overline{\alpha})}{B(k+\underline{\alpha}, T-k+\lambda\underline{\alpha})}$$

(that is, the likelihood ratio, up to a positive constant of proportionality) is increasing for integers $k \in [0, T/(1+\lambda)]$. Pick k_1 between 1 and $T/(1+\lambda)$ and let $k_2 = k_1 - 1$. We must show that

$$\frac{\Gamma(k_1+\overline{\alpha})\Gamma(T-k_1+\lambda\overline{\alpha})}{\Gamma(k_1+\underline{\alpha})\Gamma(T-k_1+\lambda\underline{\alpha})} > \frac{\Gamma(k_2+\overline{\alpha})\Gamma(T-k_2+\lambda\overline{\alpha})}{\Gamma(k_2+\underline{\alpha})\Gamma(T-k_2+\lambda\underline{\alpha})}. \tag{1}$$

[2]We caution the reader that there is a typo in the theorem as stated by Paul and Plackett (1978): the random variable is not correctly standardized.

[3]This result applies for continuous random variables, but it is straightforward to adapt the result to the discrete case which is relevant here.

As $\Gamma(z+1) = z\Gamma(z)$ for any positive real z, this reduces to

$$\frac{k_2 + \overline{\alpha}}{T - k_1 + \lambda\overline{\alpha}} > \frac{k_2 + \underline{\alpha}}{T - k_1 + \lambda\underline{\alpha}}\,,$$

which is equivalent to $k_1 + \lambda k_2 < T$. This holds because $k_1 \leq T/(1+\lambda)$ and $k_2 < k_1$. Conversely, if $k_2 = k_1 - 1 \geq \frac{T}{1+\lambda}$ then the inequality (1) reverses as then $k_1 + \lambda k_2 > T$. $\square$

1.2 Proof of Lemma 4

This section provides the Proof of Lemma 4 which we restate here:

Lemma. *Write $W(p_T)$ for an investor's wealth at time T, as a function of the price of the risky asset p_T. Suppose that $W(0) = 0$. Then terminal wealth $W(p_T)$ can be achieved by holding a portfolio of (i) $W'(K_0)$ units of the underlying asset, (ii) bonds with face value $W(K_0) - K_0 W'(K_0)$, (iii) $W''(K)\,dK$ put options on the risky asset maturing at time T with strike K, for every $K < K_0$, and (iv) $W''(K)\,dK$ call options maturing at time T with strike K, for every $K > K_0$. The constant $K_0 > 0$ can be chosen arbitrarily.*

Proof. We can write

$$\begin{aligned} W(p_T) = {} & W'(K_0)p_T + [W(K_0) - K_0 W'(K_0)] + \\ & + \left\{ \int_0^{K_0} W''(K) \max\{0, K - p_T\}\, dK + \int_{K_0}^{\infty} W''(K) \max\{0, p_T - K\}\, dK \right\}. \end{aligned}$$

To see this, integrate the right-hand side by parts. The result follows after noting that the three terms on the right-hand side are, respectively, the payoff on a position in $W'(K_0)$ units of the underlying asset; the payoff on bonds with face value $W(K_0) - K_0 W'(K_0)$; and the payoffs on the portfolios of put and call options described in the statement. $\square$

1.3 Proof of Lemma 5

This section provides the Proof of Lemma 5 which we restate here:

Lemma. *If $\theta > 1$, then the sequence $\left\{(M_t^2)^{(N)}\right\}$ (where we include superscripts to emphasize the dependence on N) is uniformly integrable.*

Proof. The proof of Result 9 showed that $(M_t^2)^{(N)}$ is asymptotically equivalent to $De^{A(\psi^{(N)})^2 + B\psi^{(N)} + C}$, where $\psi^{(N)} = \frac{m - \phi N - \frac{\eta}{\theta}\phi\sqrt{N}}{\sqrt{\phi N/2}}$, $A = \frac{\phi}{\phi+\theta}$, and B, C, and D are constants. We will show that there exists an $\varepsilon > 0$ such that $\sup_N \mathbb{E}[(e^{A(\psi^{(N)})^2 + B\psi^{(N)} + C})^{1+\varepsilon}] < \infty$, which implies that $(M_t^2)^{(N)}$ is uniformly integrable. Let us set $\varepsilon = 1 - 2A = 1 - \frac{2\phi}{\phi+\theta} > 0$.

