Memorable Events in Financial Markets

Andrés Carvajal* and Keisuke Teeple†

January 22, 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 (or, equivalently, traders failing to pay attention on 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

[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.

given period; this effectively "draws" traders independently from the cross-section and is, in fact, the only exogenous random variable in the model. This captures the realistic constraint that traders are not all simultaneously active in financial markets.

The final participant in our economy is a simple linear market maker like the one from Teeple (2022). This market maker attempts to clear markets by raising prices when there is excess demand and lowering prices when there is excess supply. When a particular parameter in the market maker's problem tends to infinity, we recover the special case of the classical Walrasian auctioneer. It is this tâtonnement process that generates price volatility in our setting. 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 we 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. We can eliminate many potential equilibrium distributions using the fact that few distributions have the stability property.[2] In our setting, 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 mar-

[2] A distribution is stable when the sum of i.i.d. random variables have that same distribution. The Normal distribution is a well-known example.

ket maker. Because equilibrium requires that past and future prices share a common distribution, stability is *necessary* for equilibrium.

Up to this point, we took the days which a particular trader remembered as given. In our main extension, 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 tighter 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 traders with a flat prior. The steady-state cross-section consists of a mixture of the youth (traders who have not traded as much, leading to dispersed beliefs) and the elderly (traders who have traded many times, leading to tighter 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 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. small traders), borrowing limits truncate trade for the tails of the distribution (i.e. large traders).

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. So while the key distinction between our model and theirs is the idiosyncratic versus fading memory, a secondary distinction is that our price volatility is purely non-fundamental and theirs originates from exogenous dividend risk.

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 includes a discussion on uniqueness and considers a richer class of settings, and Section 4 models the cross-section when traders remember active trading days. Section 5 addresses policy implications. Section 6 concludes.

2 Motivating Example

Traders. Time is discrete, infinite, and indexed by $t \in \mathbb{Z}$, 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.[3] 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, this price history will need to be consistent with prices generated within the model. Traders, indexed by i,

[3]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).

maximize mean-variance utility

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

where ρ denotes the risk aversion parameter.[4] 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*. While results will be generalized in Section 3, here we let each trader remember only one period, and furthermore 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:

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

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

$$\alpha_i^i = \gamma T, \text{ and } \sum_{t\neq i} \alpha_t^i = (1-\gamma)T$$

where $0 < \gamma \leq 1$ denotes idiosyncratic memory strength. Just as in the Bayesian case, the weights add up to the number of observations, T. However, here we distribute γ proportion of that weight to the memorable period, and the rest on all other periods. While not explicitly written above, there is an additional assumption that weight is distributed *evenly* among these remaining periods. In some sense, traders are Bayesian over remaining periods. When a memory-constrained Bayesian after T periods forms a

[4]While the trader may seem myopic in the sense that she does not look beyond period $(t+1)$, allocations agree with those from an infinite horizon setting with objective

$$\sum_{t=0}^{\infty} \beta^t \mathbb{E}^i[w_{t+1} - \frac{\rho}{2} w_{t+1}^2 | p_0] \quad \text{s.t.} \quad w_{t+1} = (p_{t+1} - p_t)x_t + w_t$$

posterior mean, it becomes a weighted average between what she remembers and what she does not

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

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.[5] 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-section of beliefs μ^i across the population approaches $N(0, \gamma^2\Sigma^2)$ as the number of traders, which we denote M, increases. More formally, the empirical distribution

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

approaches the Normal distribution 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

$$p_{t+1} = p_t + c\sum_{i=1}^{n} x^i(p_t) \tag{2}$$

ensures that prices follow a discrete time Brownian motion.[6] Normality follows from the fact that the sum of i.i.d. Normal random variables in (2) 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 (2). 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).

[5] 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.

[6] To achieve a geometric Brownian motion, we can redefine all prices in logs.

The market maker 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.[7] 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 excess demand, effectively clearing markets each period.

This objective function leads to the intuitive rule (2): 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 dynamics in a standard model. However, in our setup, we do away with

[7] 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.

dividends so that our equilibrium takes on a self-fulfilling flavor and volatility in our setting becomes purely non-fundamental. *Some* form of exogenous risk is still needed, however, and we model this by randomly drawing the identities of 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 this 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}.[8] Importantly, they do 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 (2) 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).

Discussion. 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. Furthermore, without competitive markets, traders may form beliefs about others; here, they only need to form beliefs about the statistical properties of prices.

Notice that we make no claims about how this Brownian equilibrium is initially

[8] The second half of this round trip is not included in the excess demand for period $(t+1)$.

approached. In the example there is a one-to-one mapping between time periods and traders, so we order the traders according to time and begin the model after an infinite amount of time has passed. One might ask: What, then, do traders do if their memorable period has not yet occurred? One natural approach is to assume they are Bayesian and, hence, do not trade. An alternative simple approach is to assume that the subset of size n (that trades) is selected from the set of traders who have already observed their memorable period. While the two approaches carry slightly different interpretations, ultimately the modelling choice is inconsequential for our main results.

