Memorable Events in Financial Markets

Andrés Carvajal* and Keisuke Teeple†

May 14, 2024

Abstract

We assume that financial traders *remember* certain days—for example, those when the trader was actively trading—and that these days need not coincide across traders. This disagreement leads to trade because those who recall bull markets buy and those who recall bear markets sell. In our equilibrium, which is plagued by purely non-fundamental volatility, a volatile price history generates belief dispersion and subsequent trade ensures continued price volatility. We then characterize the steady-state cross-sectional distribution when traders remember active trading days. In the cross-section, young (old) traders, who have not (have) experienced many trading days, have dispersed (concentrated) beliefs. Equilibrium prices display excess kurtosis due to the small chance that everyone trading on a particular day is young.

JEL classification: G41, G14, D84, D91

Keywords: Non-Bayesian models, non-fundamental volatility, heavy tails, experience effects

*acarvajal@ucdavis.edu, University of California Davis, One Shields Ave Davis, California 95616

†k2teeple@uwaterloo.ca, University of Waterloo, 200 University Ave West Waterloo, Ontario N2L3G1.

Teeple received financial support for this research from SSHRC Explore and Insight Development Grants.

1 Introduction

Why are asset prices so volatile? Shiller (2015) has famously shown that the majority of this volatility cannot be attributed to news alone. Here we introduce a theory of asset price volatility that relies critically on the past personal experiences of traders. Many empirical studies support this theory: Malmendier and Nagel (2011) show that individuals who experienced low stock market returns are more pessimistic about future stock returns. Kuchler and Zafar (2019) show that individuals use personal experiences to form expectations about aggregate economic outcomes like house prices and unemployment. Kaustia and Knupfer (2008) show that personally experienced returns from past IPOs, not passively observed returns, affect future IPO subscriptions. Malmendier and Nagel (2015) show that individuals overweight inflation experienced during their lifetimes when forming expectations about future inflation.

An important theoretical benchmark model has been developed by Nagel and Xu (2022), in which a representative agent has a fading memory. That is, a data point's influence on beliefs gradually fades over time. While their setup is appropriate for modeling aggregate shocks like a recession, here we focus on precisely the *residual* effects. What we have in mind are memorable events which are idiosyncratic to the trader, like active trading days. These are days (which may differ across traders) when the trader holds a nonzero position in a particular asset and personally experiences the gain or loss. Andersen et al. (2019) show that experiences on such days, as opposed to second-hand experiences, affect an individual's risk-taking behavior. Similarly, Strahilevitz et al. (2011) show that investors are reluctant to repurchase stocks previously sold for a loss. To summarize, *feeling* the gain or loss seems to impact individuals more profoundly than passively observing returns.

Aside from traders failing to recall certain days, we assume that they are Bayesian.[1] They gather what they can remember, and use this data to form expectations about future returns. Idiosyncratic recall generates expectation disagreement amongst traders, which is captured in the cross-sectional distribution over beliefs. And this disagreement is the motive for trade: pessimists who recall bear markets sell, and optimists who recall bull markets buy. We assume that a random subset of traders are active in financial markets during any given period; this effectively "draws" traders independently from

[1] Our model also resembles one of under- and overreaction that can be traced back to Grether (1980). Mullainathan (2002) also makes this connection: imperfect recall explains under- and overreactions with respect to Bayes' Rule.

the cross-section and is, in fact, the only exogenous random variable in the model. This captures the empirically-relevant constraint that traders are not all simultaneously active in financial markets.[2]

The final participant in our economy is a simple linear market maker like the one from Teeple (2022) and Teeple (2023). This market maker raises (lowers) prices when there is excess demand (supply), while providing the required liquidity throughout to ensure that markets clear.[3] The following quote from Bagehot (1971) provides an excellent summary of our model's microstructure:

> It is well known that market makers of all kinds make surprisingly little use of fundamental information. Instead they observe the relative pressure of buy and sell orders and attempt to find a price that equilibrates these pressures. The resulting market price at any point in time is not merely a consensus of the transactors in the marketplace, it is also a consensus of their mistakes. Under the heading of mistakes we may include errors in computation, errors of judgment, factual oversights and errors in the logic of analysis.

An equilibrium is established in the following sense. We begin with a history of prices drawn from a discrete time stochastic process, which forms the basis for traders' beliefs. Based on what they can recall, traders apply Bayes rule. Based on these posteriors, a subset of traders trade; and based on trades, the market maker adjusts the price. Equilibrium prices must have the property that the distribution of current prices matches that of the history of past prices. Intuitively, it is because prices have always been volatile that traders have dispersed beliefs; this, in itself, drives future price volatility. Given that there is no news or other shocks in our model, imperfect memory gives rise to purely non-fundamental volatility.

We show that the discrete time Brownian motion is the unique equilibrium price process when traders have mean-variance utility and remember a fixed, homogeneous number of periods. Importantly, this utility assumption gives our model a linear structure: demands are linear in past price observations due to the mean-variance assumption, and the sum of demands determines future prices due to the linear market maker. Then the equilibrium requirement that past and future prices share a common distribution becomes equivalent to *stability* of the equilibrium price distribution.[4] Within the stable

[2]See Graham et al. (2009) and Richards and Willows (2018).

[3]The classical Walrasian auctioneer appears as a limiting case of our market maker.

[4]A distribution is stable when the sum of i.i.d. random variables have that same distribution.

family, the Normal distribution is the only distribution with a finite variance.

Up to this point, we took the days which a particular trader remembered as given. In our main application, we assume that traders remember the days on which they were actively trading. This gives rise to learning, and the cross-sectional distribution over subjective returns becomes more concentrated over time as traders receive more information. In order to make the model stationary, we assume that traders die at a constant rate and are consequently replaced with newborns who have yet to trade. The steady-state cross-section consists of a mixture of the youth (traders who have traded less, leading to dispersed beliefs) and the elderly (traders who have traded more, leading to concentrated beliefs). This is consistent with survey evidence from Giglio et al. (2021) that older individuals have tighter subjective distributions over future stock returns. Unlike in the baseline setting, where belief dispersion is fixed across the cross-section and equilibrium prices are Normal, in this extension, there is a mixture of belief dispersion across demographics which leads to excess kurtosis in equilibrium prices. The intuition is as follows: extreme prices result from the small chance that all traders trading on a given day are young. To the knowledge of the authors, this mechanism for generating heavy-tailed prices—which we term *youthful volatility*—is novel. We then provide several testable implications of our model: an older population of traders leads to lower price volatility, and a more dispersed population of traders (in terms of age) leads to larger excess kurtosis in prices.

We consider three policies to reduce price variance, and immediately rule one out: a capital gains tax. While it does dampen capital gains, a gains tax also dampens wealth volatility, and the net effect goes in the wrong direction: it *increases* trade volume. We then compare the distributional consequences of the two remaining effective policies: a tax on trade volume and tightened borrowing limits. While transaction taxes truncate trade for the center of the cross-sectional distribution (i.e. older traders), borrowing limits truncate trade for the tails of the distribution (i.e. younger traders).

Related Literature. Recent papers by Malmandier et al. (2020) and Schraeder (2016), which are the closest studies to ours, augment the fading memory assumption with trader heterogeneity using an overlapping generations setup. In Schraeder (2016), the young generation is assumed to be rational while the adult generation is assumed to overweight the previous period's observation. In Malmandier et al. (2020), the young generation reacts more strongly than the old to recent observations, as these make up a larger part of their lifetimes. To compare models, we assume idiosyncratic instead of fading memory; and to compare results, our paper is the only one that connects memory to excess price kurtosis. Also closely related are papers by Collin-Dufresne et al. (2017) and Ehrling et al. (2018). In Collin-Dufresne et al. (2017), younger generations update beliefs more so than older generations with respect to macroeconomic shocks. In Ehrling et al. (2018), the young generation, who act as trend chasers, disregard the price history and lose money to the old generation, who act as contrarians.

The difference between the fading memory setup and ours is the over-weighting of the *previous* period observation versus an *idiosyncratic* observation. First, our assumption is well-supported by studies of human memory. According to the textbook treatment of strength theory from Kahana (2012), each memory has a fixed numerical value representing the degree to which that memory evokes a sense of familiarity. Kahana (2012) comments that it is "reasonable to suppose that items vary in their strength, with some items being stronger than others." In fact, strength theory models typically assume that strength values are drawn from a Gaussian distribution. In our model, we simplify matters by assuming that observations are either remembered or not (we also extend to high versus low memory strength). Second, our idiosyncratic assumption is reinforced by recent economic survey evidence. For example, Dominitz and Manski (2011) find that traders' return expectations are interpersonally variable but intrapersonally stable. Similarly, Giglio et al. (2021) find that beliefs are mostly characterized by large and persistent individual heterogeneity.

The remainder of the paper is organized as follows. Section 2 provides a motivating example demonstrating the Brownian equilibrium. Section 3 models the cross-section when traders remember active trading days, and Section 4 presents the main theoretical results in a generalized setting. Section 5 addresses policy implications, and Section 6 concludes.

2 Motivating Example

Traders. Time is discrete, infinite, and indexed by $t \in \mathbb{Z}$ (not $\mathbb{N}$), and there is a countable number of traders of a single long-lived financial asset that pays no dividends. These individuals should be thought of as retail investors who use technical analysis before placing market orders.[5] Suppose that there is a history of past price realizations − even at the beginning of the model ($t = 0$) − and that these past realizations were drawn from a discrete time Brownian motion with zero drift. That is, $(p_{t+1-s} - p_{t-s})_{s\geq 1}$ were independently drawn from a $N(0, \Sigma^2)$ distribution. In equilibrium, we will require that this price history be consistent with prices generated within the model (for $t \geq 0$). Traders, indexed by i, maximize mean-variance utility

$$u^i(x_t) = (\mathbb{E}^i[p_{t+1}|p_t] - p_t)x_t - \frac{\rho}{2}\mathrm{Var}(p_{t+1} - p_t|p_t)x_t^2 \tag{1}$$

where ρ denotes the risk aversion parameter and x_t denotes the trader's position.[6] Demand is chosen to maximize utility,

$$x^i(p_t) = \frac{\mathbb{E}^i[p_{t+1}|p_t] - p_t}{\rho\Sigma^2}$$

What makes each trader unique is that she believes that price increments $(p_{t+1} - p_t)$ have mean μ^i, because certain past price realizations are more *memorable*. In this simple example, we let each trader remember only one period, and we assume that each trader has a unique memorable time period. Following Nagel and Xu (2022), we assume that traders apply geometric weights to observations $(p_{t+1-s} - p_{t-s})_{s\geq 1}$ when applying a memory-constrained version of Bayes' Rule:

$$f(\mu|p_t - p_{t-1},, ..., p_{t-T+1} - p_{t-T}) \propto \prod_{s=1}^{T} f(p_{t+1-s} - p_{t-s}|\mu)^{\alpha_s^i} \tag{2}$$

where T denotes the length of the price history and μ denotes the true price drift. Note that, when $\alpha_s^i = 1$ for all s, this reduces to Bayesian updating with a flat prior. Instead of the equal weighting case à la Bayes or α_s^i fading over time à la Nagel and Xu (2022), we assume that traders distinctly remember one period i by setting weights

$$\alpha_i^i = \gamma T, \text{ and } \alpha_s^i = \frac{(1-\gamma)T}{T-1} \text{ for } s \neq i$$

where $0 < \gamma \leq 1$ denotes idiosyncratic memory strength. Because there is a one-to-one mapping between traders and memorable periods in this example, we let the trader's index i also denote the period she remembers. Hence α_i^i denotes the weight trader i puts

[5] Empirical evidence suggests that the scarring effects of personal experiences affect not only retail investors, but also highly specialized individuals. See Malmandier and Wachter (2022).