By Hoeffding's inequality,

$$\mathbb{P}\left(\left|\frac{m-\phi N}{\sqrt{\phi N}}\right| \geq k\right) \leq 2e^{-k^2} \tag{2}$$

for any $k > 0$. For $x > 0$, this implies that $\mathbb{P}\left(e^{\frac{1}{1+\varepsilon^2}\frac{(m-\phi N)^2}{\phi N}} \geq x\right) \leq \frac{2}{x^{1+\varepsilon^2}}$. As $e^{\frac{1}{1+\varepsilon^2}\frac{(m-\phi N)^2}{\phi N}} \geq 1$, we have

$$\mathbb{E}[e^{\frac{1}{1+\varepsilon^2}\frac{(m-\phi N)^2}{\phi N}}] \leq \int_0^\infty \mathbb{P}\left(e^{\frac{1}{1+\varepsilon^2}\frac{(m-\phi N)^2}{\phi N}} \geq x\right) dx \leq 1 + \int_1^\infty \frac{2}{x^{1+\varepsilon^2}} dx < \infty\,.$$

Finally, note that $\psi = \sqrt{2}(\frac{m-\phi N}{\sqrt{\phi N}} - \sqrt{\phi}\frac{\eta}{\theta})$ and $2(1+\varepsilon)A < 1/(1+\varepsilon^2)$. Thus there is a constant, K, such that $(1+\varepsilon)(A\psi^2 + B\psi + C) < \frac{1}{1+\varepsilon^2}\frac{(m-\phi N)^2}{\phi N} + K$, and therefore $\mathbb{E}[(e^{A(\psi^{(N)})^2 + B\psi^{(N)} + C})^{1+\varepsilon}] < \mathbb{E}[e^{\frac{1}{1+\varepsilon^2}\frac{(m-\phi N)^2}{\phi N} + K}] < \infty$. $\square$

2 Result 1 via dynamic completeness

This section presents an alternative proof of Result 1 which exploits the fact that the market is dynamically complete. We thank an anonymous referee for suggesting this approach.

Proof. We write q_m for the risk neutral probability of reaching node (m, T), $\pi_m^{(h)}$ for the corresponding probability from the perspective of investor h, $w_{m,T}^{(h)}$ for the wealth of investor h at node (m, T), and $u(\cdot)$ for utility over terminal wealth. As the market is dynamically complete, we can equivalently solve for the static Arrow–Debreu equilibrium, in which investors choose their terminal consumption $w_{m,T}^{(h)}$ to solve the problem

$$\max_{w_{m,T}^{(h)}} \sum_{m=0}^{T} \pi_m^{(h)} u(w_{m,T}^{(h)}) \quad \text{subject to} \quad \sum_{m=0}^{T} q_m w_{m,T}^{(h)} = p_0\,.$$

The first-order conditions for this problem are that

$$\pi_m^{(h)} u'(w_{m,T}^{(h)}) = \lambda^{(h)} q_m \quad \text{for all } m\,,$$

where $\lambda^{(h)}$ is the Lagrange multiplier on agent h's budget constraint. We can now use the budget constraint to solve for $\lambda^{(h)}$, giving

$$\sum_m \frac{\pi_m^{(h)} u'(w_{m,T}^{(h)})}{\lambda^{(h)}} w_{m,T}^{(h)} = p_0 \, .$$

With log utility, this implies that $\lambda^{(h)} = \frac{1}{p_0}$ for all agents h, and hence that

$$w_{m,T}^{(h)} = \frac{\pi_m^{(h)} p_0}{q_m} \, .$$

As aggregate wealth at node (m, T) is $p_{m,T}$, it follows that

$$\frac{\int_0^1 \pi_m^{(h)} p_0 f(h) dh}{q_m} = p_{m,T} \, .$$

The net riskless rate is zero, so $\sum_m q_m = 1$, and hence

$$\sum_{m=0}^{T} \left(\int_0^1 \pi_m^{(h)} p_0 f(h) \, dh \right) z_{m,T} = 1 \, ,$$

where we write $z_{m,T} = 1/p_{m,T}$ as in the main text. Hence the price at time 0 satisfies

$$p_0 = \left[\sum_0^T \left(\int_0^1 \pi_m^{(h)} f(h) \, dh \cdot z_{m,T} \right) \right]^{-1} \, .$$