Prices are volatile because they have always been volatile; past volatility is what creates the required cross-sectional dispersion in beliefs. A natural next question is whether the Brownian motion is the unique equilibrium price process. This question will be addressed in the next section, along with relaxations of the requirement that each trader remember one period disjoint from any other trader.

3 Generalizations

Before any generalizations, 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 initally drawn. Formally, an equilibrium is a price history $\mathcal{H}_t$, trader demands $x^i(p_t)$, and a price distribution F_t satisfying:

(a) Given the recollection $\mathcal{H}_t^i \subset \mathcal{H}_t$ and conjecture Σ_t^2, $x^i(p_t)$ solves each trader's optimization problem if they are in the active group, $i \in \mathcal{A}_t$. Otherwise, $x^i(p_t) = 0$.

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

Furthermore, we focus our attention on steady-state equilibria where $F_t = F$ ($\Sigma_t^2 = \Sigma^2$) for all t. Consider the benchmark Brownian motion from Section 2. 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 (2)

$$p_{t+1} = p_t + c \left(\sum_i x^i(p_t) - S \right) \tag{3}$$

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 (3). With some algebra, this condition reduces to

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

The intuition is that positive drift generates excess demand; this, in turn, requires a positive asset supply.[9] Importantly, $\mu = 0$ when $S = 0$, establishing the uniqueness of zero drift in the baseline setting.

Since both variance (pinned down in Section 2) and drift are unique, we say that the equilibrium is unique. However, we can say more by making use of the following well-known statistical fact: 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 are linear in demand due to the nature of the market maker objective. 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

[9] $S > 0$ does not change the variance calculation, hence $\Sigma^2 = c\gamma\sqrt{n}/\rho$ like before.

distribution.

Next we generalize several restrictive requirements of the example. First, we examine the effect of memorable periods being shared by multiple traders.[10] We allow for heterogeneity in the sense that each period may be remembered by a different (but finite) number of traders. All proofs are in Appendix A.

Proposition 1. ***(Overlapping Memory)*** *Say that each period is remembered by k traders for $k \in \{1, ..., K\}$ and, furthermore, say that each trader remembers only one period. Then prices converge in distribution to a discrete time Brownian motion as $T \to \infty$.*[11]

First imagine that each period is remembered by a large − but identical − number of traders. This makes each historical price realization equally likely to be drawn from the cross-sectional distribution over beliefs. Although there are more traders, the dispersion of beliefs remains precisely the same as in the case with no overlap. Now consider adding heterogeneity, which does not change the result for the following reason. First partition the historical price realizations into ones remembered by one trader, ones remembered by two traders, and so on. But because traders only remember one period by assumption, this partition of past realizations also partitions the trader population. When a trader is drawn from the cross-section, there is some chance that she is from the k-th group (i.e. the group of traders who remember a time period remembered by k traders). In other words, the cross-section is a mixture of these k groups for $k = 1, ..., K$. Previous arguments for the homogeneous case now apply group-by-group. Prices converge to a Brownian motion with the same equilibrium variance as in Section 2.

Next we revert back to mutually exclusive memorable events, as in the baseline setting, but we assume that each trader remembers a finite set of periods, which generalizes the previous requirement that each trader remember *one* period. Analogous to Proposition 1, by allowing for any finite set, we allow for heterogeneity: one trader may remember two periods, while another may remember three.

Proposition 2. ***(Expanded Memory)*** *Say that each trader remembers k periods for $k \in \{1, ..., K\}$ and, furthermore, say that no two traders' memorable periods overlap.*

[10] Like in Section 2, there is an implicit assumption that the indexes of memorable prices cannot be measurable with respect to the realization. This, for example, rules out extreme prices being remembered by more traders. This comment also applies to Proposition 2.

[11] Note that sending the history $T \to \infty$ implies that the number of traders $M \to \infty$ at the same rate. This comment also applies to Proposition 2.

Then prices converge in distribution to a discrete time Brownian motion as $n \to \infty$ and $T \to \infty$.

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 of memorable events hold less dispersed, or tighter, beliefs. This, in itself, is not problematic 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).