[6] Upcoming assumptions about the *timing* of trades justify the trader's myopic objective.

on her memorable period. Now consider the total weight put on non-memorable periods, $\sum_{s\neq i} \alpha_s^i = (1-\gamma)T$. Adding this to the memorable weight, α_i^i, we see that the weights add up to the number of observations, T, just like in the Bayesian case. However, here we distribute γ proportion of that weight to the memorable period, and the rest on all other periods. Because this remaining weight is distributed *evenly* across periods, in some sense, traders are Bayesian over non-memorable periods. When a memory-constrained Bayesian after T periods forms a posterior mean, it becomes a weighted average between what she remembers and what she does not[7]

$$\mu^i = \gamma(p_i - p_{i-1}) + \frac{1-\gamma}{T-1}\sum_{s\neq i}(p_s - p_{s-1})$$

As we send the number of observations in the history, T, to infinity, by the strong law of large numbers we have that

$$\mu^i \to \gamma(p_i - p_{i-1}) \text{ a.s.}$$

so that non-memorable events effectively wash out.[8] If $\gamma = 0$ and there are no memorable periods, the expression above shows that traders eventually learn that the price drift is truly zero. Since each trader recalls a different time period by assumption, the cross-sectional distribution of beliefs μ^i approaches the $N(0, \gamma^2\Sigma^2)$ distribution as the number of traders, which we denote M, tends to infinity. More formally, by the strong law of large numbers, the empirical distribution

$$F_M(t) = \frac{1}{M}\sum_{i=1}^{M} 1_{\mu^i \leq t}$$

approaches the Normal distribution function for each t almost surely as $M \to \infty$.

Market Maker. If a randomly selected subset of traders of size n trades each period, then a linear market maker

$$p_{t+1} = p_t + c\sum_{i=1}^{n} x^i(p_t) \tag{3}$$

[7] This formula relies on the Normality of price observations.

[8] Unlike in Nagel and Xu (2022), here there is no residual subjective uncertainty. That is, the variance of the posterior tends to zero. Hence, volatility in this model comes from objective uncertainty over prices, not subjective uncertainty over μ.

ensures that prices follow a discrete time Brownian motion.[9] Normality follows from the fact that the sum of i.i.d. Normal random variables in (3) is also Normal. Furthermore, $\Sigma^2 = c\gamma\sqrt{n}/\rho$ is the unique value that makes the variance of historical prices match that of current prices. It is an equilibrium object. When Σ^2 is too large (small), past price volatility makes traders trade too little (much), leading to prices characterized by too little (much) volatility according to (3). Note that the indexes of traders who trade in any given period are exogenous random variables (in fact, the only ones), and are assumed to be independently drawn across periods. This modeling assumption is motivated by recent survey findings that beliefs do not predict *when* investors trade, only the direction and magnitude of trades (see Giglio et al. (2021)). Furthermore, in a dataset of UK investors, Richards and Willows (2018) find that while the top decile of traders trades 69 times per year, the second decile trades 23 times per year and the remaining 80% trade 6 times per year. Similarly, Graham et al. (2009) find that the average number of days between trading is 89 using data from a UBS/Gallup general investor survey. These data support a model where not all traders are simultaneously active in financial markets.[10]

The market maker (3) has the following microfoundation also used in Teeple (2022). We set up the market maker objective function according to the three published objectives of the NYSE's designated market maker: prioritize price discovery, lower volatility, and provide liquidity.

$$\max_{p_{t+1}} \; p_{t+1} \sum_i x^i(p_t) - \frac{1}{2c}(p_{t+1} - p_t)^2$$

Note that the first term above corresponds to price discovery, and is precisely the objective of the classical Walrasian auctioneer.[11] The incentive to raise prices on days with excess demand and lower prices on days with excess supply is evident in this term. In the classical general equilibrium setting, the fixed point of such a maximization problem is studied; here we do not abstract away from preceding dynamics. The second term above corresponds to the second objective of the designated market maker (lowering volatility). The parameter c controls the relative weight allocated between these first two objectives. Such an objective function implies that the market maker *must* provide liquidity (its third goal); she injects extra liquidity into the market by taking the opposite side to

[9] To achieve a geometric Brownian motion, we redefine prices in logs.

[10] That all traders are memory-constrained and trade on random days may seem extreme. We could extend the model to also include Bayesians with access to news who trade on all days. In the spirit of simplicity, we have refrained from this exercise.

[11] This market maker does not rely on limit orders to infer the shape of trader demand functions. Instead, she only observes the value of aggregate demand.

excess demand, effectively clearing markets each period.

This objective function leads to the intuitive rule (3): the market maker maps excess demand into higher prices and excess supply into lower prices at elasticity c. This is consistent with empirical evidence from Chordia et al. (2002), who confirm that excess buy (sell) orders drive up (down) returns, even at lagged time periods. For familiarity, let us map this unconventional microstructure into a more standard one, maintaining the same population of memory-constrained traders. Assume that the same group of n traders trades each period (instead of n randomly drawn), that the asset pays random dividends d_{t+1} (instead of zero dividends), and that the interest rate is $R > 0$ (instead of zero interest rates). Demand would then be given by

$$x^i(p_t) = \frac{\mathbb{E}^i[p_{t+1} + d_{t+1}] - (1+R)p_t}{\rho\Sigma^2} = \frac{\gamma d_i + (1-\gamma)\mathbb{E}[d_{t+1}] - Rp_t}{\rho\Sigma^2}$$

where Σ^2 now denotes volatility of exogenous dividends, and the second equality holds when prices are at a steady state. The fixed point of the market maker's rule, which clears markets, is given by $p = \frac{1}{R}\left(\frac{\gamma}{n}\sum_i d_i + (1-\gamma)\mathbb{E}[d_{t+1}]\right)$. The point of this exercise is to emphasize that the market maker is a sensible one, which simply generates off-equilibrium, or tâtonnement, dynamics in a standard model. However, in our setup, we do away with dividends so that our equilibrium takes on a self-fulfilling flavor and volatility in our setting becomes purely non-fundamental.[12] The random variable that takes the place of exogenous dividends is the identities of active traders.

It is important to point out that the timing assumptions differ significantly from standard market order models like that of Kyle (1985). There, traders face uncertainty about the price at which their orders are executed. Here, traders are able to trade unlimited positions at the fixed price p_t. By ignoring price impact, we are effectively ignoring the bid-ask spread, which simplifies our model and is a reasonable assumption when spreads are small. Teeple (2023) introduces the spread in a similar setting but with no memory constraints. There, a revenue-maximizing market maker is shown to have the same qualitative properties as the one described here.

Because traders trade on random days, it is assumed that they buy (sell) on such days at price p_t then sell (buy) on the very next day at price p_{t+1}.[13] Importantly, they do

[12] 21% of the S&P500 does not pay dividends. However, this number grows to over 50% when considering all US publicly traded stocks.

[13] The second half of this round trip is not included in the excess demand for period $(t+1)$. In Teeple (2023), where the market maker's problem is explicitly modeled, market makers have an incentive to

not hold the asset between two randomly chosen trading days. This has implications for both earnings and inventories of the market maker. In terms of earnings, prices according to (3) rise (fall) when the market maker takes a short (long) position. In other words, the market maker loses money each period. This observation is consistent with empirical evidence from Sofianos (1995) that market makers incur positioning losses on their inventory, which are compensated by revenues from spreads (not modeled here). In terms of inventories, when the market maker takes a long (short) position, she buys (sells) at price p_t then sells (buys) at price p_{t+1}. Hence her inventories in any given period are $-\sum_i x^i(p_t)$, but importantly, she does not accumulate inventories across periods. This is observationally consistent with mean-reversion theories of market maker inventories. For example, see Hasbrouck and Sofianos (1993).

Equilibrium Definition. Here we define our equilibrium concept. The price history is initially drawn from some distribution. Based on their imperfect memory, traders disagree and trade; and based on the market maker, prices adjust. We require that the distribution of prices generated by the model be precisely the one from which the price history was initially drawn. Formally, an equilibrium is a price history $\mathcal{H}_t$, trader beliefs μ^i and demands $x^i(p_t)$, and a price distribution F_t satisfying:

(a) Given the recollection $\mathcal{H}_t^i \subset \mathcal{H}_t$, traders form beliefs μ^i according to (2).

(b) Given beliefs μ^i and conjecture Σ_t^2, $x^i(p_t)$ maximizes (1) if they are in the active group, $i \in \mathcal{A}_t$. Otherwise, $x^i(p_t) = 0$.

(c) The market maker (3) generates prices drawn from F_t, where F_t is consistent with history $\mathcal{H}_t$ and $\mathrm{Var}_{F_t}(p_{t+1} - p_t) = \Sigma_t^2$.

Furthermore, we focus our attention on steady-state equilibria where $F_t = F$ ($\Sigma_t^2 = \Sigma^2$) for all t. While markets do not clear in the traditional sense (they do clear when the market maker's position is accounted for), we maintain the standard assumption that markets are competitive. With few traders, each has a non-negligible own price impact. Optimizing over this price impact, even with memory constraints, could allow traders to do better than the demand functions described here.

move prices based on initial trades to elicit trader participation. However, this incentive disappears when traders unwind positions.

