Experience Effects Explain Five Facts about Beliefs and Portfolios

Andrés Carvajal* Keisuke Teeple†

May 30, 2025

Abstract

Giglio et al. (2021) uncover a series of regular features related to financial investors' beliefs and portfolios; this paper proposes a model that incorporates these features. In our model, investors trade infrequently, and when forming expectations, they overweigh the information of previous periods when they traded. As a result, our cross-section of traders contains pessimists (optimists) who recall bear (bull) markets, as well as the youth (elderly) who have fewer (more) data points in their samples. As required by Giglio et al. (2021), expected returns and the subjective probability of rare disasters are negatively related in our model.

JEL classification: G41, G14, D84, D91
Keywords: Non-Bayesian models, retail traders, heavy tails, experience effects

* University of California Davis and EPGE–FGV, 1 Shields Ave Davis, California 95616, USA, acarvajal@ucdavis.edu.
† University of Waterloo, 200 University Ave West Waterloo, Ontario N2L3G1, Canada, k2teeple@uwaterloo.ca.
Teeple received financial support for this research from SSHRC Explore and Insight Development Grants.

The starting point of this paper is Giglio et al. (2021), who use a survey administered to retail investors to pin down five facts about beliefs and portfolios. The authors suggest that these facts may be used as guidance for the design of future financial models.

1. Beliefs are reflected in portfolio allocations. Furthermore, the sensitivity of portfolios to beliefs is weaker than what is predicted by frictionless models.
2. Beliefs affect the direction and magnitude, but not timing, of trades.
3. Beliefs are characterized by persistent heterogeneity. Demographics do not explain optimism versus pessimism.
4. Expected GDP growth and expected returns are positively related.
5. Expected returns and the subjective probability of rare disasters are negatively related.

While not listed as one of their main findings, we additionally use a sixth fact from their findings to guide our theory: older individuals have tighter subjective distributions over future stock returns.[1] We then build a model of retail investors subject to experience effects that is consistent with, and in a large part guided by, these six facts. Consider the following variable: the subjective expected return of a stock. As required by fact three, there are pessimists (optimists) who believe that this variable is negative (positive). As required by fact six, the distribution over subjective expected returns tightens with age. The visual being suggested by the combination of these two facts is depicted in Figure 1. In the figure, μ^i denotes the subjective expected return and k is a measure of investor age.

Notice that the second requirement of fact three—that demographics do not explain optimism versus pessimism—is satisfied by Figure 1. That is, age has no explanatory power in predicting whether a given investor in this cross-section is a pessimist or optimist. Next consider fact two: that beliefs affect the portfolio, but are orthogonal to trade timing. Motivated by this fact, we build a model where the timing of investor trades is an exogenous random variable unrelated to their beliefs (beliefs will be formed endogenously).[2] More formally, we independently draw n investors from the cross-section of beliefs in Figure 1 to trade in any given period. Taken quite literally, these retail investors login to their brokerage account (or call their broker) at random times.

What creates the motive to trade for these retail investors? On top of the exogenous trade timing assumption, we assume that traders use their experiences on—and only on—past active trading days to form their beliefs.[3] Consequently, those who previously experienced bear (bull) markets

[1] This fact can be found in the Online Appendix A.8 of Giglio et al. (2021).

[2] In fact, Giglio et al. (2021) suggest a model with exogenous trade probabilities.

[3] Using survey evidence, Jiang et al. (2023) show that experiences explain a large part of the cross-sectional dispersion in beliefs.