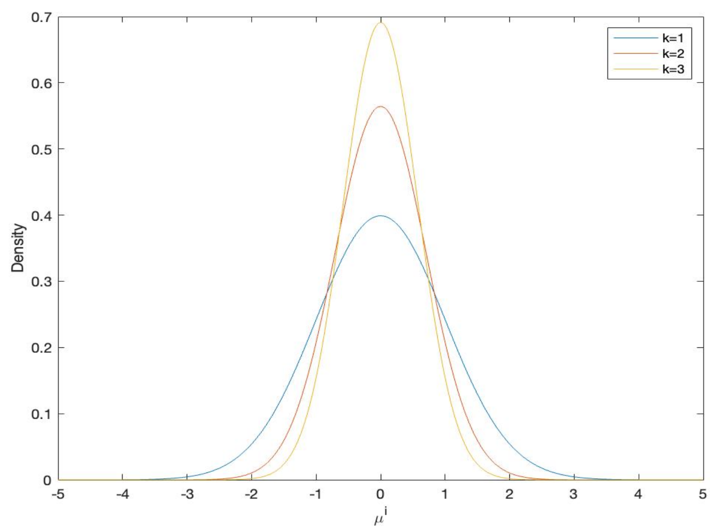

Figure 1: Cross-Section with Expanded Memory

Now consider adding heterogeneity in memory length. Interestingly, the cross-sectional distribution of beliefs is not Normal due to heterogeneity in the population, 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 tight 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. Sending $n \to \infty$ is with loss of generality, because conclusions from Section 2 hold for any n.[12] A similar line of reasoning can be applied to

[12] 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.

heterogeneous memory strength γ^i and risk aversion ρ^i. As in Proposition 2, the cross-sectional distribution of beliefs need not be Normal. However, prices are Normal due to the central limit theorem.

Next, we recast this last result not as a negative one (i.e. failure of Normality) 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.

4 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 changes. When the random subset of size n trade, they remember that period's return (after they trade and the return is realized). The memory parameter, γ, is set to one so that only such days are remembered.[13] In the background, we assume that there is a cost of paying attention on non-active days that makes it prohibitively costly for traders to do so. This is consistent with empirical 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)). With this new setup, 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.

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

[13]When $\gamma < 1$, we previously assumed that traders were Bayesian before their memorable period occurred. When $\gamma = 1$ and their memorable period has yet to occur, traders put full weight on their (flat) prior. We assume this has mean zero so they do not trade.

$\mathcal{P}_1^{t+1}$. 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:

$$\mathcal{P} = \left(\frac{\mathcal{D}}{\mathcal{D}+n}, \frac{\mathcal{D}n}{(\mathcal{D}+n)^2}, \frac{\mathcal{D}n^2}{(\mathcal{D}+n)^3}, ... \right) \tag{4}$$

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 $\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, ...\right)$. Now consider the equilibrium conjecture, Σ^2. Recall that the market maker generates prices according to the rule (2). If the trader drawn from the cross section remembers k periods, the variance of demand would be

$$\text{Var}\left(\frac{\frac{1}{k}\sum_{t\in\mathbb{K}^i}(p_t - p_{t-1})}{\rho\Sigma^2} \right) = \frac{1}{k\rho^2\Sigma^2} \tag{5}$$

where $\mathbb{K}^i$ denotes the set of memorable periods for trader i. In (5), when traders remember k periods, they take the average of those k periods when forming their posterior mean. While this is consistent with Section 2, the observant reader will have noticed that traders may not use the sample average as their Bayesian posterior mean when prices are not Normal. Indeed, the upcoming theorem states that prices have heavy tails (i.e. non-Normal). Despite the well-known difficulties in deriving a closed-form solution for the posterior mean of arbitrary distributions, our first lemma addresses this issue.

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

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

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.

Recall that μ denotes the mean of the price distribution (zero in our setting) 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. So how do we circumvent the usual algebraic difficulties when dealing with non-Normal posteriors? The key is in the qualifier, $T \to \infty$. Asymptotically, there are results from probability theory that guarantee consistency of the Bayesian posterior mean. In order to invoke these, we consider a mixture distribution that generates precisely the same frequency of observations as in (1). With probability $\gamma/|\mathbb{K}^i|$, $(p_t - p_{t-1})$ is drawn for each $t \in \mathbb{K}^i$. With probability $(1-\gamma)$, prices are independently drawn from the price distribution. The mean of such a distribution must equal the value μ^i described in Lemma 1. What makes this work is that the weights in (6) have been selected so that it is *as-if* the trader not only observes an infinite number of independent observations from the price distribution, but also an infinite number of 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 is μ^i (i.e. the posterior variance tends to zero).

Going back to the variance calculation (5), we must proceed with caution because an implicit assumption in this calculation is that prices are independent. Price independence is implied by two further assumptions. First, the same trader can never be drawn twice in a span of k periods. This guarantees that all kn traders who generated these k prices are unique. Second, there can be no overlap in beliefs between a set of kn traders. This furthermore guarantees that the k prices in (5) were generated by traders whose beliefs are based on mutually exclusive past price realizations. These two points are confirmed in the next lemma.

Lemma 2. ***(Price Independence)*** *Assume that the maximum number of periods remembered in the economy is K. As the number of traders $M \to \infty$, the probability of the same trader trading twice in a span of k periods becomes zero. As the history $T \to \infty$, the probability that kn traders' beliefs overlap becomes zero.*

The key to the lemma is that we truncate the cross-section of the economy, (4), at a large number K. That is, we ignore traders who remember an arbitrarily large number of 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, such traders have small mass according to (4). Hence, the economy considered in Lemma 2 becomes arbitrarily close to the true cross-section when K is large. Note that this need not be true if, for example, traders who remembered more periods traded proportionately *more* (not less). Beyond this important observation, Lemma 2 follows from a direct computation of probabilities.