Uniqueness. Our next question is whether the Brownian motion is the unique equilibrium price process in this baseline setting. Interestingly, the drift of the Brownian motion (or lack thereof) is unique. To understand why, consider the following extension of the model. Instead of assuming that the asset is in net zero supply, say that the asset has an exogenous, deterministic supply $S > 0$. The natural extension of the market maker's objective is

$$\max_{p_{t+1}} p_{t+1} \left(\sum_i x^i(p_t) - S \right) - \frac{1}{2c}(p_{t+1} - p_t)^2$$

which yields the following generalization of (3)

$$p_{t+1} = p_t + c \left(\sum_i x^i(p_t) - S \right) \tag{4}$$

so that, just like before, prices rise (fall) when excess demand is positive (negative). Assuming that the price history has a drift of μ, beliefs converge almost surely to

$$\mu^i = \gamma(p_i - p_{i-1}) + (1 - \gamma)\mu$$

and the cross-section of beliefs μ^i across the population then approaches $N(\mu, \gamma^2\Sigma^2)$. The equilibrium drift μ can then be found by taking the expectation of the market maker's rule (4). With some algebra, this condition reduces to

$$\mu = \frac{cS}{cn/(\rho\Sigma^2) - 1} \tag{5}$$

The intuition is that positive drift generates excess demand; this, in turn, requires a positive asset supply.[14] Importantly, $\mu = 0$ when $S = 0$, establishing the uniqueness of zero drift in the baseline setting.

To summarize so far, both variance and drift are unique. However, we can say more by making use of the following well-known statistical fact from Durrett (2017): if a linear combination of two independent random variables with some distribution has that same distribution, it is said to be stable. Let us consider the implications for our model, where demand is linear in past prices due to the mean-variance assumption, and prices

[14] $S > 0$ does not change the variance calculation, hence $\Sigma^2 = c\gamma\sqrt{n}/\rho$ like before. Then $n > 1$ ensures that the denominator of (5) is positive.

are linear in demand due to the nature of the market maker objective. Altogether,

$$p_{t+1} - p_t = \underbrace{c \sum_i \frac{\gamma \overbrace{(p_i - p_{i-1})}^{\text{Linear in prices}}}{\rho \Sigma^2}}_{\text{Linear in demand}}$$

Equilibrium requires that the past price distribution match the current one, so the only candidate for the equilibrium distribution is the stable-α distribution. Furthermore, there is only one distribution in the stable-α family that has a finite variance: the Normal distribution. Hence the Brownian motion (with drift and variance pinned down) is not just an equilibrium distribution. It is the *unique* equilibrium distribution.

3 Active Trading Days

Up until this point, we have taken the days which each trader remembers as given. The setting we consider now is that of Section 2 except for the following change. When the random subset of size n trade, they remember that period's return (after they trade and the return is realized). Empirically, this is consistent with evidence that traders disproportionately use active trading days (more than inactive ones) to form beliefs about future returns (see Andersen et al. (2019) and Strahilevitz et al. (2011)). Theoretically, this is consistent with the context-dependent memory model of Bordalo et al. (2020), where the context here would be whether or not the trader was actively trading.[15]

Expanded Memory. The first difficulty with this setup is characterizing the Bayesian posterior mean when a trader has multiple memorable periods. Despite the well-known difficulties in deriving a closed-form solution for the posterior mean of arbitrary distributions (equilibrium prices may not be Normal), our first lemma addresses this issue. All proofs are in Appendix A.

Lemma 1. ***(Posterior Consistency)*** *Say that traders calculate their posterior according to memory-constrained Bayes' Rule* (2)*, with weights given by*

$$\alpha_t^i = \begin{cases} \gamma T/|\mathbb{K}^i| \text{ for } t \in \mathbb{K}^i \\ (1-\gamma)T/(T - |\mathbb{K}^i|) \text{ otherwise} \end{cases} \tag{6}$$

[15]See also Wachter and Kahana (2024).

As the history $T \to \infty$, their posterior mean approaches

$$\mu^i = \frac{\gamma}{|\mathbb{K}^i|} \sum_{t \in \mathbb{K}^i} (p_t - p_{t-1}) + (1-\gamma)\mu$$

irrespective of the price distribution.

Note that μ denotes the mean of the price distribution, γ denotes idiosyncratic memory strength, and $\mathbb{K}^i$ denotes the memorable set for trader i. When this set consists of only one period, the weights in (6) collapse precisely to those described in Section 2. Now consider $|\mathbb{K}^i| > 1$. Just like in Section 2, the total weight adds up to T, with γ proportion of that weight distributed evenly across memorable periods. The way that we circumvent the usual algebraic difficulties in dealing with non-Normal posteriors is by sending $T \to \infty$. Asymptotically, there are results from probability theory that guarantee consistency of the Bayesian posterior mean. Weights in (6) are *as-if* the trader not only observes an infinite number of independent non-memorable prices, but also an infinite number of repeated observations of $(p_t - p_{t-1})$ for each $t \in \mathbb{K}^i$. As $T \to \infty$, they become sure that their posterior mean equals the μ^i defined in Lemma 1 (i.e. their posterior variance tends to zero).

To get a sense of an economy with expanded memory, Proposition 1 extends the economy from Section 2 with multiple memorable periods. What we do *not* allow for (yet) is overlapping memory; each trader still remembers a subset of the price history disjoint from any other trader. What we *do* allow for here is heterogeneity in memory length; one trader may remember two periods, while another may remember three. With heterogeneity in memory length, equilibrium prices need not be Normal. However, in one special case we recover the Brownian equilibrium.[16]

Proposition 1. ***(Expanded Memory)*** *Say that each trader remembers k^i disjoint periods for $k^i \in \{1, ..., K\}$. Equilibrium prices converge in distribution to a discrete time Brownian motion as $n \to \infty$ and $T \to \infty$.*[17]

First consider the homogeneous case, where all traders remember the same number, k, of disjoint periods. If multiple events are memorable, good days and bad days begin to cancel each other out within the memorable set. Hence, traders with a larger set

[16] Note that we exclude the case where indexes of memorable prices are measurable with respect to the realization. This, for example, rules out extreme prices being remembered by more traders.

[17] Note that sending the history $T \to \infty$ implies that the number of traders $M \to \infty$ at the same rate.

of memorable events hold less dispersed, or tighter, beliefs. This, in itself, does not destroy Normality and simply lowers the equilibrium variance, Σ^2. See Figure 1 for the cross-sectional distribution over μ^i for traders who remember 1, 2, and 3 disjoint periods (memory parameter γ is set to one).

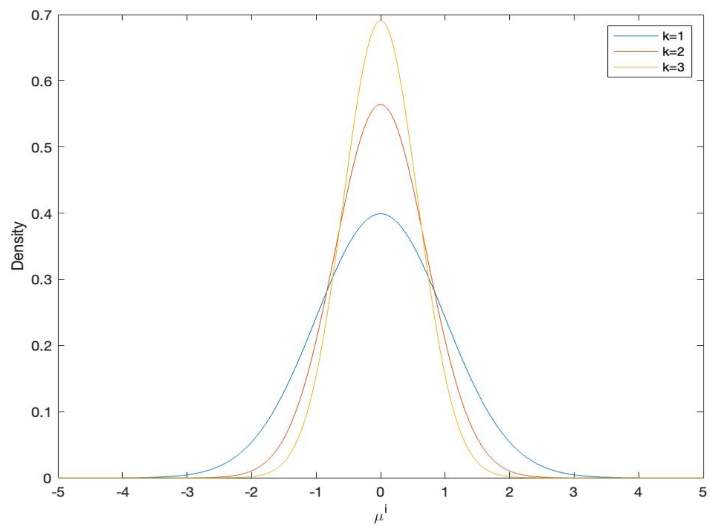

Figure 1: Cross-Section with Expanded Memory

Now consider adding heterogeneity in memory length. The cross-sectional distribution of beliefs is no longer Normal, because the mixture of Normal distributions need not be Normal. That is, when drawing a trader from the cross-section, there is some chance she is inexperienced (few memorable periods, with dispersed beliefs) and some chance she is experienced (many memorable periods, with concentrated beliefs). So while the cross-section need not be Normal, the Lindeberg-Lévy theorem applies when n is large and the market maker ensures that prices continue to follow a Brownian motion. In contrast to Section 2, sending $n \to \infty$ is with significant loss of generality.[18]

Compared to the baseline case of Section 2, this last discussion suggests that there will be two notable changes in the active trading days setup. First, the equilibrium variance will be reduced because the older generations have less dispersed beliefs. Second, equilibrium prices need not be Normally distributed due to the mixing of younger and older generations (when n is finite).

[18]As n grows, so does the equilibrium variance. To deal with this fact, we normalize the equilibrium variance to one and find a market maker constant, c, consistent with that normalization.

Cross-Section. When traders remember active trading days, the cross-section of traders would continuously learn. To regain stationarity in our model, we assume that $\mathcal{D} > 0$ traders die each period and are consequently replaced by $\mathcal{D}$ new traders.[19] The steady-state cross-section of traders is characterized by a sequence $(\mathcal{P}_j)_{j=0}^{\infty}$, which represents the proportion of the population that remembers j days. We first analyze the case where the number of traders, M, is finite, then take limits. When M is finite but large (in particular, $M > n + \mathcal{D}$), the law of motion for $\mathcal{P}_j$ is given by

$$\mathcal{P}_j^{t+1} = \begin{cases} \mathcal{P}_j^t \left(1 - \frac{\mathcal{D}}{M} - \frac{n}{M}\right) + \frac{\mathcal{D}}{M}, \text{ if } j = 0 \\ \mathcal{P}_j^t \left(1 - \frac{\mathcal{D}}{M} - \frac{n}{M}\right) + \mathcal{P}_{j-1}^t \frac{n}{M}, \text{ otherwise} \end{cases}$$

When $j = 0$, the proportion decreases by a factor $\frac{\mathcal{D}}{M}$ due to death, and by another factor $\frac{n}{M}$ due to active trading, learning, and traders consequently leaving group $\mathcal{P}_0^t$ for group $\mathcal{P}_1^{t+1}$. The proportion increases by the amount $\frac{\mathcal{D}}{M}$ due to the introduction of new traders. When $j > 0$, the proportion decreases for the exact same reasons. The proportion increases by the amount $\mathcal{P}_{j-1}^t \frac{n}{M}$ due to active trading, learning, and traders consequently entering group $\mathcal{P}_j^{t+1}$ from group $\mathcal{P}_{j-1}^t$. The steady state of such a system is geometric:

$$\mathcal{P} = \left(\frac{\mathcal{D}}{\mathcal{D}+n}, \frac{\mathcal{D}n}{(\mathcal{D}+n)^2}, \frac{\mathcal{D}n^2}{(\mathcal{D}+n)^3}, \ldots\right) \tag{7}$$

and, importantly, is independent of M. As an example, consider the special case when $\mathcal{D} = n$ so that traders die at the same rate that they learn. Then the steady state is $\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots\right)$.

The cross section (7) forces us to take a stance on trader beliefs in the absence of any memorable period. We have yet to encounter this issue because each trader remembered one day in Section 2 and one or more days in Proposition 1. Our approach here is to extend the weights in (6) so that, when $|\mathbb{K}^i| = 0$, $\alpha_t^i = 1$ for all t. In other words, they are Bayesian before their memorable period arrives. Note that this does not preclude newborns from observing the price history; it simply means that no period from that history stands out, or is memorable, to them.