As $\pi_m^{(h)} = \binom{T}{m} h^m (1-h)^{T-m}$, this gives the result. $\square$

3 Survival

In this section we assume that there is some true probability of an up-move h_{true}, and we are interested in the evolution of the wealth distribution after many periods in the baseline model in which agents do not learn. (We discuss what happens when agents learn in the next section.)

For convenience, we will refer to the share of wealth held by type-h agents as $\Omega(h, m, n)$. As shown in the main text,

$$\Omega(h, m, n) = \frac{h^m (1-h)^n f(h)}{\int_0^1 h^m (1-h)^n f(h) \, dh} \, .$$

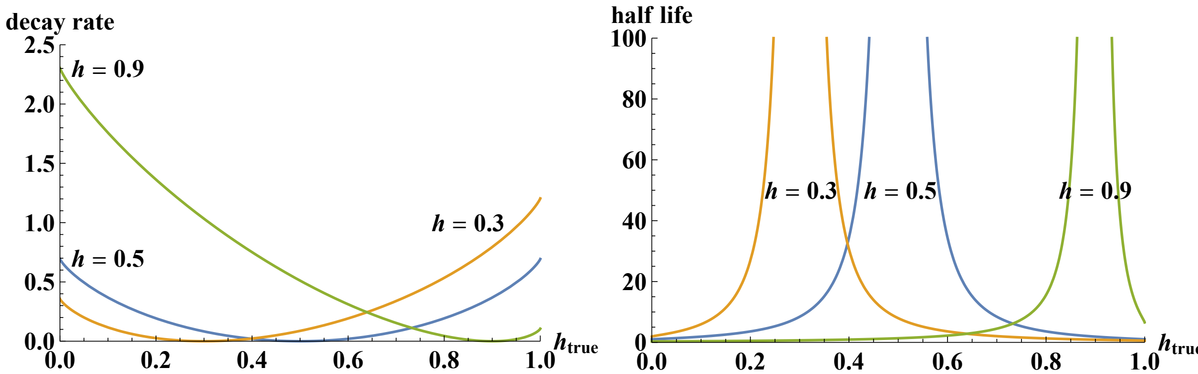


Figure 1: Decay rate (left panel) and half-life (right panel) plotted against the true up-probability, h_{true}, for investors $h = 0.3, 0.5, 0.9$.

We are interested in the case $m = h_{\text{true}} t$ and $n = (1 - h_{\text{true}})t$, where the elapsed number of periods, t, is large. For $h \neq h_{\text{true}}$, the share of wealth will (asymptotically) decay exponentially. Thus we focus on the asymptotic rate of exponential decay in agent h's share of wealth, namely,

$$\begin{aligned}
\lim_{t\to\infty} -\frac{1}{t}\log\Omega(h,m,n) &= \lim_{t\to\infty} -\frac{1}{t}\log\frac{h^{h_{\text{true}}t}(1-h)^{(1-h_{\text{true}})t}f(h)}{\int_0^1 h^{h_{\text{true}}t}(1-h)^{(1-h_{\text{true}})t}f(h)\,dh} \\
&= -h_{\text{true}}\log h - (1-h_{\text{true}})\log(1-h) + \lim_{t\to\infty}\frac{\log f(h)}{t} + \\
&\quad + \lim_{t\to\infty}\frac{1}{t}\log\int_0^1 h^{h_{\text{true}}t}(1-h)^{(1-h_{\text{true}})t}f(h)\,dh \\
&= -h_{\text{true}}\log h - (1-h_{\text{true}})\log(1-h) + \lim_{t\to\infty}\frac{\log f(h)}{t} + \\
&\quad + \lim_{t\to\infty}\frac{1}{t}\log\int_0^1 \exp\left\{\left[h_{\text{true}}\log h + (1-h_{\text{true}})\log(1-h)\right]t + \log f(h)\right\}\,dh \\
&= -h_{\text{true}}\log h - (1-h_{\text{true}})\log(1-h) + 0 + \\
&\quad + \sup_h\left\{h_{\text{true}}\log h + (1-h_{\text{true}})\log(1-h)\right\} \\
&= h_{\text{true}}\log\frac{h_{\text{true}}}{h} + (1-h_{\text{true}})\log\frac{1-h_{\text{true}}}{1-h}\,. \qquad (3)
\end{aligned}$$