Continuing the variance calculation, because the cross-section is distributed according to the geometric distribution (4), each demand has variance

$$\mathrm{Var}(x^i(p_t)) = \frac{\mathcal{D}}{\mathcal{D}+n}\sum_{k=1}^{\infty}\left(\frac{n}{\mathcal{D}+n}\right)^k \frac{1}{k\rho^2\Sigma^2} = \frac{\mathcal{D}}{(\mathcal{D}+n)\rho^2\Sigma^2} ln\left(\frac{\mathcal{D}+n}{\mathcal{D}}\right)$$

Now taking the variance of the market maker condition (2), we have that

$$\Sigma^2 = \frac{c^2 n}{\rho^2\Sigma^2}\frac{\mathcal{D}}{\mathcal{D}+n} ln\left(\frac{\mathcal{D}+n}{\mathcal{D}}\right)$$

which can be solved for Σ^2. Note that, in this calculation above, we have again invoked the assumption − proved in Lemma 2 − that demands are independent (i.e. no overlap in beliefs). Next we consider price tails. The fourth price moment is[14]

$$\mu^4 = c^4\mathbb{E}\left[\left(\sum_i x^i(p_t)\right)^4\right] = c^4\left(n\mathbb{E}[x^i(p_t)^4] + 3n(n-1)\mathbb{E}[x^i(p_t)^2]^2\right) \tag{7}$$

assuming that $\mathbb{E}[x^i(p_t)] = 0$.[15] The next theorem shows that this object (7), normalized by the squared variance, is always greater than three. In other words, equilibrium prices characterized by a mixture of traders distributed according to the cross-section (4) have heavy tails.[16]

Theorem 1. ***(Heavy Tails)*** *$\mu^4/(\Sigma^2)^2 \geq 3$, and the inequality holds with equality if and only if $n \to \infty$.*

We conclude that prices are heavy-tailed, and that 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. 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

[14] This formula is derived from an induction argument.

[15] That prices have zero drift can be shown using an argument like the one in Section 3.

[16] Because the upcoming theorem considers the case $n \to \infty$ (alongside the finite case), the model is no longer stationarity when $\mathcal{D}$ is a fixed number. Hence we assume that $\mathcal{D} = \delta n$ for some $\delta > 0$. This is without loss in the finite ($n < \infty$) case.

kurtosis due to the central limit theorem. Theorem 1 can be understood as characterizing the net effect between these two competing forces (mixing and summing).

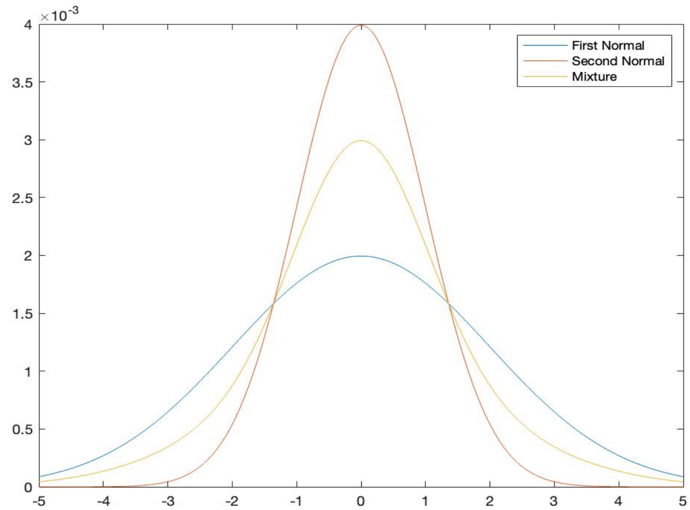

Figure 2: Heavy-Tailed Mixture Distributions

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.

5 Policy

We consider three policies to reduce price variance in the baseline model from Section 2: a transaction tax, a capital gains tax, and tightened borrowing constraints. Certain policies may make equilibrium prices non-Normal. However, Lemma 1 guarantees that this is not problematic, at least from the perspective of Bayesian estimation.

Transaction Tax. While mean-variance utility was previously used without formal derivation, here we begin 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) 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. Instead, the trader's wealth with a transaction tax $r > 0$ is

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

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:

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

Traders who remain active in financial markets trade smaller quantities. While it may seem immediate that the price variance in the following period 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.[17]

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

[17] The same comment applies to Proposition 4.

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 3 for an illustration of this truncated cross-section of demands x_t^i. With this lowered dispersion of demands, we find that equilibrium price variance is reduced when transaction taxes are reasonably small.