Overlapping Memory. The active trading days setup introduces an additional complication in the form of overlapping memory. That is, when $n > 1$ traders remember an

[19]Due to an absence of wealth effects in the model, we conjecture that our results would not change qualitatively under an alternate survival story where traders die off when wealth hits zero.

active trading day, their beliefs are no longer mutually independent.[20] In general, this problem of overlapping beliefs makes the model intractable. However, in the special case when traders remember past active trading days, the degree of overlap is "limited" and we have the following lemma. It states that, with enough traders, this limited overlap does not preclude the law of large numbers from holding and, consequently, the cross-section of trader beliefs converges in distribution much like in Section 2.

Lemma 2. ***(Limited Dependence)*** *Assume that the maximum number of periods remembered in the economy is K. Then, as the number of traders $M \to \infty$, the empirical distribution over beliefs $F_M(t)$ converges in probability to its expected value for each t.*

A natural next question is: What is the expected value of $F_M(t)$? However, we hold off on this investigation until Section 4. Instead, the point of Lemma 2 is to emphasize that the cross-section converges to some distribution despite the presence of traders whose beliefs overlap (albeit in a limited way).

For intuition, consider whether the weak law of large numbers holds (instead of whether the empirical distribution converges). By Chebyshev's Inequality,

$$Pr(|\overline{\mu}_M - \mu| \geq \varepsilon) \leq \frac{\text{Var}(\overline{\mu}_M)}{\varepsilon^2} \tag{8}$$

where $\overline{\mu}_M$ denotes the average belief in the cross-section of size M, and μ denotes the actual price drift. In the standard i.i.d. case, the variance term above collapses into something that is $\mathcal{O}(1/M)$ and so the weak law holds. However, here, posterior means μ^i are neither independent nor identically distributed. Although the younger (older) generation has more dispersed (concentrated) beliefs, this is not problematic because there exists a most-dispersed belief, which serves as an upper bound for the variance of any belief in (8). And although there is some overlap in beliefs, each trader's belief can only overlap with $K(n-1)$ others' because they trade on K days where there were $(n-1)$ other traders on each day. Hence the variance of the average belief in (8) decomposes into some (bounded) variances and (a finite number of) covariances, which all disappear as $M \to \infty$. This same idea is employed in the actual proof, although the empirical distribution, $\frac{1}{M}\sum_{i=1}^{M} 1_{\mu^i \leq t}$, is considered.

The key to the lemma is that we truncate the number of remembered periods at a large number K. That is, we ignore traders who remember an arbitrarily large number of

[20]Note that any overlap in beliefs was explicitly assumed away in Section 2 and Proposition 1, but is unavoidable in the active trading days setup.

past periods. This is justified for two reasons. First, such traders do not contribute to our economy in the sense that they trade close to zero. Second, we focus on cross-sectional distributions (7) where such traders have a small mass. Hence, the economy considered in Lemma 2 will become arbitrarily close to the true cross-section when K is large.

4 Equilibrium Prices

Section 2 was the simplest possible example in our setting, and Section 3 suggested a viable microfoundation for heterogeneous memory. To nest both cases, we now consider a general steady-state sequence $(\mathcal{P}_j)_{j=0}^K$, representing proportions of the population remembering j days. If this sequence were geometric, we recover the special case of Section 3; if it were degenerate, we recover the special case of Section 2. We emphasize that, while this setup may seem very general, we can only guarantee that upcoming results apply to the following two settings:

1. An extension of Section 2 where each trader remembers k^i disjoint periods for $k^i \in \{1, ..., K\}$.
2. The active trading days setting from Section 3, so that Lemma 2 continues to apply.[21]

As discussed in Section 3, it is the problematic overlap in beliefs which necessitates these conditions above. These two requirements are sufficient but not necessary for upcoming results.[22]

Next, we recast the failure of Normality from Proposition 1 (when $n < \infty$) not as a negative one, but as a positive one (i.e. interesting properties of equilibrium prices). In particular, we show that mixing distributions with different variances in the cross-section results in *excess kurtosis* in equilibrium prices.

Theorem 1. ***(Heavy Tails)*** *Let μ^4 denote the fourth (central) moment of prices. Then*

1. $\mathbb{E}[p_{t+1} - p_t | p_t] = 0$.

[21] In Section 3 the cross-section $(\mathcal{P}_j)_{j=0}^K$ depended on n, while here it does not. This is not critical for our results.

[22] For example, consider Section 2 but with heterogeneity in memory strength γ^i or risk aversion ρ^i. Upcoming results apply in these settings as well.

2. *Assume that the cross-section* $(\mathcal{P}_j)_{j=0}^K$ *is nondegenerate and that* $n > N(\mathcal{P})$*. Then* $\mu^4/(\Sigma^2)^2 \geq 3$*, and the inequality holds with equality if and only if* $n \to \infty$*.*

Analogous to the no drift discussion in Section 2, equilibrium prices generally are a martingale. Importantly, prices are heavy-tailed and the effect disappears only when $n \to \infty$ and central limit theorems apply. Figure 2 shows how mixing leads to heavy tails, using two Normal distributions with mean zero and differing variances ($\sigma^2 = 1$ and $\sigma^2 = 2$). For visual clarity, the mixture fractions here are set to $\frac{1}{2}$ and $\frac{1}{2}$. In words, heavy tails result from a small chance of the trader being drawn from the high variance distribution. However, this argument only guarantees that the cross-section has heavy tails, not prices. Prices arise from the sum of n traders' demands, and summing random variables reduces excess kurtosis due to the central limit theorem. Theorem 1 Condition 2 can be understood as characterizing the net effect between these two competing forces (mixing and summing).

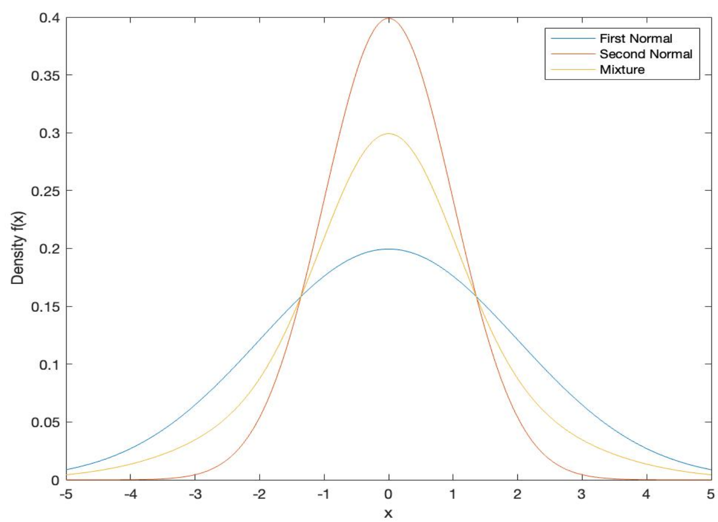

Figure 2: Heavy-Tailed Mixture Distributions

To get a sense of how the proof of Theorem 1 works, consider the simpler case when $n \to \infty$. Calculating the fourth moment of the market maker's rule (3) and dividing by $(\Sigma^2)^2$,

$$\frac{\mu^4}{(\Sigma^2)^2} = \frac{c^4(n\mathbb{E}[x^i(p_t)^4] + 3n(n-1)\mathbb{E}[x^i(p_t)^2]^2)}{(\Sigma^2)^2} \tag{9}$$

which holds true when demands are independently drawn from the cross-section and have mean zero. Note the factor of three in the numerator of (9), which arises mathematically when expanding the expression $\mathbb{E}[\{\sum_i x^i(p_t)\}^4]$. By taking the variance of the market maker's rule (3), we have that $\Sigma^2 = c^2 n \mathbb{E}[x^i(p_t)^2]$. Plugging this into the denominator of (9), only $\mathcal{O}(n^2)$ terms in the numerator of (9) survive, and hence the price kurtosis tends to three as n tends to infinity.

Interestingly, the heavy tail result can fail for small values of n below $N(\mathcal{P})$. This notation is meant to emphasize that the cutoff value of n depends only on the cross-sectional distribution $\mathcal{P} \equiv (\mathcal{P}_j)_{j=0}^K$, not any other parameters. While so far we have emphasized the latter, there are two reasons why equilibrium prices might have heavy tails. First, there are historical effects in the sense that a heavy-tailed history tends to generate heavy-tailed prices (or vice versa). Second, due to the mixing of different generations in the cross-section, there is a small chance of a youthful draw. This breakdown can be explicitly seen after manipulating condition (9):

$$\text{Excess kurtosis} = \frac{N(\mathcal{P}) \cdot (\text{Excess kurtosis}) + M(\mathcal{P}) \cdot (\text{Variance of cross-section})}{n}$$

where $N(\mathcal{P}) > 0$ and $M(\mathcal{P}) > 0$, whose explicit expressions are relegated to Appendix A. For excess kurtosis to inherit the sign of the variance of the cross-section (positive when the cross-section is nondegenerate), we require that $n > N(\mathcal{P})$. In Figure 3, we plot this cutoff N but not as a function of the distribution $\mathcal{P}$. Instead, we take all distributions with support of size K, then find a sufficient cutoff for that family of distributions. Mathematically, $N(K) = \max_{\text{supp}(\mathcal{P}) \leq K} N(\mathcal{P})$. We emphasize that this cutoff is a small number (hence $n > N(\mathcal{P})$ is not restrictive) even when traders are allowed to recall $K = 50$ memorable periods.

The explanation for heavy tails proposed here is distinct from those previously proposed in the literature. For example, Gabaix et al. (2006) cite heavy tails in wealth, Cont and Bouchaud (2000) cite imitation among traders, and Thurner et al. (2012) cite leverage effects. Here, heavy tails come from the mixing of random variables with different variances. Extreme price events occur when *young* traders, understood as those who have experienced only a few past trades, all trade in a given period. This is consistent with survey evidence in Giglio et al. (2021), who find that older individuals' subjective distributions over future stock returns have lower standard deviations (than those of young individuals) and, furthermore, they assign smaller probabilities to extreme events such as large stock market declines.