This quantity is strictly positive when $h \neq h_{\text{true}}$, indicating that the wealth share of all types other than the objectively correct type decays exponentially fast for large t.

The left panel of Figure 1 shows how the asymptotic decay rate (3) varies, as a function of the true probability h_{true}, for investors with $h = 0.3$, 0.5, and 0.9. The right panel shows the asymptotic half-life, $\frac{\log 2}{\text{decay rate}}$; this represents the number of time periods required for an investor's wealth share to halve. Although investors who are incorrect,

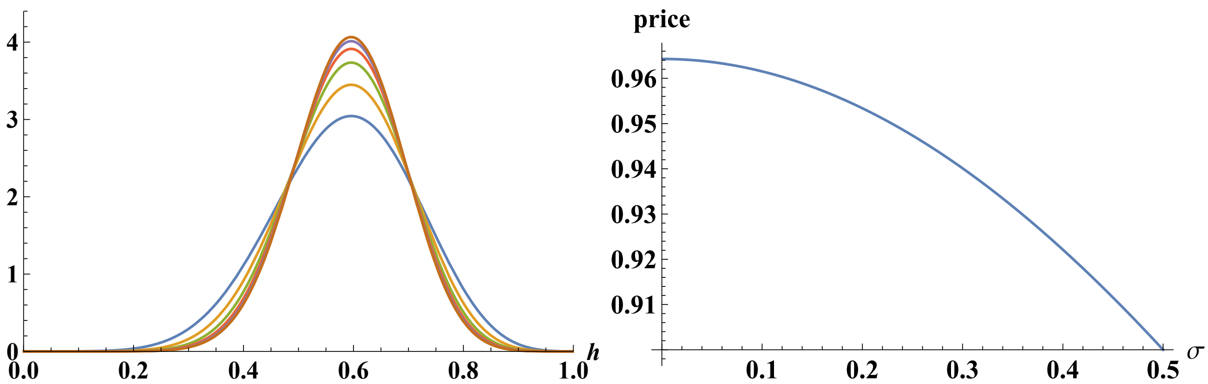


Figure 2: Left: The wealth distribution, plotted against agent type h, after $t = 25$, 50, 100, 200, 400, and 800 periods. The figure assumes that 60% of moves are up-moves and sets $\zeta = 24$, as in the example illustrated in Figure 3 of Section 2.1. Right: The time 0 price of the risky asset in the example studied in Figure 2 of the main paper, if agents learn over time, as a function of $\sigma = \frac{1}{\sqrt{4(1+\zeta)}}$, the standard deviation of the median agent's prior belief.

$h \neq h_{\text{true}}$, experience exponential decay in wealth share, it takes several periods for investors who are roughly correct to become irrelevant. For example, the half-life for investor $h = 0.5$ is more than 34 periods for all values of h_{true} between 0.4 and 0.6.

4 Learning

As noted in the main text, all agents survive asymptotically in the case in which agents learn, as the truth is in the support of every agent's prior, and every agent's posterior belief will converge to the truth (Blume and Easley, 2006). The left panel of Figure 2 illustrates shows the wealth distribution as a function of agent type in the case $\zeta = 24$, as in the example illustrated in Figure 3 of Section 2.1. The figure assumes that 60% of moves are up-moves and plots the wealth distribution after 25, 50, 100, 200, 400, and 800 periods.