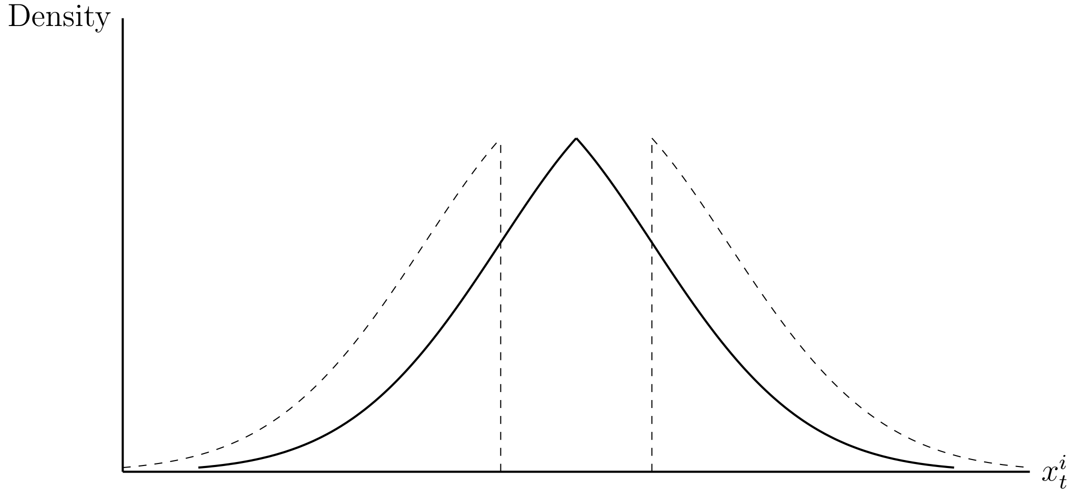

Figure 3: Truncated Cross-Section of Demands

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

$$\Sigma^2 = c^2 \text{Var} \left[\sum_{i=1}^{n} \frac{(1-\tau)\gamma(p_i - p_{i-1})}{\rho(1-\tau)^2 \Sigma^2} \right]$$

where the term inside the brackets is the capital gains-adjusted aggregate demand. This condition is solved by $\Sigma^2 = \frac{c\gamma\sqrt{n}}{\rho(1-\tau)}$, so that taxes are actually *counterproductive*. 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. In addition to the standard budget constraint, we introduce two additional constraints

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

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 applies and represents a standard borrowing constraint on the riskless asset. When traders take a short position on the risky asset, $x_t < 0$, the second constraint 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

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

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 $\frac{\tau}{\rho\Sigma^2}$. 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. Beyond our setting, where the distribution simply captures belief dispersion, this discussion may have consequences in settings where the distribution is also related to trader wealth.

While both transaction taxes and borrowing limits are effective policy tools, capital gains taxes are not. This is due to the additive nature of the two effective policies. The two 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 classified as purely non-fundamental.

We then formally modeled one reason why traders might have idiosyncratic memory: paying attention on 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 tight 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.

The model could be extended in several directions. First, we have assumed that traders exogenously die in Section 4. This could be extended to a survival story, where only traders who have sufficient funds continue to trade. Because of the lack of wealth effects in our model, our conjecture is that most results would go through in this new setting. Second, we assumed that traders are independently drawn from the cross-section. Introducing persistence here could help the model better match empirical price moments. We have abstained from this exercise in this paper, however, due to the similarities with fading memory models à la Malmandier et al. (2020).

Appendix A

Proof of Proposition 1. Like in Section 2, by the strong law of large numbers we have that

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

Let M_k denote the set of traders who remember a time period remembered by k traders (so that $M = \sum_k |M_k|$), let m_k denote that same set removing those with identical beliefs (so that $|M_k| = k|m_k|$), and let β_k denote their proportion in the population (so $\beta_k = \frac{|M_k|}{M}$). Now consider the empirical distribution over the cross-section of beliefs μ^i in the population:

$$F_M(t) = \frac{1}{M}\sum_{i=1}^{M} 1_{\mu^i \le t} = \sum_{k=1}^{K} \frac{\beta_k}{|M_k|} \sum_{i \in M_k} 1_{\mu^i \le t} = \sum_{k=1}^{K} \frac{\beta_k}{|M_k|} \sum_{i \in m_k} k 1_{\mu^i \le t} = \sum_{k=1}^{K} \frac{\beta_k}{|m_k|} \sum_{i \in m_k} 1_{\mu^i \le t}$$

where the third equality follows from the fact that beliefs agree for sets of traders of size k. The empirical distribution then converges to the $N(0, \gamma^2\Sigma^2)$ almost surely as $M \to \infty$, so the market maker ensures that prices converge in distribution to the discrete time Brownian motion with the same equilibrium variance as the Motivating Example, $\Sigma^2 = c\gamma\sqrt{n}/\rho$. $\square$