Our next proposition considers comparative statics of equilibrium price moments

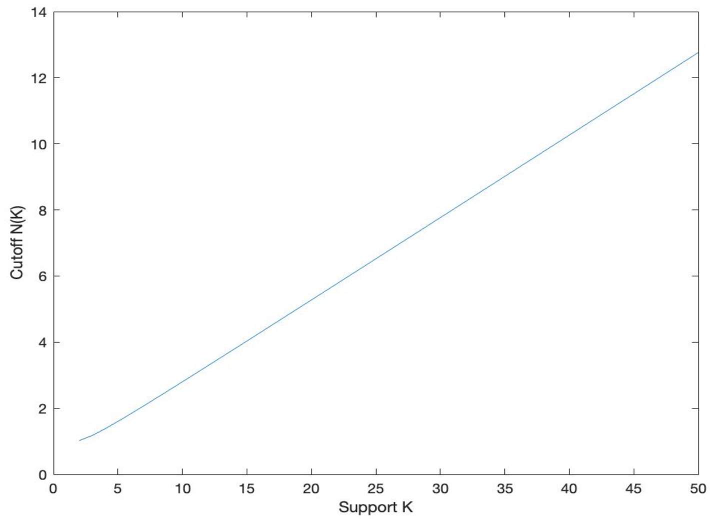

Figure 3: Cutoff Value $N(K)$

with respect to the cross-sectional distribution $(\mathcal{P}_j)_{j=0}^{K}$. As Proposition 1 suggested, increasing the "age" of the population decreases the equilibrium variance. And as Theorem 1 alluded to, increasing the dispersion of age within the population increases excess kurtosis.

Proposition 2. ***(Comparative Statics)*** *Let Σ^2 and μ^4 denote the variance and fourth (central) moment of prices, respectively. Then*

1. *Σ^2 is increasing in $\mathbb{E}_{\mathcal{P}}[1/X]$.*
2. *$\mu^4/(\Sigma^2)^2$ is increasing in $Var_{\mathcal{P}}(1/X)$ when $n > N(\mathcal{P})$.*

First let us understand the random variable, $1/X$. X here denotes age, or more specifically, when $X = k$, the proportion $\mathcal{P}_k$ of the population has k memorable periods. Hence $1/X$ is the inverse of age, so when this object is larger (smaller), the trader trades more (less). Let us refer to $1/X$ as the trader's *youthfulness*. Condition 1 of Proposition 2 then says that more youthful economies are characterized by more volatility. And as intuition suggests, excess kurtosis in Condition 2 is increasing in the variance of the cross-section. Proposition 2 presents empirically testable implications of our theory.

5 Policy

We consider three policies to reduce price variance in the general model from Section 4: a transaction tax, a capital gains tax, and tightened borrowing constraints.

Transaction Tax. The trader's wealth w_{t+1} with a transaction tax $r > 0$ is

$$w_{t+1} = (p_{t+1} - p_t)x_t - r|x_t|$$

where $r = 0$ corresponds to the baseline setting. In terms of implications for trade, the tax creates an inaction zone, where traders for whom $\mathbb{E}^i[p_{t+1}|p_t] \approx p_t$ choose not to trade. This can be seen formally in the new demand function:

$$x^i(p_t) = \begin{cases} \frac{\mathbb{E}^i[p_{t+1}|p_t] - p_t - r}{\rho\Sigma^2}, & \text{if } \mathbb{E}^i[p_{t+1}|p_t] - p_t > r \\ \frac{\mathbb{E}^i[p_{t+1}|p_t] - p_t + r}{\rho\Sigma^2}, & \text{if } \mathbb{E}^i[p_{t+1}|p_t] - p_t < -r \\ 0, & \text{otherwise} \end{cases}$$

Traders who remain active in financial markets trade smaller quantities. While it may seem immediate that the equilibrium price variance is reduced, we must be mindful of equilibrium effects. Lowered price volatility can make financial markets look more attractive, thereby increasing trade. Note that the upcoming Proposition 3 describes a new equilibrium, presumably a sufficient amount of time after the policy change has been enacted. This qualifier is needed, because the price *history* must also reflect the effects of the policy.[23]

Proposition 3. ***(Transaction Tax)*** $\frac{\partial\Sigma}{\partial r} < 0$ *for small values of* r*.*

Consider the cross-section of demands. The tails have been shifted inwards toward the origin, because potential market gains − in either direction − are reduced by the amount of the tax. The middle range of the distribution has been removed altogether, because the cost of trading entirely outweighs the benefits for such traders. See Figure 4 for an illustration of this truncated cross-section of demands $x^i(p_t)$. With this lowered dispersion of demands, we find that equilibrium price variance is reduced when transaction taxes are reasonably small.

[23]The same comment applies to Proposition 4.

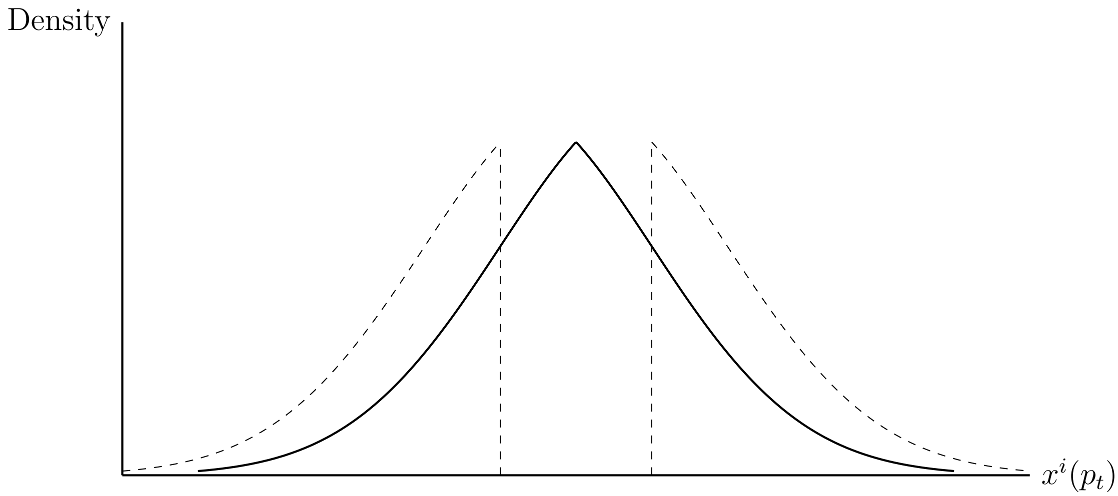

Figure 4: Truncated Cross-Section of Demands

Capital Gains Tax. We model the capital gains tax with an adjusted budget set:

$$w_{t+1} = (1-\tau)(p_{t+1} - p_t)x_t$$

where the tax is in the range $0 < \tau < 1$. Again we must consider the equilibrium effects of the tax. Price variance must solve a fixed point condition

$$\Sigma^2 = c^2 n \sum_{k=1}^{K} \mathcal{P}_k \mathrm{Var}\left[\frac{(1-\tau)\gamma \sum_{t \in \mathbb{K}^i}(p_t - p_{t-1})/k}{\rho(1-\tau)^2 \Sigma^2}\right]$$

where the term inside the brackets is the capital gains-adjusted demand. This condition is solved by $\Sigma^2 = \frac{c\gamma}{\rho(1-\tau)}\sqrt{n\sum_{k=1}^{K}\mathcal{P}_k/k}$, so that taxes are actually *counterproductive*.[24] To understand why, consider two competing forces: dampened capital gains versus dampened wealth variance. The calculation above shows that the latter force outweighs the former, leading to more trade in equilibrium and (unintentionally) higher price volatility. We conclude that a capital gains tax is an ineffective policy tool in this setting.

Borrowing Constraints. The final policy that we consider is a borrowing constraint. While mean-variance utility was previously used without formal derivation, here we begin

[24]From the proof of Theorem 1, without the tax, $\Sigma^2 = \frac{c\gamma}{\rho}\sqrt{n\sum_{k=1}^{K}\mathcal{P}_k/k}$.

with the trader's budget set today:

$$x_t p_t + a_t = 0$$

where a_t denotes the quantity saved (or borrowed) of a riskless asset. Notice that we have normalized today's wealth (RHS of equation above) to zero, which would be with loss of generality if we were to consider an infinite horizon objective. The budget set tomorrow is:

$$w_{t+1} = x_t p_{t+1} + a_t$$

where w_{t+1} denotes wealth. Plugging the first budget set into the second, we recover the objective of the trader from Section 2. In addition to this standard budget constraint, we introduce two additional constraints

$$a_t \geq -p_t \bar{b}, \quad x_t \geq -\bar{b} \tag{10}$$

where $\bar{b} > 0$ denotes the real borrowing constraint. When traders take a long position on the risky asset, $x_t > 0$, the first constraint in (10) applies and represents a borrowing constraint on the riskless asset. When traders take a short position on the risky asset, $x_t < 0$, the second constraint in (10) applies and limits the quantity borrowed of the risky asset. Substituting in today's budget constraint, the two new constraints reduce to:

$$|x_t| \leq \bar{b}$$

This effectively truncates demand tails. Demand functions become

$$x^i(p_t) = \begin{cases} \bar{b}, & \text{if } \frac{\mathbb{E}^i[p_{t+1}|p_t]-p_t}{\rho\Sigma^2} > \bar{b} \\ -\bar{b}, & \text{if } \frac{\mathbb{E}^i[p_{t+1}|p_t]-p_t}{\rho\Sigma^2} < -\bar{b} \\ \frac{\mathbb{E}^i[p_{t+1}|p_t]-p_t}{\rho\Sigma^2}, & \text{otherwise} \end{cases}$$

Proposition 4. ***(Borrowing Constraints)*** $\frac{\partial \Sigma}{\partial \bar{b}} > 0$ *for large values of* $\bar{b}$*.*

As we tighten the constraint (decrease $\bar{b}$), price variance is reduced as desired. And the result holds when $\bar{b}$ is large (i.e. the borrowing constraint is relatively loose). Lastly, we discuss differences between the two effective policies in terms of their distributional effects. In the case of transaction taxes, traders with moderate beliefs trade zero, and those with tail beliefs reduce, or shift, their amount of trade by the quantity $\frac{r}{\rho \Sigma^2}$. In the case of borrowing limits, demand is truncated for traders with extreme beliefs. Hence transaction taxes effectively target the *center* of the distribution, while borrowing limits effectively target the *tails* of the distribution. In the context of the active trading days setting, this means that transaction taxes (borrowing limits) are more likely to target older (younger) individuals.

While both transaction taxes and borrowing limits are effective policy tools, capital gains taxes are not. The two effective policies truncate expected gains, without directly affecting the wealth variance (there are indirect effects that work through prices, which were the focus of Propositions 3 and 4). In contrast, capital gains taxes are multiplicative. This multiplicative property not only leads to dampened capital gains, but also directly dampens wealth volatility; the net effect is (inadvertent) increased trade.

6 Conclusion

We have proposed a simple mechanism – idiosyncratic memory – to explain non-fundamental volatility and heavy tails in financial markets. When some traders recall bear markets and others recall bull markets, this creates belief dispersion in the cross-section and, hence, a motive for trade. The first key takeaway from the paper is that we live, in some sense, in a bad equilibrium: past volatility is what guarantees enough belief dispersion to generate current volatility. The Brownian motion is the unique equilibrium price process, and resulting volatility can be categorized as purely non-fundamental.