The right panel of Figure 2 shows how the time 0 price of the risky asset in the illustrative example provided in Figure 2 of the main paper varies if agents learn over time. The variable on the x-axis is σ, the standard deviation of the median agent's prior belief; in the example in the main paper, $\sigma = 0$. The maximum possible standard deviation is $\sigma = 1/2$. As $1/p_{m,t}$ is convex in this example (which could be checked directly, but in this case can be seen immediately using the sufficient condition that $\log p_{m,t}$ is weakly concave), the price declines as investors' prior uncertainty increases (i.e., as ζ decreases), as shown more generally in Result 4.

5 Volume and leverage

Recall that the *leverage ratio* of investor h, which we define as the ratio of funds borrowed, $x_h p - w_h$, to wealth, w_h, is

$$\frac{x_h p - w_h}{w_h} = \frac{h - H_{m,t}}{H_{m,t} - h^*}.$$

If $p_u > p_d$ then $h^* < H_{m,t}$, by equation (8) of the paper; in this case the above equation shows that people who are optimistic relative to the representative investor borrow from pessimists. We can define *gross leverage* as the total dollar amount these optimists borrow, scaled by aggregate wealth:

$$\begin{aligned}\frac{\int_{H_{m,t}}^1 (x_h p - w_h) f(h)\, dh}{p} &= \int_{H_{m,t}}^1 \frac{w_h f(h)}{p} \frac{h - H_{m,t}}{H_{m,t} - h^*}\, dh \\ &= \frac{H_{m,t}^{m+\alpha}(1 - H_{m,t})^{n+\beta}}{(m + \alpha + n + \beta) B(\alpha + m, \beta + n)(H_{m,t} - h^*)},\end{aligned}$$

where we use the fact that $H_{m,t}^{m+\alpha}(1 - H_{m,t})^{n+\beta} = -\int_{H_{m,t}}^1 \left(h^{m+\alpha}(1-h)^{n+\beta}\right)' dh$ to derive the final expression. Conversely, if $p_u < p_d$ then optimists are lenders and pessimists borrowers. In either case, we can define gross leverage as the absolute value of the above expression,

$$\frac{H_{m,t}^{m+\alpha}(1 - H_{m,t})^{n+\beta}}{(m + \alpha + n + \beta) B(\alpha + m, \beta + n)\, |H_{m,t} - h^*|}. \tag{4}$$

5.1 Volume and leverage in the risky bond and bubbly asset examples

The left panel of Figure 3 shows the time series of volume and gross leverage in the risky bond example with $\varepsilon = 0.3$, assuming bad news arrives each period. (If good news arrives at any stage, volume drops permanently to zero.) There is a burst of trade at first: volume substantially exceeds the total supply of the asset initially, as agents with extreme views undertake highly leveraged trades, but declines rapidly over time as wealth becomes concentrated in the hands of investors with similar beliefs. The right panel shows the corresponding series if $\varepsilon = 0.9$.

Assuming a down-movement in the transition from time t to time $t+1$, the volume of trade (in terms of the number of units of the risky asset transacted) is

$$\frac{1}{2}\int_0^1 \left| \frac{(1-h)^t}{\frac{1}{1+t}} \frac{h - h_t^*}{H_{0,t} - h_t^*} - \frac{(1-h)^{t+1}}{\frac{1}{2+t}} \frac{h - h_{t+1}^*}{H_{0,t+1} - h_{t+1}^*} \right| dh = \frac{4(1+t)^{1+t}}{(3+t)^{3+t}} \cdot \left| 1 + t + \frac{1 + \varepsilon T}{1 - \varepsilon} \right|,$$