Proof of Proposition 2. 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 $|\mathbb{K}^i| = k$. The weights associated with memory-constrained Bayes' Rule adjust to:

$$\sum_{t \in \mathbb{K}^i} \alpha_t^i = \gamma T, \text{ and } \sum_{t \notin \mathbb{K}^i} \alpha_t^i = (1-\gamma)T$$

for $0 < \gamma \le 1$. Then, after T periods, the posterior mean becomes

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

which nests the Motivating Example as a special case when $k = 1$. Like before, the strong law of large numbers holds as $T \to \infty$, and we have that

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

Let M_k denote the set of traders who remember k periods (so $\sum_k |M_k| = M$), and let β_k denote their proportion in the population (so $\frac{|M_k|}{M} = \beta_k$). The empirical distribution can be written

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

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 β_k, which need not be Normal. However, a straightforward calculation yields a mean of zero, and variance

$$\sum_{k=1}^{K} \beta_k (\mathbb{E}[X_k^2]) = \sum_{k=1}^{K} \beta_k \frac{\gamma^2}{k}$$

because $X_k \sim N(0, \gamma^2/k)$. If a randomly selected subset of size n trades each period, then the market maker (2) ensures that prices follow a discrete time Brownian motion with variance one when $c = \frac{\rho}{\gamma \sqrt{n \sum_k \beta_k / k}}$ as $n \to \infty$. □

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

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

As $T \to \infty$, this posterior approaches the one that would have been generated from i.i.d. draws from a distribution – call it F – 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. Mixture distributions have an expected value that is equal to the mixture of the expected values:

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

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

$\square$

Proof of Lemma 2. Let M be large but finite (in particular, $M > kn$). The probability of the same trader trading twice or more in a span of k periods is

$$\mathcal{P}(M,n,k) = 1 - \underbrace{\left(\frac{M-n}{M}\right)\left(\frac{M-n-1}{M-1}\right)\cdots\left(\frac{M-n-[n-1]}{M-[n-1]}\right)}_{\text{Second period}}\cdots$$
$$\cdots\underbrace{\left(\frac{M-[k-1]n}{M}\right)\left(\frac{M-[k-1]n-1}{M-1}\right)\cdots\left(\frac{M-[k-1]n-[n-1]}{M-[n-1]}\right)}_{k\text{-th period}}$$

which assumes that traders are independently drawn. Importantly, we have that

$$\lim_{M\to\infty}\mathcal{P}(M,n,k) = 0$$

Now let T be large but finite (in particular, $T > knK$). The probability that kn traders' beliefs have any amount of overlap is bounded from above by the following, assuming the worst-case scenario where all kn traders remember K periods

$$\mathcal{Q}(T,n,k,K) = 1 - \underbrace{\left(\frac{T-K}{T}\right)\left(\frac{T-K-1}{T-1}\right)\cdots\left(\frac{T-K-[K-1]}{T-[K-1]}\right)}_{\text{Second trader}}\cdots$$
$$\cdots\underbrace{\left(\frac{T-[kn-1]K}{T}\right)\left(\frac{T-[kn-1]K-1}{T-1}\right)\cdots\left(\frac{T-[kn-1]K-[K-1]}{T-[K-1]}\right)}_{kn\text{-th trader}}$$

which assumes that traders' beliefs are independently drawn. Like before, we have that $\lim_{T\to\infty}\mathcal{Q}(T,n,k,K) = 0$. $\square$

Proof of Theorem 1. If the trader drawn from the cross section remembers k periods, the fourth moment of demand would be

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

Since the cross section is distributed according the geometric distribution (4), each demand has fourth moment

$$\begin{aligned}
\mathbb{E}[x^i(p_t)^4] =&\frac{\mathcal{D}}{\mathcal{D}+n}\sum_{k=1}^{\infty}\left(\frac{n}{\mathcal{D}+n}\right)^k \frac{k\mu^4+3k(k-1)(\Sigma^2)^2}{k^4\rho^4(\Sigma^2)^4} \\
=&\frac{\mathcal{D}}{\mathcal{D}+n}\left[\frac{\mu^4-3(\Sigma^2)^2}{\rho^4(\Sigma^2)^4}\sum_{k=1}^{\infty}\left(\frac{n}{\mathcal{D}+n}\right)^k\frac{1}{k^3}+\frac{3}{\rho^4(\Sigma^2)^2}\sum_{k=1}^{\infty}\left(\frac{n}{\mathcal{D}+n}\right)^k\frac{1}{k^2}\right] \\
=&\frac{\mathcal{D}}{\mathcal{D}+n}\left[\frac{\mu^4-3(\Sigma^2)^2}{\rho^4(\Sigma^2)^4}Li_3\left(\frac{n}{\mathcal{D}+n}\right)+\frac{3}{\rho^4(\Sigma^2)^2}Li_2\left(\frac{n}{\mathcal{D}+n}\right)\right]
\end{aligned}$$