We then formally modeled one reason why traders might have idiosyncratic memory: remembering active trading days. The cross-section of traders consists of the young (those who remember few trading days and have dispersed beliefs) and the old (those who remember many trading days and have concentrated beliefs). The mixture distribution over such demographics is not Normal, which leads to equilibrium prices that are not Normal. This is the second key takeaway from the paper: we have identified a novel mechanism for generating prices with excess kurtosis. Extreme prices result from the small chance that all traders trading on a given day are young.

Appendix A

Proof of Lemma 1. The posterior according to memory-constrained Bayes' Rule (2) and weights (6) is given by

$$f(\mu|p_t - p_{t-1}, , ..., p_{t-T+1} - p_{t-T}) \propto \underbrace{\prod_{t\in\mathbb{K}^i} f(p_t - p_{t-1}|\mu)^{\gamma T/|\mathbb{K}^i|}}_{\text{Memorable}} \underbrace{\prod_{t\notin\mathbb{K}^i} f(p_t - p_{t-1}|\mu)^{(1-\gamma)T/(T-|\mathbb{K}^i|)}}_{\text{Non-memorable}}$$

As $T \to \infty$, this posterior approaches the one that would have been generated from i.i.d. draws from a distribution − call it G − where, with probability $\gamma/|\mathbb{K}^i|$, $(p_t - p_{t-1})$ is drawn for each $t \in \mathbb{K}^i$; and with probability $(1-\gamma)$, prices are drawn independently from the price distribution. To see why, first consider the log of the non-memorable terms,

$$\frac{(1-\gamma)T}{T - |\mathbb{K}^i|} \sum_{t\notin\mathbb{K}^i} ln[f(p_t - p_{t-1}|\mu)]$$

and divide by (the log of) the product of $(1-\gamma)T$ densities evaluated at independent prices.[25] We claim that, as $T \to \infty$, this ratio satisfies

$$\frac{\frac{(1-\gamma)T}{T-|\mathbb{K}^i|}\sum_{t\notin\mathbb{K}^i} ln[f(p_t - p_{t-1}|\mu)]}{\sum_{t=1}^{(1-\gamma)T} ln[f(p_t - p_{t-1}|\mu)]} = \frac{\frac{1}{T-|\mathbb{K}^i|}\sum_{t\notin\mathbb{K}^i} ln[f(p_t - p_{t-1}|\mu)]}{\frac{1}{(1-\gamma)T}\sum_{t=1}^{(1-\gamma)T} ln[f(p_t - p_{t-1}|\mu)]} \to 1 \text{ a.s.}$$

where the equality follows from algebra and the convergence follows from the strong law of large numbers. This line of reasoning justifies replacing non-memorable terms in the posterior with the product of $(1-\gamma)T$ densities evaluated at independent prices. Next consider the log of each memorable term,

$$\frac{\gamma T}{|\mathbb{K}^i|} ln[f(p_t - p_{t-1}|\mu)], \text{ for } t \in \mathbb{K}^i$$

which trivially equals (the log of) the product of $\frac{\gamma T}{|\mathbb{K}^i|}$ densities evaluated at the same memorable price, $(p_t - p_{t-1})$ for $t \in \mathbb{K}^i$. Putting arguments altogether, the frequencies $(1-\gamma)$ and $\gamma/|\mathbb{K}^i|$ observed in the posterior precisely match those of the proposed mixture G. Mixture distributions have an expected value that is equal to the mixture of the

[25] If $(1-\gamma)T$ is not an integer, the same argument applies to floor$[(1-\gamma)T]$. This comment also applies to the number of densities associated with memorable prices, $\frac{\gamma T}{|\mathbb{K}^i|}$.

expected values:

$$\mathbb{E}_G[\mu] = \frac{\gamma}{|\mathbb{K}^i|} \sum_{t \in \mathbb{K}^i} (p_t - p_{t-1}) + (1-\gamma)\mu$$

By the theorem of Doob (1949), the posterior mean μ^i converges to $\mathbb{E}_G[\mu]$ almost surely. $\square$

Proof of Proposition 1. As n grows, so does the equilibrium variance. As mentioned in the main text, we normalize the equilibrium variance to one, and find a market maker constant, c, consistent with that normalization. Let $\mathbb{K}^i$ denote the set of memorable periods for trader i, so $|\mathbb{K}^i| = k^i$. By Lemma 1, as $T \to \infty$, we have that

$$\mu^i \to \frac{\gamma}{k^i} \sum_{t \in \mathbb{K}^i} (p_t - p_{t-1}) \text{ a.s.}$$

In equilibrium we will confirm that, indeed, the price drift $\mu = 0$. Let M_k denote the set of traders who remember k periods (so $\sum_k |M_k| = M$), and let $\mathcal{P}_k$ denote their proportion in the population (so $\frac{|M_k|}{M} = \mathcal{P}_k$). The empirical distribution of beliefs can be written

$$F_M(t) = \frac{1}{M} \sum_{i=1}^{M} 1_{\mu^i \leq t} = \sum_{k=1}^{K} \frac{\mathcal{P}_k}{|M_k|} \sum_{i \in M_k} 1_{\mu^i \leq t}$$

For each k, the empirical distribution approaches $N(0, \gamma^2/k)$ almost surely. The entire empirical distribution is a mixture of these Normal distributions, weighted by $\mathcal{P}_k$, which need not be Normal. However, a direct calculation yields a mean of zero and variance $\sum_{k=1}^{K} \mathcal{P}_k \frac{\gamma^2}{k}$. If a randomly selected subset of size n trades each period, then the Lindeberg-Lévy theorem can be applied to the market maker (3) as $n \to \infty$, ensuring that prices follow a discrete time Brownian motion with mean zero and variance one when $c = \frac{\rho}{\gamma\sqrt{n \sum_k \mathcal{P}_k / k}}$. $\square$

Proof of Lemma 2. By Chebyshev's Inequality, for each t

$$Pr\left(\left| \frac{1}{M} \sum_{i=1}^{M} 1_{\mu^i \leq t} - F(t) \right| \geq \varepsilon \right) \leq \frac{\text{Var}(\sum_{i=1}^{M} 1_{\mu^i \leq t})}{M^2 \varepsilon^2} \tag{11}$$

where $F(t)$ denotes the expected value of the empirical distribution. But now consider the variance of an arbitrary $1_{\mu^i \leq t}$ in the cross section. It is

$$\mathbb{E}[1^2_{\mu^i \leq t}] - \mathbb{E}[1_{\mu^i \leq t}]^2 = Pr(\mu^i \leq t) - Pr(\mu^i \leq t)^2 \leq 1$$

where the inequality follows from probabilities being between zero and one. Repeat the exercise but for an arbitrary covariance.

$$\mathbb{E}[1_{\mu^i \leq t} 1_{\mu^j \leq t}] - \mathbb{E}[1_{\mu^i \leq t}]\mathbb{E}[1_{\mu^j \leq t}] = Pr(\mu^i \leq t \text{ and } \mu^j \leq t) - Pr(\mu^i \leq t)Pr(\mu^j \leq t) \leq 1$$

With these two bounds, we continue the line of reasoning in (11)

$$\begin{aligned}
\frac{\mathrm{Var}(\sum_{i=1}^M 1_{\mu^i \leq t})}{M^2 \varepsilon^2} &= \frac{\sum_{i=1}^M \mathrm{Var}(1_{\mu^i \leq t}) + \sum_{i \neq j} \mathrm{Cov}(1_{\mu^i \leq t}, 1_{\mu^j \leq t})}{M^2 \varepsilon^2} \\
&\leq \frac{M + M(n-1)K}{M^2 \varepsilon^2} \\
&\to 0
\end{aligned}$$

as $M \to \infty$. The inequality above firstly uses the previously derived variance and covariance bounds. Secondly, it uses an upper bound for the number of covariance terms, which is attained when all traders remember K periods. When traders remember K periods, their beliefs overlap with up to $(n-1)K$ other traders. $\square$

Proof of Theorem 1. To prove the first claim, the mean of demand (assuming the trader remembers k periods) is

$$\mathbb{E}\left[\frac{\frac{\gamma}{k}\sum_{t \in \mathbb{K}^i}(p_t - p_{t-1}) + (1-\gamma)\mu}{\rho \Sigma^2}\right] = \frac{\mu}{\rho \Sigma^2}$$

where μ denotes the equilibrium price drift. Note that this value does not differ across generations, and hence proportions $\mathcal{P}_j$ do not enter into the calculation. Taking the expectation of the market maker (3):[26]

$$\frac{cn\mu}{\rho \Sigma^2} = \mu \iff \mu = 0 \tag{12}$$

[26]The equilibrium variance is given by $\Sigma^2 = \frac{c\gamma}{\rho}\sqrt{n \sum_{k=1}^K \mathcal{P}_k / k}$. Then the constant that multiplies μ in expression (12), $\frac{cn}{\rho \Sigma^2}$, is always greater than one when $n > 1$.

Next we consider excess kurtosis. If the trader drawn from the cross section remembers k periods, the fourth moment of demand would be

$$\mathbb{E}\left[\left(\frac{\frac{\gamma}{k}\sum_{t\in\mathbb{K}^i}(p_t-p_{t-1})}{\rho\Sigma^2}\right)^4\right]=\gamma^4\frac{k\mu^4+3k(k-1)(\Sigma^2)^2}{k^4\rho^4(\Sigma^2)^4}$$

because the price history is independently drawn and prices have mean zero. Since the cross section is distributed according to $(\mathcal{P}_j)_{j=0}^K$, each demand has fourth moment

$$\begin{aligned}\mathbb{E}[x^i(p_t)^4]=&\sum_{k=1}^K\mathcal{P}_k\gamma^4\frac{k\mu^4+3k(k-1)(\Sigma^2)^2}{k^4\rho^4(\Sigma^2)^4}\\ =&\gamma^4\frac{\mu^4-3(\Sigma^2)^2}{\rho^4(\Sigma^2)^4}\sum_{k=1}^K\frac{\mathcal{P}_k}{k^3}+\frac{3\gamma^4}{\rho^4(\Sigma^2)^2}\sum_{k=1}^K\frac{\mathcal{P}_k}{k^2}\end{aligned}$$

First we confirm equality in the case $n\to\infty$. The fourth price moment is given by the expression

$$\mu^4=c^4(n\mathbb{E}[x^i(p_t)^4]+3n(n-1)\mathbb{E}[x^i(p_t)^2]^2) \tag{13}$$

because demands are independently drawn from the cross-section and have mean zero. We divide (13) by $(\Sigma^2)^2$ and plug in two terms: first, the fourth moment of demand from above and, second, the demand variance, $\sum_{k=1}^K\mathcal{P}_k\mathbb{E}\left[\left(\frac{\frac{\gamma}{k}\sum_{t\in\mathbb{K}^i}(p_t-p_{t-1})}{\rho\Sigma^2}\right)^2\right]$:

$$3=\lim_{n\to\infty}\frac{c^4\left(\frac{3n\gamma^4}{\rho^4(\Sigma^2)^2}\sum_{k=1}^K\frac{\mathcal{P}_k}{k^2}+3n(n-1)\left(\sum_{k=1}^K\mathcal{P}_k\frac{\gamma^2}{k\rho^2\Sigma^2}\right)^2\right)}{(\Sigma^2)^2} \tag{14}$$