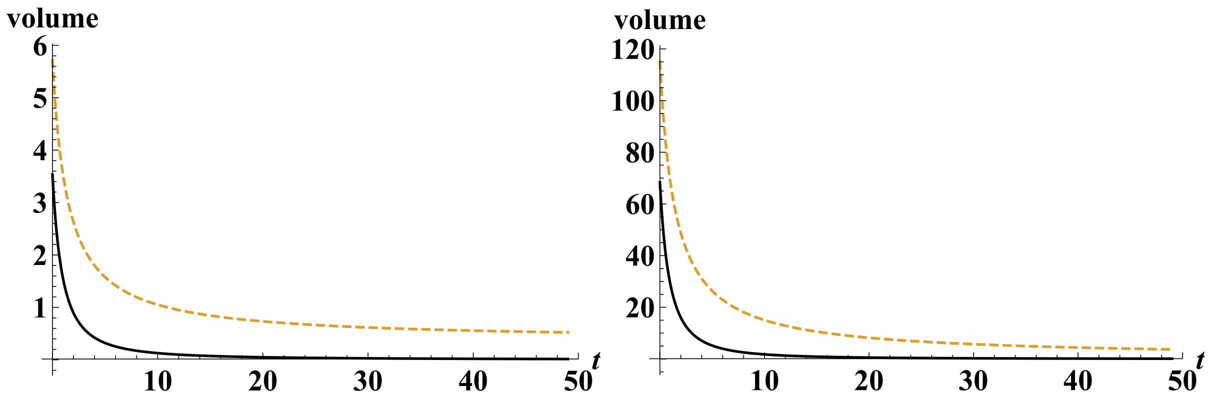


Figure 3: Volume (solid) and gross leverage (dashed) over time in the risky bond example, with $\varepsilon = 0.3$ (left) or $\varepsilon = 0.9$ (right).

where we include the factor of 1/2 to avoid double-counting.

Gross leverage, in the same transition, calculated from (4), is

$$\left(\frac{1+t}{2+t}\right)^{2+t}\left|1+\frac{1+T}{1+t}\frac{\varepsilon}{1-\varepsilon}\right|.$$

Note that the above formulas hold both for the risky bond case ($\varepsilon < 1$) and for the bubbly asset case ($\varepsilon > 1$). (For simplicity, we assume here that in the bubbly asset case, the good outcome $\varepsilon > 1$ occurs at the bottom node, rather than at the top node as in the paper. This makes no substantive difference and allows us to use the same formula as for the risky bond case, but means that throughout this subsection a down-move must be interpreted as good news in the bubbly asset case.) In the former case volume and gross leverage are each increasing in ε and in T. The safer the bond is, the more aggressively agents trade on their disagreement without risking ruin, as the relative safety of the asset permits agents to take on more leverage: extremists on both sides of the market are trying to "pick up nickels in front of a steamroller." Similarly, in the bubbly asset case both volume and gross leverage are decreasing in ε.

6 Static trade in the risky bond example

This section contains some further calculations in the risky bond example, assuming $f(h) = 1$. Suppose agents are not allowed to trade dynamically. Agent h perceives a probability $1-(1-h)^T$ that the bond pays 1, and $(1-h)^T$ that the bond pays ε, so solves

$$\max_{x_h} \left(1-(1-h)^T\right) \log\left(w_h - x_h p + x_h\right) + (1-h)^T \log\left(w_h - x_h p + x_h \varepsilon\right).$$

The first-order condition (after setting $w_h = p$ to account for the fact that all agents are initially endowed with a unit of the risky asset) is

$$x_h = p\left(\frac{1-(1-h)^T}{p-\varepsilon} - \frac{(1-h)^T}{1-p}\right).$$

If T is reasonably large, most agents will have $(1-h)^T \approx 0$, and so will choose $x_h \approx \frac{p}{p-\varepsilon}$; their wealth in the bad state of the world is then approximately zero. Thus, if forced to trade statically most agents will lever up (almost) as much as possible without risking bankruptcy.

For the market to clear, we require $\int_0^1 x_h\, dh = 1$, which implies that $p = \frac{(1+T)\varepsilon}{1+T\varepsilon}$. This is the same as the time-0 price in the case with dynamic trade. It follows that agent h's demand for the asset is

$$x_h = 1 + \left(1-(1+T)(1-h)^T\right) \frac{1+T\varepsilon}{T(1-\varepsilon)}.$$

This is the quantity plotted as the "fundamental" position in Figure 5 of the paper.

7 Two calibrations in the Brownian limit

Figure 4 shows the two calibrations discussed in the main paper.