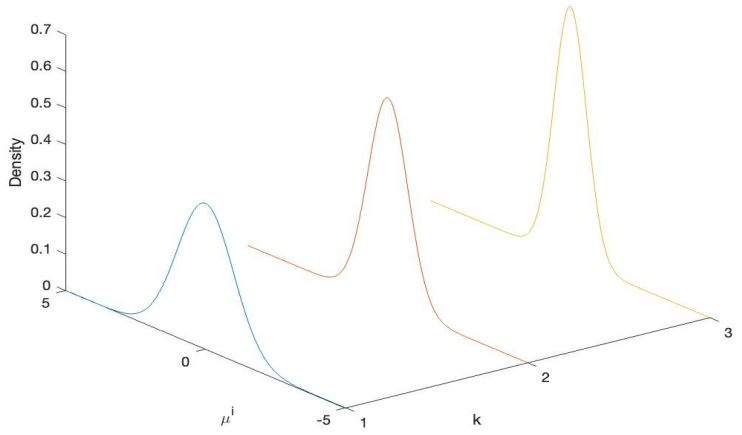

Figure 1: Cross-Section over Subjective Expected Returns

become pessimists (optimists), leading to an endogenous cross-sectional distribution over beliefs like the one shown in Figure 1. In support of this setup, Andersen et al. (2019) show that experiences on active trading 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, *experiencing* the gain or loss seems to impact individuals more profoundly than passively observing returns.

An equilibrium is established in the following sense. We begin with a history of prices drawn independently from some price distribution, which forms the basis for traders' beliefs. Based only on experiences from days on which they were actively trading, traders apply Bayes' rule. Based on these posteriors, a subset of traders trade; and based on trades, prices adjust. 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.

As a benchmark, we first investigate an economy with a static *representative-age* trader. Visually, we integrate over the k-dimension in Figure 1. In this representative-age setting, we show that the Normal distribution is the unique equilibrium price distribution when traders have mean-variance utility. Importantly, this utility assumption gives our model a linear structure: demands are linear in past price observations due to the mean-variance assumption, and prices are linear in demand because we focus on linear equilibria. 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 family, the Normal distribution is the only distribution with a finite variance.

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

We then move on to an investigation of our full economy with a rich cross-section over both beliefs and age. Like in the representative-age case, there are pessimists and optimists. But when we formally model traders learning from their experiences, young traders (those who have traded less, leading to dispersed beliefs) and old traders (those who have traded more, leading to concentrated beliefs) arise based on their trade history. Unlike in the representative-age setting, where belief dispersion is constant across the cross-section and equilibrium prices are Normal, in this generalized setting, there is a mixture of belief dispersion across demographics which leads to excess kurtosis in equilibrium prices.[5] The intuition is as follows: extreme prices result from the small chance that all traders trading on a given day are young.

These results allow us to address the three remaining facts from Giglio et al. (2021). It turns out that facts one (beliefs reflected in portfolios) and four (relationship between GDP growth and returns) are immediate in our setup.[6] So we focus on fact five (relationship between returns and rare disasters). To understand why this relationship holds in our model, we make two observations in the context of Figure 1. First, what are plotted are subjective expected *returns*; we assume that traders hold correct beliefs over all other price moments. Hence, each trader drawn in Figure 1 holds beliefs over an entire price distribution, simply shifted by the value μ^i. Second, we claim that our model is characterized by rare disasters because of the *additional* probability of tail events above and beyond that of the Normal distribution. Because our equilibrium price distribution is characterized by endogenously heavy tails, this distance from the Normal distribution is nonzero. Altogether, pessimists (optimists) have a larger (smaller) probability of such tail events in their shifted, subjective distribution over returns.

1 Related Literature

Many empirical studies support the use of past personal experiences when forming beliefs: 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.

[5] Alternative theoretical explanations for excess price kurtosis include heavy tails in wealth by Gabaix et al. (2006), imitation among traders by Cont and Bouchaud (2000), and leverage effects by Thurner et al. (2012).

[6] For fact one, we model mean-variance traders so beliefs are *linearly* reflected in portfolios; furthermore, transaction costs weaken the relationship between portfolios and beliefs. For fact four, our baseline setting has zero GDP growth. However, Appendix A shows that, whenever our model has positive GDP growth, GDP growth is positively related to returns.

An important set of theoretical models use *memory* to capture individuals' experiences and their effects on beliefs. For example, in Nagel and Xu (2022), a representative agent has a fading memory in the sense that a data point's influence on beliefs gradually fades over time. 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. Most recently, Azeredo da Silveira et al. (2024) use Shannon mutual information to model memory costs, leading to idiosyncratic random variation in forecasts. 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.[7]