Using the fact that $(\Sigma^2)^2=\frac{c^2\gamma^2n}{\rho^2}\sum_{k=1}^K\frac{\mathcal{P}_k}{k}$,

$$3=\lim_{n\to\infty}\frac{c^4\left(\frac{3n\gamma^4}{\rho^4}\sum_{k=1}^K\frac{\mathcal{P}_k}{k^2}+3n(n-1)\left(\sum_{k=1}^K\mathcal{P}_k\frac{\gamma^2}{k\rho^2}\right)^2\right)}{\frac{c^4\gamma^4n^2}{\rho^4}\left(\sum_{k=1}^K\mathcal{P}_k/k\right)^2}$$

Because the denominator is $\mathcal{O}(n^2)$, from the numerator, only $\mathcal{O}(n^2)$ terms remain which proves the desired equality. Next we show that price kurtosis, $\mu^4/(\Sigma^2)^2$, is always greater than 3 when n is finite. We follow similar steps as in the infinite case, recalculating the

RHS of (14)

$$\frac{\mu^4}{(\Sigma^2)^2} = \frac{c^4\left(\frac{\gamma^4 n\varepsilon}{\rho^4(\Sigma^2)^2}\sum_{k=1}^K \frac{\mathcal{P}_k}{k^3} + \frac{3n\gamma^4}{\rho^4(\Sigma^2)^2}\sum_{k=1}^K \frac{\mathcal{P}_k}{k^2} + 3n(n-1)\left(\sum_{k=1}^K \mathcal{P}_k \frac{\gamma^2}{k\rho^2\Sigma^2}\right)^2\right)}{(\Sigma^2)^2}$$

where

$$\varepsilon \equiv \frac{\mu^4 - 3(\Sigma^2)^2}{(\Sigma^2)^2}$$

Again plugging in $(\Sigma^2)^2 = \frac{c^2\gamma^2 n}{\rho^2}\sum_{k=1}^K \frac{\mathcal{P}_k}{k}$ and simplifying,

$$\begin{aligned}\frac{\mu^4}{(\Sigma^2)^2} =& \frac{c^4\left(\frac{\gamma^4 n\varepsilon}{\rho^4}\sum_{k=1}^K \frac{\mathcal{P}_k}{k^3} + \frac{3n\gamma^4}{\rho^4}\sum_{k=1}^K \frac{\mathcal{P}_k}{k^2} + 3n(n-1)\left(\sum_{k=1}^K \mathcal{P}_k \frac{\gamma^2}{k\rho^2}\right)^2\right)}{\left(\frac{c^2\gamma^2 n}{\rho^2}\sum_{k=1}^K \frac{\mathcal{P}_k}{k}\right)^2} \\ =& 3 + \frac{n\varepsilon\sum_{k=1}^K \frac{\mathcal{P}_k}{k^3} + 3n\sum_{k=1}^K \frac{\mathcal{P}_k}{k^2} - 3n\left(\sum_{k=1}^K \mathcal{P}_k \frac{1}{k}\right)^2}{\left(n\sum_{k=1}^K \frac{\mathcal{P}_k}{k}\right)^2}\end{aligned}$$

Plugging in the value of ε and solving for the excess kurtosis,

$$\frac{\mu^4}{(\Sigma^2)^2} = 3 + \frac{3\left(\sum_{k=1}^K \frac{\mathcal{P}_k}{k^2} - \left[\sum_{k=1}^K \frac{\mathcal{P}_k}{k}\right]^2\right)}{n\left(\sum_{k=1}^K \frac{\mathcal{P}_k}{k}\right)^2 - \sum_{k=1}^K \frac{\mathcal{P}_k}{k^3}} \tag{15}$$

and we call

$$N(\mathcal{P}) \equiv \frac{\sum_{k=1}^K \frac{\mathcal{P}_k}{k^3}}{\left(\sum_{k=1}^K \frac{\mathcal{P}_k}{k}\right)^2}$$

so that $n > N(\mathcal{P})$ guarantees that the denominator on the RHS of (15) is positive. Then for the desired result, it suffices to show that

$$\sum_{k=1}^K \frac{\mathcal{P}_k}{k^2} - \left(\sum_{k=1}^K \frac{\mathcal{P}_k}{k}\right)^2 \geq 0$$

with inequality strict when $(\mathcal{P}_j)_{j=0}^K$ is nondegenerate. But this inequality can be rewritten $\mathbb{E}[X^2] - \mathbb{E}[X]^2 \geq 0$, where probabilities are given by $\mathcal{P}_k$ and the random variable takes on values $\frac{1}{k}$. The variance is, indeed, strictly positive when the distribution is nondegenerate. $\square$

Proof of Proposition 2. We begin with the first claim, which follows from the closed-form solution for the equilibrium variance from the proof of Theorem 1:

$$\Sigma^2 = \frac{c\gamma}{\rho}\sqrt{n\sum_{k=1}^{K}\frac{\mathcal{P}_k}{k}}$$

The second claim follows immediately from the closed form solution (15) from Theorem 1. $\square$

Proof of Proposition 3. The price variance Σ^2 must solve the following fixed point condition

$$\Sigma^2 = \frac{c^2\gamma^2 n}{\rho^2\Sigma^4}\sum_{k=1}^{K}\frac{\mathcal{P}_k}{k^2}\left[\int_{-\infty}^{0} x^2 f_k(x-kr/\gamma)dx + \int_{0}^{\infty} x^2 f_k(x+kr/\gamma)dx\right],$$

where $f(x)$ denotes the price density and $f_k(x)$ denotes its k-fold convolution. The RHS of the condition above is the variance of a distribution where the middle range $[-kr/\gamma, kr/\gamma]$ of beliefs has been removed, and the tails $(-\infty, -kr/\gamma)\cup(kr/\gamma,\infty)$ have been shifted inwards toward the origin. Note that the term kr/γ follows from the following rewriting of demand (assuming that the trader remembers k periods)

$$x^i(p_t) = \begin{cases} \frac{\gamma}{k}\frac{\sum_{t\in\mathbb{K}^i}(p_t-p_{t-1})-kr/\gamma}{\rho\Sigma^2}, & \text{if } \sum_{t\in\mathbb{K}^i}(p_t-p_{t-1}) > kr/\gamma \\ \frac{\gamma}{k}\frac{\sum_{t\in\mathbb{K}^i}(p_t-p_{t-1})+kr/\gamma}{\rho\Sigma^2}, & \text{if } \sum_{t\in\mathbb{K}^i}(p_t-p_{t-1}) < -kr/\gamma \\ 0, & \text{otherwise} \end{cases}$$

Denote the function $g(r,\Sigma)$ as

$$g(r,\Sigma) \equiv \frac{c^2\gamma^2 n}{\rho^2}\sum_{k=1}^{K}\frac{\mathcal{P}_k}{k^2}\left[\int_{-\infty}^{0} x^2 f_k(x-kr/\gamma)dx + \int_{0}^{\infty} x^2 f_k(x+kr/\gamma)dx\right] - \Sigma^6$$

First we make the change of variable, $u = x - kr/\gamma$ (for the second integral, $u = x + kr/\gamma$)

$$g(r,\Sigma) = \frac{c^2\gamma^2 n}{\rho^2}\sum_{k=1}^{K}\frac{\mathcal{P}_k}{k^2}\left[\int_{-\infty}^{-kr/\gamma}(u+kr/\gamma)^2 f_k(u)du + \int_{kr/\gamma}^{\infty}(u-kr/\gamma)^2 f_k(u)du\right] - \Sigma^6$$

and, applying the Leibniz integral rule, we have that

$$\frac{\partial g}{\partial r} = \frac{c^2 \gamma n}{\rho^2} \sum_{k=1}^{K} \frac{\mathcal{P}_k}{k} \left[\int_{-\infty}^{-kr/\gamma} 2(u + kr/\gamma) f_k(u) du - \int_{kr/\gamma}^{\infty} 2(u - kr/\gamma) f_k(u) du \right] < 0$$

where the inequality follows from the fact that, in the first (second) integral, u is always evaluated below $-kr/\gamma$ (above kr/γ). When r is near zero, we also have that

$$\begin{aligned}\frac{\partial g}{\partial \Sigma} \approx & \frac{c^2 \gamma^2 n}{\rho^2} \sum_{k=1}^{K} \frac{\mathcal{P}_k}{k^2} \frac{\partial}{\partial \Sigma} \left[\int_{-\infty}^{\infty} u^2 f_k(u) du \right] - 6\Sigma^5 \\ = & \frac{c^2 \gamma^2 n}{\rho^2} \sum_{k=1}^{K} \frac{\mathcal{P}_k}{k} \frac{\partial}{\partial \Sigma} \left[\int_{-\infty}^{\infty} u^2 f(u) du \right] - 6\Sigma^5 \\ = & \frac{c^2 \gamma^2 n}{\rho^2} \sum_{k=1}^{K} \frac{\mathcal{P}_k}{k} 2\Sigma - 6\Sigma^5\end{aligned}$$

where the second line follows from the fact that the variance of the k-fold convolution is k times the price variance. But because the derivative is evaluated near $r = 0$, we know that the equilibrium $\Sigma^2 \approx \frac{c\gamma}{\rho} \sqrt{n \sum_{k=1}^{K} \mathcal{P}_k / k}$. Squaring this term then plugging in,

$$\frac{\partial g}{\partial \Sigma} \approx 2\Sigma^5 - 6\Sigma^5 < 0$$

Altogether, by the implicit function theorem, we have that $\frac{\partial \Sigma}{\partial r} = -\frac{\partial g / \partial r}{\partial g / \partial \Sigma} < 0$. □