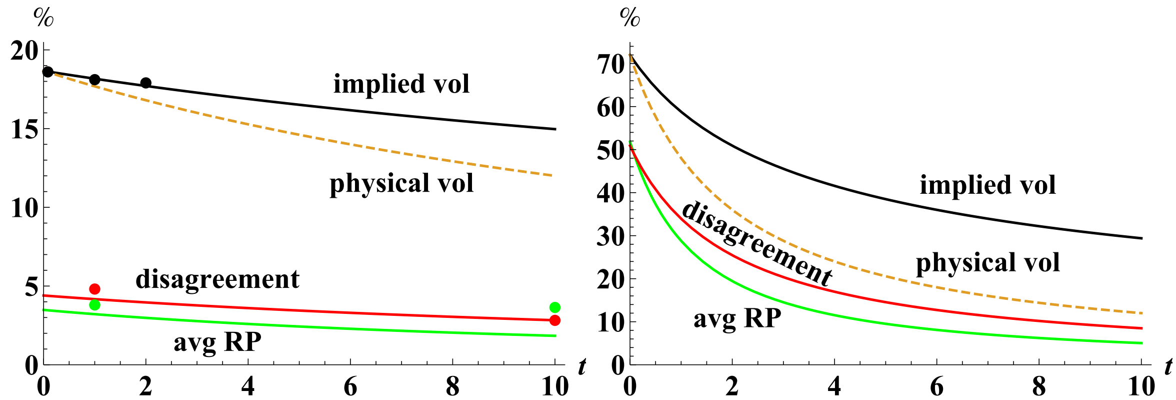
8 Option prices in the Poisson limit

To state our results in an economical way, we write

$$A = e^{-J}\left[1 - \frac{\omega^2\lambda(T-t)(e^J-1)}{1+\omega^2\lambda t}\right] < 1 \quad \text{and} \quad B = \left[1 - \frac{\omega^2\lambda(T-t)(e^J-1)}{1+\omega^2\lambda t}\right]^{1/\omega^2}, \tag{5}$$

so that the price of the risky asset, at time t, if q jumps have taken place, is BA^q.

The option pricing result is most cleanly stated when the strike $K = BA^k$, where $k \geq 0$ is an integer. For strikes not of this form, options are priced by interpolating linearly in strike: that is, if $\underline{K} = BA^{k+1}$ and $\overline{K} = BA^k$ and $K = \chi\underline{K} + (1-\chi)\overline{K}$ for $\chi \in (0,1)$, then the price of an option with strike K is the convex combination of the



(a) $\theta = 1.8$.

(b) $\theta = 0.2$.

Figure 4: Term structures of implied and physical volatility, mean expected returns and disagreement in the baseline (left) and crisis (right) calibrations.

prices of options with strikes $\underline{K}$ and $\overline{K}$, with weights χ and $1-\chi$, respectively. (To see this, note that the price of a butterfly spread constructed using options of all three strikes is zero, because the probability of the underlying asset's price lying strictly between $\underline{K}$ and $\overline{K}$ at expiry is zero by definition of $\underline{K}$ and $\overline{K}$.)

Result 1. *The time 0 price of a put option, expiring at time t, with strike BA^k is*[4]

$$p_0 \left(1-C\right)^{1/\omega^2} C^{1+k} \binom{k+1/\omega^2}{k+1} \left[\frac{1}{A} F\left(1, 1+k+1/\omega^2, 2+k, C/A\right) - F\left(1, 1+k+1/\omega^2, 2+k, C\right)\right],$$

where the price of the underlying asset is $p_0 = \left[1-\omega^2 \lambda T\left(e^J-1\right)\right]^{1/\omega^2}$, $k \geq 0$ is an integer, A and C are defined as in (5), $C = \omega^2\lambda t/(1+\omega^2\lambda t)$, and $F(\cdot,\cdot,\cdot,\cdot)$ is Gauss's hypergeometric function.

In the case $\omega = 1$, this simplifies to $\left(\frac{\lambda t}{1+\lambda t}\right)^{k+1}(1+\lambda T)(e^J-1)$, whereas if beliefs are homogeneous ($\omega = 0$) then the price of the put option is $e^{-\lambda t - \lambda T\left(e^J-1\right)} \sum_{q>k} \frac{(\lambda t)^q}{q!}\left(e^{J(q-k)}-1\right)$.