While most of these aforementioned papers assume (or derive) memory models that emphasize *recent* observations, another literature emphasizes past observations experienced in a similar *context* as that of the current decision maker. According to the textbook treatment of associative memory from Kahana (2012), contextual stimuli act as cues that trigger recall of similar past experiences. In our model, the trader finds herself in a particular context: being active in financial markets. She recalls all past experiences within that same context, and emphasizes such data points when forming future expectations. Similar to Bordalo et al. (2020), who formalize context-dependent memory in an economic setting, we arrive at an expected return formula that is a weighted average of two terms: the expected return conditional on the relevant context and the unconditional average past returns (reflecting out-of-context memories):

$$\mathbb{E}[r_{t+1}] = \gamma \mathbb{E}[r_{t+1} \mid \text{context}] + (1-\gamma)\overline{r}$$

Because we assume that traders trade at random times, there are also connections between our work and that which attempts to endogenize noise trading. Further emphasizing this connection, recent evidence such as that proposed in Peress and Schmidt (2021) indicate that noise traders *are* retail investors (as opposed to institutional investors). Most relevant to our paper within this literature is Admati and Pfleiderer (1988), who model noise traders with exogenous positions—potentially reflecting the liquidity needs of their clients—but who can coordinate the timing of their liquidation. We assume the opposite: traders with endogenous positions trade at exogenous times. Other relevant papers include Palomino (1996) and De Long et al. (1990b), who model noise traders as those with stochastic beliefs, and Wang (2010), who models noise traders as those

[7] More generally, our paper is related to the large literature on extrapolation. For example, see De Jong et al. (1990a).

who trade on noise as if it were information.

The remainder of the paper is organized as follows. Section 2 sets up the model and defines our equilibrium concept. Section 3 investigates two consequences of experiential learning: overlapping experiences (across traders) and expanded recollection (within traders). Section 4 characterizes the equilibrium both in the representative-age trader case and in the case of the full model. Section 5 proves the excess kurtosis result and relates it back to rare disasters, and Section 6 concludes.

2 Model

2.1 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.[8] These individuals should be thought of as retail investors who use technical analysis before placing market orders.[9] Suppose that there is a history of past price realizations—even at the beginning of the model ($t = 0$)—and that these past realizations, $(p_{t+1-s} - p_{t-s})_{s\geq 1}$, were drawn independently from a distribution F. 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

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

where $\rho > 0$ denotes the risk aversion parameter, x_t denotes the trader's position, and $\tau > 0$ denotes the transaction cost parameter.[10] Transaction costs are assumed to be quadratic and deterministic. The reason that the expectation in (1) has a superscript i is because we assume that traders do not know the price drift; instead, they will solve an idiosyncratic learning problem to estimate this object.[11] Demand is chosen to maximize utility,

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

where Σ^2 is the variance of distribution F. Consider fact one from Giglio et al. (2021): while portfolios reflect beliefs, their sensitivity to beliefs is smaller than that predicted by a standard mean-variance model. These authors suggest several interpretations including capital gains taxes

[8] See Appendix A for an extension with positive dividends. Also in the background, there is a zero-interest riskless asset, against which the risky asset is traded.

[9] 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).

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

[11] They do, however, have rational conjectures about the price variance.

and low confidence. Here, this fact follows from quadratic transaction costs τ.

While traders choose endogenous quantities to maximize (1), we assume that they only trade at random *times*.[12] This modeling assumption is motivated by fact number two from Giglio et al. (2021) that beliefs do not predict *when* investors trade, only the direction and magnitude of trades. In fact, these authors suggest that "one parsimonious way to model such behavior is to introduce infrequent random trading, whereby an agent is selected at random based on a memory-less distribution to have the possibility of trading in a given period." (To support this assumption empirically, 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. Note that we are silent on the reason for infrequent trading. However, much of the research has focused on either attention/mental costs associated with monitoring and re-optimizing portfolio positions or fixed transaction costs associated with executing trades.[13])