where $Li_n(\cdot)$ denotes the polylogarithm function. First we confirm equality in the case $n\to\infty$. Based on the discussion in the main text, we write $\mathcal{D}=\delta n$ so that $\frac{n}{\mathcal{D}+n}=\frac{1}{\delta+1}$ and $\frac{\mathcal{D}}{\mathcal{D}+n}=\frac{\delta}{\delta+1}$. Plugging in our expressions into the fourth price moment (7) and dividing by $(\Sigma^2)^2$

$$3=\lim_{n\to\infty}\frac{c^4\left(\frac{\delta n}{\delta+1}\frac{3}{\rho^4(\Sigma^2)^2}Li_2\left(\frac{1}{\delta+1}\right)+3n(n-1)\left(\frac{\delta ln([\delta+1]/\delta)}{(\delta+1)\rho^2\Sigma^2}\right)^2\right)}{(\Sigma^2)^2} \tag{8}$$

Using the fact that $\Sigma^2=\frac{c^2n}{\rho^2\Sigma^2}\frac{\delta}{\delta+1}ln\left(\frac{\delta+1}{\delta}\right)$ and simplifying,

$$3=\lim_{n\to\infty}\frac{\frac{3n(\delta+1)}{\delta}Li_2\left(\frac{1}{\delta+1}\right)+3n(n-1)ln\left(\frac{\delta+1}{\delta}\right)^2}{n^2ln\left(\frac{\delta+1}{\delta}\right)^2}$$

From this expression, only $\mathcal{O}(n^2)$ terms remain and this 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, calculating the RHS of (8)

$$RHS=\frac{c^4\left(\frac{\delta n}{\delta+1}\left[\frac{\varepsilon}{\rho^4(\Sigma^2)^2}Li_3\left(\frac{1}{\delta+1}\right)+\frac{3}{\rho^4(\Sigma^2)^2}Li_2\left(\frac{1}{\delta+1}\right)\right]+3n(n-1)\left(\frac{\delta ln([\delta+1]/\delta)}{(\delta+1)\rho^2\Sigma^2}\right)^2\right)}{(\Sigma^2)^2}$$

where

$$\varepsilon\equiv\frac{\mu^4-3(\Sigma^2)^2}{(\Sigma^2)^2}>0$$

Simplifying,

$$\begin{aligned} RHS =& \frac{\frac{n(\delta+1)}{\delta}\left[\varepsilon Li_3\left(\frac{1}{\delta+1}\right) + 3Li_2\left(\frac{1}{\delta+1}\right)\right] + 3n(n-1)ln\left(\frac{\delta+1}{\delta}\right)^2}{n^2 ln\left(\frac{\delta+1}{\delta}\right)^2} \\ =& 3 + \frac{\frac{n(\delta+1)}{\delta}\left[\varepsilon Li_3\left(\frac{1}{\delta+1}\right) + 3Li_2\left(\frac{1}{\delta+1}\right)\right] - 3n ln\left(\frac{\delta+1}{\delta}\right)^2}{n^2 ln\left(\frac{\delta+1}{\delta}\right)^2} \end{aligned}$$

Because $Li_3\left(\frac{1}{\delta+1}\right) > 0$, for our desired result it suffices to show that

$$\frac{\delta+1}{\delta} Li_2\left(\frac{1}{\delta+1}\right) \geq ln\left(\frac{\delta+1}{\delta}\right)^2 \tag{9}$$

Using the fact that $Li_2(1-x) = \int_1^x \frac{ln(u)}{1-u} du$, we can rewrite (9) as

$$\int_1^{\frac{\delta}{\delta+1}} \frac{ln(u)}{1-u} du \geq \int_1^{\frac{\delta}{\delta+1}} [ln(u)^2 + 2ln(u)] du$$

where the integrand on the RHS is precisely the function that, when integrated, equals $\frac{\delta}{\delta+1} ln\left(\frac{\delta+1}{\delta}\right)^2$. With some algebra, we get

$$\int_{\frac{\delta}{\delta+1}}^1 ln(u)\left(\frac{1}{1-u} - ln(u) - 2\right) du \leq 0$$

A sufficient condition for this is

$$f(u) \equiv \frac{1}{1-u} - ln(u) - 2 \geq 0, \text{ for } u \in (0,1)$$

The function on the LHS of the inequality is strictly convex in this interval, and hence its minimum is defined by the first-order condition

$$\frac{1}{(1-u^*)^2} - \frac{1}{u^*} = 0 \iff u^* = \frac{3-\sqrt{5}}{2}$$

Checking the function value $f(u^*) > 0$ completes the argument. □

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

$$\Sigma^2 = \frac{c^2\gamma^2 n}{\rho^2\Sigma^4}\int_{-\infty}^{0} x^2 f(x - r/\gamma)dx + \frac{c^2\gamma^2 n}{\rho^2\Sigma^4}\int_{0}^{\infty} x^2 f(x + r/\gamma)dx,$$