Proof of Result 1. Following the same logic as in the proof of Result 7 of the paper, the put price is $p_0\,\mathbb{E}\left[\left(A^{k-q}-1\right)^+\right]$. As $n \to \infty$, a beta binomial distribution with parameters n, α, ωn approaches a negative binomial distribution with parameters α and $1/(1+\omega)$, so q is asymptotically distributed NegativeBinomial$\left(1/\omega^2, C\right)$. Thus the put price is

$$p_0 \sum_{q>k} \binom{q+1/\omega^2-1}{q} \left(1-C\right)^{1/\omega^2} C^q \left(A^{k-q}-1\right)$$

[4]We write $\binom{n}{k}$ for the generalized binomial coefficient $\frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}$, which is defined for $k \in \mathbb{N}$ and arbitrary n.

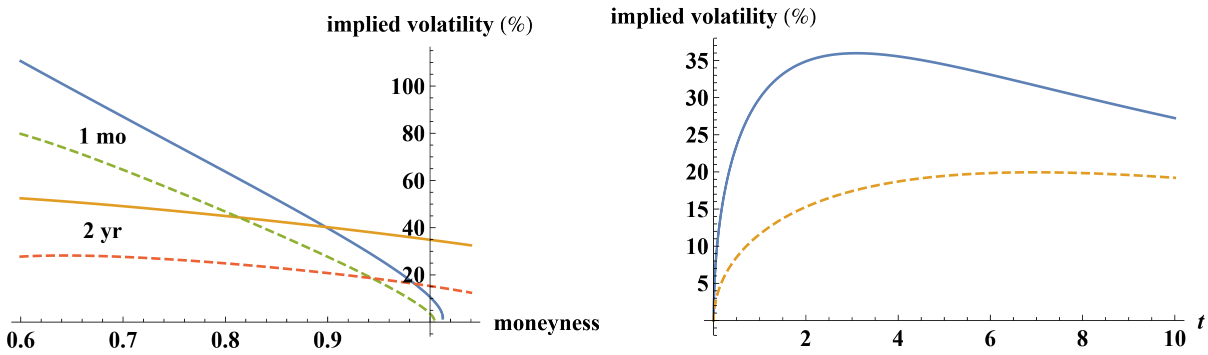


Figure 5: Left: The volatility smile for options of maturity 1 month and 2 years, in heterogeneous (solid) and homogeneous (dashed) belief economies. Right: The term structure of at-the-money implied volatility plotted against time-to-expiry, t, in heterogeneous (solid) and homogeneous (dashed) belief economies. Both panels use the baseline calibration.

which reduces to the given formulas in the cases $\omega > 0$ and $\omega = 1$.

As $\omega \to 0$, the negative binomial random variable q converges to a Poisson random variable with mean λt, and A approaches e^{-J}. Thus the put price is

$$p_0 \sum_{q=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^q}{q!} \left(e^{J(q-k)} - 1\right)^+ = e^{-\lambda t - \lambda T\left(e^J - 1\right)} \sum_{q>k} \frac{(\lambda t)^q}{q!} \left(e^{J(q-k)} - 1\right) . \qquad \square$$

Figure 5 illustrates. The left panel plots the Black–Scholes implied volatility for short-dated ($t = 1/12$) and long-dated ($t = 2$) options across a range of strikes, both with and without heterogeneity. Short-dated options exhibit a steeper smirk than long-dated options. Heterogeneity increases the level of volatility and further steepens the smirk relative to the homogeneous economy. The right panel plots the term structure of implied volatility for at-the-money options, which exhibits a hump shape in the presence of heterogeneous beliefs.

References

Blume, L. and Easley, D. (2006). If you're so smart, why aren't you rich? Belief selection in complete and incomplete markets. *Econometrica*, 74(4):929–966.

Chung, K. L. (2012). *Elementary probability theory with stochastic processes*. Springer Science & Business Media.

Osius, G. (1989). *Some results on convergence of moments and convergence in distribution with applications in statistics*. Fachber. Mathematik/Informatik, Univ.

Paul, S. R. and Plackett, R. L. (1978). Inference sensitivity for Poisson mixtures. *Biometrika*, 65(3):591–602.

Ramos, H. M., Ollero, J., and Sordo, M. A. (2000). A sufficient condition for generalized Lorenz order. *Journal of Economic Theory*, 90(2):286–292.