To model experience effects, we assume that traders only recall past periods where they were actively trading; this makes each trader unique and believe that price increments $(p_{t+1} - p_t)$ have mean μ_t^i. As mentioned above, this assumption is consistent with empirical evidence from Andersen et al. (2019) and Strahilevitz et al. (2011), as well as with the context-dependent memory model of Bordalo et al. (2020).[14] For the purpose of setting up the model, we simply assume that traders are *frequentist* over these past experienced price realizations.[15] Formally, we denote the set of traders chosen to trade in period t by $\mathcal{A}_t$, and assume each i believes that the price drift is equal to

$$\mu_t^i = \frac{1}{|\mathbb{K}_t^i|} \sum_{s \in \mathbb{K}_t^i} (p_{s+1} - p_s) \tag{3}$$

where $\mathbb{K}_t^i = \{s < t \mid i \in \mathcal{A}_s\}$ denotes the set of time periods in which the trader has previously traded. We will refer to prices corresponding to this set of time periods as the experienced, or recalled, prices. We denote the cross-sectional distribution of beliefs as

$$G_t^N(m) = \frac{1}{N} \sum_{i=1}^{N} 1_{\mu_t^i \leq m}$$

where N denotes the total number of traders in the economy, which we will let tend to ∞ in our

[12] In fact, this is the only exogenous random variable in the model.
[13] For example, see Abel et al. (2007) and Duffie and Sun (1990).
[14] See also Wachter and Kahana (2024).
[15] For a more microfounded Bayesian posterior based on recent models of imperfect memory, see Section 2.4.

arguments. We assume that $n > 1$ traders are independently drawn to trade in any given period,[16] so their beliefs are a sample of size n from this cross-section G_t^N. 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}. Importantly, they do not hold the asset between two randomly chosen trading days.[17]

When traders recall active trading days, the cross-section of traders continuously learns. To make our model stationary, we assume that traders who have traded $K > 1$ times die and are subsequently replaced by the same mass of newborn traders.[18] To model newborn traders, we assume that there is a reserve of traders who have yet to trade in any period. Due to the lack of past experience, we simply set $\mu_t^i = 0$ for such traders.[19] When replacing deceased traders, we assume that the same mass of traders from this reserve joins the cross-section G_t^N and recalls the period in which they join. It is *as-if* they trade, but they trade zero quantities because $\mu_t^i = 0$ and they trade only for the purpose of recalling this period and updating beliefs. This is nothing more than an accounting exercise, making G_t^N the *conditional* cross-section, conditional on having previously traded at least once (and, hence, having non-trivial beliefs).

2.2 Market Maker

We assume that there is a market maker who sets prices in such a way that zero expected profits are made. This zero profit condition can be written

$$\mathbb{E}\left[\frac{\tau}{2}\sum_{i=1}^{N} x^i(p_t)^2 - (p_{t+1} - p_t)\sum_{i=1}^{N} x^i(p_t)\right] = 0 \tag{4}$$

As is well-documented in studies like Sofianos (1995), market makers make financial gains through transaction costs, corresponding to the first term in (4), and financial losses via inventory, corresponding to the second term in (4). Because the market maker must hold the position $-\sum_{i=1}^{N} x^i(p_t)$ to effectively clear markets, returns from this inventory position are $-(p_{t+1} - p_t)\sum_{i=1}^{N} x^i(p_t)$. We emphasize that expectations in (4) are taken over *demands*, because beliefs μ_t^i are independently drawn from the cross-section, not prices, which are set by the market maker. We abstract from the more difficult problem of setting optimal spreads in this paper, and τ is taken as a parameter

[16] More formally, the probability space $(\Omega, \mathcal{F}, P)$ consists of the product space $\Omega = \mathbb{R}^n$, $\mathcal{F}$ is the product of Borel σ-algebras, and P is the measure defined by n independent draws from the distribution G_t^N as $N \to \infty$.

[17] Because traders trade on random days, we exclude the case where indexes of recalled prices are measurable with respect to the realization. This, for example, rules out extreme prices being recalled by more traders.

[18] While this baseline setting is the simplest one, additional specifications can be appended to the model for realism. For example, in Section 3.1, we consider a fixed probability that each trader dies in any given period—regardless of how many times they have traded—and all of our results go through.

[19] As Lemma 2 will show, this assumption is also consistent with such traders simply being Bayesian.

of the model.[20]

Finally, in the spirit of Kyle (1985), we restrict our equilibrium price-setting rules to those that are linear in excess demand:

$$p_{t+1} = p_t + c \sum_{i=1}^{N} x^i(p_t) \tag{5}$$

for some slope c that will be determined in equilibrium.[21] A market maker that maps excess demand into higher prices and excess supply into lower prices 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.