where $f(x)$ denotes the price density. The RHS of the condition above is the variance of a distribution where the middle range $[-r/\gamma, r/\gamma]$ of beliefs has been removed, and the tails $(-\infty, -r/\gamma) \cup (r/\gamma, \infty)$ have been shifted inwards toward the origin. Note that the factor γ follows from the following rewriting of demand

$$x_t^i = \begin{cases} \gamma\frac{(p_i - p_{i-1}) - r/\gamma}{\rho\Sigma^2}, & \text{if } (p_i - p_{i-1}) > r/\gamma \\ \gamma\frac{(p_i - p_{i-1}) + r/\gamma}{\rho\Sigma^2}, & \text{if } (p_i - p_{i-1}) < -r/\gamma \\ 0, & \text{otherwise} \end{cases}$$

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

$$g(r, \Sigma) = \frac{c^2\gamma^2 n}{\rho^2}\int_{-\infty}^{0} x^2 f(x - r/\gamma)dx + \frac{c^2\gamma^2 n}{\rho^2}\int_{0}^{\infty} x^2 f(x + r/\gamma)dx - \Sigma^6$$

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

$$g(r, \Sigma) = \frac{c^2\gamma^2 n}{\rho^2}\int_{-\infty}^{-r/\gamma} (u + r/\gamma)^2 f(u)du + \frac{c^2\gamma^2 n}{\rho^2}\int_{r/\gamma}^{\infty} (u - r/\gamma)^2 f(u)du - \Sigma^6$$

and applying the Leibniz integral rule, we have that

$$\frac{\partial g}{\partial r} = \frac{c^2\gamma n}{\rho^2}\int_{-\infty}^{-r/\gamma} 2(u + r/\gamma)f(u)du - \frac{c^2\gamma n}{\rho^2}\int_{r/\gamma}^{\infty} 2(u - r/\gamma)f(u)du < 0$$

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

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

But because the derivative is evaluated near $r = 0$, we know that the equilibrium $\Sigma^2 \approx c\gamma\sqrt{n}/\rho$. Plugging this 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

$$\Sigma^2 = \frac{c^2\gamma^2 n}{\rho^2\Sigma^4}\int_{-\bar{b}\rho\Sigma^2/\gamma}^{\bar{b}\rho\Sigma^2/\gamma} x^2 f(x)dx + c^2 n\int_{\bar{b}\rho\Sigma^2/\gamma}^{\infty} \bar{b}^2 f(x)dx + c^2 n\int_{-\infty}^{-\bar{b}\rho\Sigma^2/\gamma} \bar{b}^2 f(x)dx$$

where $f(x)$ denotes the price density. 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

$$x_t^i = \begin{cases} \bar{b}, & \text{if } (p_i - p_{i-1}) > \bar{b}\rho\Sigma^2/\gamma \\ -\bar{b}, & \text{if } (p_i - p_{i-1}) < -\bar{b}\rho\Sigma^2/\gamma \\ \frac{\gamma(p_i - p_{i-1})}{\rho\Sigma^2}, & \text{otherwise} \end{cases}$$

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

$$g(\bar{b}, \Sigma) = \frac{c^2\gamma^2 n}{\rho^2}\int_{-\bar{b}\rho\Sigma^2/\gamma}^{\bar{b}\rho\Sigma^2/\gamma} x^2 f(x)dx + c^2 n\Sigma^4\int_{\bar{b}\rho\Sigma^2/\gamma}^{\infty} \bar{b}^2 f(x)dx + c^2 n\Sigma^4\int_{-\infty}^{-\bar{b}\rho\Sigma^2/\gamma} \bar{b}^2 f(x)dx - \Sigma^6$$

and we apply the Leibniz integral rule.

$$\frac{\partial g}{\partial \bar{b}} = c^2 n\Sigma^4\left[\int_{\bar{b}\rho\Sigma^2/\gamma}^{\infty} 2\bar{b}f(x)dx + \int_{-\infty}^{-\bar{b}\rho\Sigma^2/\gamma} 2\bar{b}f(x)dx\right] > 0$$

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

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

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 (and, hence, the truncation of the distribution by $\bar{b}$ plays little role). 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 $\Sigma^2 \approx c\gamma\sqrt{n}/\rho$. Plugging this in,

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

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

$$\varepsilon \equiv 4c^2 n \Sigma^3 \left[\int_{\bar{b}\rho\Sigma^2/\gamma}^{\infty} \bar{b}^2 f(x)dx + \int_{-\infty}^{-\bar{b}\rho\Sigma^2/\gamma} \bar{b}^2 f(x)dx \right] \tag{10}$$

The entire term (10) tends to zero because

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

where the inequality holds because $x \geq \bar{b}\rho\Sigma^2/\gamma$. The same argument applies to the second term on the RHS of (10), except now $x \leq -\bar{b}\rho\Sigma^2/\gamma$. □

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.

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

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.

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.

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.

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.