Proof of Proposition 4. The price variance solves the following fixed point condition

$$\Sigma^2 = c^2 n \sum_{k=1}^{K} \mathcal{P}_k \left[\frac{\gamma^2}{k^2 \rho^2 \Sigma^4} \int_{-k\bar{b}\rho\Sigma^2/\gamma}^{k\bar{b}\rho\Sigma^2/\gamma} x^2 f_k(x) dx + \int_{k\bar{b}\rho\Sigma^2/\gamma}^{\infty} \bar{b}^2 f_k(x) dx + \int_{-\infty}^{-k\bar{b}\rho\Sigma^2/\gamma} \bar{b}^2 f_k(x) dx \right]$$

where $f(x)$ denotes the price density and $f_k(x)$ denotes its k-fold convolution. The first term on the RHS corresponds to traders unconstrained by the borrowing constraint. The second and third terms on the RHS correspond to all extreme traders whose demands have been truncated by the borrowing constraint. Note that the limits of integration follow from the following rewriting of demand (assuming that the trader remembers k

periods)

$$x^i(p_t) = \begin{cases} \bar{b}, & \text{if } \sum_{t\in\mathbb{K}^i}(p_t - p_{t-1}) > k\bar{b}\rho\Sigma^2/\gamma \\ -\bar{b}, & \text{if } \sum_{t\in\mathbb{K}^i}(p_t - p_{t-1}) < -k\bar{b}\rho\Sigma^2/\gamma \\ \frac{\gamma\sum_{t\in\mathbb{K}^i}(p_t-p_{t-1})}{k\rho\Sigma^2}, & \text{otherwise} \end{cases}$$

Define the function $g(\bar{b},\Sigma)$ as

$$\begin{aligned} & g(\bar{b},\Sigma) \\ \equiv & c^2 n \sum_{k=1}^{K} \mathcal{P}_k \left[\frac{\gamma^2}{k^2\rho^2} \int_{-k\bar{b}\rho\Sigma^2/\gamma}^{k\bar{b}\rho\Sigma^2/\gamma} x^2 f_k(x)dx + \Sigma^4 \int_{k\bar{b}\rho\Sigma^2/\gamma}^{\infty} \bar{b}^2 f_k(x)dx + \Sigma^4 \int_{-\infty}^{-k\bar{b}\rho\Sigma^2/\gamma} \bar{b}^2 f_k(x)dx \right] - \Sigma^6 \end{aligned}$$

and we apply the Leibniz integral rule.

$$\frac{\partial g}{\partial \bar{b}} = c^2 n \sum_{k=1}^{K} \mathcal{P}_k \Sigma^4 \left[\int_{k\bar{b}\rho\Sigma^2/\gamma}^{\infty} 2\bar{b} f_k(x)dx + \int_{-\infty}^{-k\bar{b}\rho\Sigma^2/\gamma} 2\bar{b} f_k(x)dx \right] > 0$$

where the inequality follows from direct computation of the integrals. Applying the Leibniz integral rule once again,

$$\begin{aligned} & \frac{\partial g}{\partial \Sigma} \\ = & c^2 n \sum_{k=1}^{K} \mathcal{P}_k \frac{\partial}{\partial \Sigma} \left[\frac{\gamma^2}{k^2\rho^2} \int_{-k\bar{b}\rho\overline{\Sigma}^2/\gamma}^{k\bar{b}\rho\overline{\Sigma}^2/\gamma} x^2 f_k(x)dx + \overline{\Sigma}^4 \int_{k\bar{b}\rho\overline{\Sigma}^2/\gamma}^{\infty} \bar{b}^2 f_k(x)dx + \overline{\Sigma}^4 \int_{-\infty}^{-k\bar{b}\rho\overline{\Sigma}^2/\gamma} \bar{b}^2 f_k(x)dx \right] \\ & - 6\Sigma^5 + \varepsilon \\ \approx & \frac{c^2\gamma^2 n}{\rho^2} \sum_{k=1}^{K} \frac{\mathcal{P}_k}{k^2} \frac{\partial}{\partial \Sigma} \left[\int_{-\infty}^{\infty} x^2 f_k(x)dx \right] - 6\Sigma^5 + \varepsilon \\ = & \frac{c^2\gamma^2 n}{\rho^2} \sum_{k=1}^{K} \frac{\mathcal{P}_k}{k} \frac{\partial}{\partial \Sigma} \left[\int_{-\infty}^{\infty} x^2 f(x)dx \right] - 6\Sigma^5 + \varepsilon \\ = & \frac{c^2\gamma^2 n}{\rho^2} \sum_{k=1}^{K} \frac{\mathcal{P}_k}{k} 2\Sigma - 6\Sigma^5 + \varepsilon \end{aligned}$$

where $\overline{\Sigma}^2$ denotes the equilibrium variance but held constant (with respect to the partial derivative), and the approximate equality holds when $\bar{b}$ is large. Note that the second-to-last line follows from the fact that the variance of the k-fold convolution equals k times the price variance. The variable ε denotes an error term, which we will show disappears when $\bar{b}$ is large. Before that, because the expression above is evaluated for a large value of $\bar{b}$, we know that the equilibrium $\Sigma^2 \approx \frac{c\gamma}{\rho}\sqrt{n\sum_{k=1}^{K}\mathcal{P}_k/k}$. Plugging in the square of this expression,

$$\frac{\partial g}{\partial \Sigma} \approx 2\Sigma^5 - 6\Sigma^5 + \varepsilon < 0$$

when ε is small. By the implicit function theorem, we have that $\frac{\partial \Sigma}{\partial \bar{b}} = -\frac{\partial g/\partial \bar{b}}{\partial g/\partial \Sigma} > 0$. The last step is to confirm that $\varepsilon \to 0$ when $\bar{b} \to \infty$. Without limits, it is

$$\varepsilon \equiv 4c^2 n\Sigma^3 \sum_{k=1}^{K} \mathcal{P}_k \left[\int_{k\bar{b}\rho\Sigma^2/\gamma}^{\infty} \bar{b}^2 f_k(x)dx + \int_{-\infty}^{-k\bar{b}\rho\Sigma^2/\gamma} \bar{b}^2 f_k(x)dx \right] \tag{16}$$

The entire term (16) tends to zero because

$$\int_{k\bar{b}\rho\Sigma^2/\gamma}^{\infty} \bar{b}^2 f_k(x)dx \leq \int_{k\bar{b}\rho\Sigma^2/\gamma}^{\infty} \left(\frac{x\gamma}{k\rho\Sigma^2}\right)^2 f_k(x)dx \to 0, \text{ as } \bar{b} \to \infty$$

where the inequality holds because $x \geq k\bar{b}\rho\Sigma^2/\gamma$. The same argument applies to the second term on the RHS of (16), except now $x \leq -k\bar{b}\rho\Sigma^2/\gamma$. $\square$

References

S. Andersen, T. Hanspal, and K. Nielsen. Once bitten, twice shy: the power of personal experiences in risk taking. *Journal of Financial Economics*, 132:97–117, 2019.

W. Bagehot. The only game in town. *Financial Analysts Journal*, 27(2):12–14, 1971.

P. Bordalo, N. Gennaioli, and A. Shleifer. Memory, attention, and choice. *Quarterly Journal of Economics*, 135(3):1399–1442, 2020.

T. Chordia, R. Roll, and A. Subrahmanyam. Order imbalance, liquidity, and market returns. *Journal of Financial Economics*, 65:111–130, 2002.

P. Collin-Dufresne, M. Johannes, and L. Lochstoer. Asset pricing when this time is different. *Review of Financial Studies*, 30(2):505–538, 2017.

R. Cont and J. Bouchaud. Herd behavior and aggregate fluctuations in financial markets. *Macroeconomic Dynamics*, 4(2):170–196, 2000.

J. Dominitz and C. Manski. Measuring and interpreting expectations of equity returns. *Journal of Applied Econometrics*, 26:352–370, 2011.

J. Doob. Application of the theory of martingales. *Actes du Colloque International Le Calcul des Probabilites et ses applications*, Paris CNRS:23–27, 1949.

R. Durrett. *Probability: Theory and Examples*. Cambridge University Press, 2017.

P. Ehrling, A. Graniero, and C. Heyerdahl-Larsen. Asset prices and portfolio choice with learning from experience. *Review of Economic Studies*, 85(3):1752–1780, 2018.

X. Gabaix, P. Gopikrishnan, V. Plerou, and E. Stanley. Institutional investors and stock market volatility. *Quarterly Journal of Economics*, 121(2):461–504, 2006.

S. Giglio, M. Maggiori, J. Stroebel, and S. Utkus. Five facts about beliefs and portfolios. *American Economic Review*, 111(5):1481–1522, 2021.

J. Graham, H. Campbell, and H. Huang. Investor competence, trading frequency, and home bias. *Management Science*, 55(7):1094–1106, 2009.

D. Grether. Bayes rule as a descriptive model: The representativeness heuristic. *Quarterly Journal of Economics*, 95(3):537–557, 1980.

J. Hasbrouck and G. Sofianos. The trades of market makers: An empirical analysis of nyse specialists. *Journal of Finance*, 48(5):1565–1593, 1993.

M. Kahana. *Foundations of Human Memory*. Oxford University Press, 2012.

M. Kaustia and S. Knupfer. Do investors overweight personal experience? evidence from ipo subscriptions. *Journal of Finance*, 63(6):2679–2702, 2008.

T. Kuchler and B. Zafar. Personal experiences and expectations about aggregate outcomes. *Journal of Finance*, 74(5):2491–2542, 2019.

A. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335, 1985.

U. Malmandier and J. Wachter. Memory of past experiences and economic decisions. *Oxford Handbook of Human Memory*, 2022.

U. Malmandier, D. Pouzo, and V. Vanasco. Investor experiences and financial market dynamics. *Journal of Financial Economics*, 136(3):597–622, 2020.

U. Malmendier and S. Nagel. Depression babies: Do macroeconomic experiences affect risk taking? *Quarterly Journal of Economics*, 126(1):373–416, 2011.

U. Malmendier and S. Nagel. Learning from inflation experiences. *Quarterly Journal of Economics*, 131(1):53–87, 2015.

S. Mullainathan. A memory-based model of bounded rationality. *Quarterly Journal of Economics*, 117(3):735–774, 2002.

S. Nagel and Z. Xu. Asset pricing with fading memory. *Review of Financial Studies*, 35 (5):2190–2245, 2022.

D. Richards and G. Willows. Who trades profusely? the characteristics of individual investors who trade frequently. *Global Finance Journal*, 35:1–11, 2018.

S. Schraeder. Information processing and non-bayesian learning in financial markets. *Review of Finance*, 20(2):823–853, 2016.

R. Shiller. *Irrational Exuberance*. Princeton University Press, 2015.

G. Sofianos. Specialist gross trading revenues at the new york stock exchange. *Working Paper*, 1995.

M. Strahilevitz, T. Odean, and B. Barber. Once burned, twice shy: how naive learning, counterfactuals, and regret affect the repurchase of stocks previously sold. *Journal of Marketing Research*, 48:102–120, 2011.

K. Teeple. Support, resistance, and technical trading. *Working Paper*, 2022.

K. Teeple. Optimally stochastic volatility. *Working Paper*, 2023.

S. Thurner, J. Farmer, and J. Geanakoplos. Leverage causes fat tails and clustered volatility. *Quantitative Finance*, 12(5):695–707, 2012.

J. Wachter and M. Kahana. A retrieved-context theory of financial decisions. *Quarterly Journal of Economics*, 139(2):1095–1147, 2024.