2.3 Equilibrium

With this setup, an equilibrium is a price distribution F such that,[22] given any initial history $\{p_{s+1} - p_s \mid s = -1, -2, \ldots\}$ of i.i.d. realizations of it, for all $t \geq 1$: (a) price increments $p_{t+1} - p_t$ given by (5) are distributed according to F; (b) individual demands $x^i(p_t)$ are given by (2) if $i \in \mathcal{A}_t$, and $x^i(p_t) = 0$ otherwise; (c) beliefs μ_t^i are given by (3); and (d) expected profits for the market maker are zero, (4).

A visual depiction of the equilibrium concept is shown in Figure 2. On the LHS of this diagram, we have the cross-section over μ^i, similar to Figure 1. On the RHS of this diagram, we have the distribution of prices. Both are endogenous, steady-state objects.

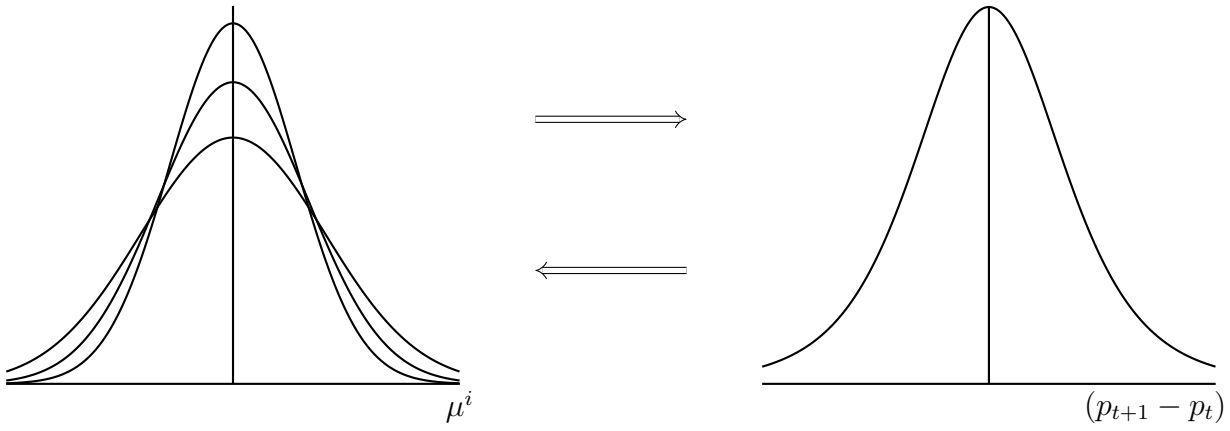

Figure 2: Equilibrium Cross-Section and Price Distribution

[20] Teeple (2023) solves for the optimal spread in a similar setting but with no experience effects.

[21] Note that this rule captures the price impact of the entire round trip—trading at p_t then unwinding that position at p_{t+1}—in period $(t+1)$. Current price impact is captured in τ.

[22] In Appendix B, we explore the implications of an alternative limit order model (instead of a market order model).

First consider the arrow pointing right (how beliefs affect prices). Beliefs enter into demands via (1), which get compounded into prices via (5). Because there is some chance the trader drawn from the cross-section is young (old) with a dispersed (concentrated) belief distribution, the entire cross-section is a mathematical *mixture* over these "generations." That is, we mix over the young and old when drawing traders to generate prices. Next consider the arrow pointing left (how prices affect beliefs). Random past prices are recalled, then beliefs are formed according to (3). If the trader is young (old), they take an average over few (many) past prices when forming beliefs; such a belief μ^i then corresponds to the dispersed (concentrated) curve on the LHS of Figure 2.

We have already assumed that markets are competitive in the standard sense that traders do not optimize over own price impact. This was implicit in the optimization problem (1). Furthermore, we assume that markets are competitive so that traders do not use their knowledge of their own demand to better forecast prices. Instead, traders simply consider the statistical properties of prices, which are inherited from (i.e. measurable with respect to) independently-drawn demands from the cross-section.

2.4 Connection to Memory Models

In this section, we provide a memory-based microfoundation for the frequentist formula (3). Following Nagel and Xu (2022), we assume that traders apply weights α_s^i to past observations $(p_{t+1-s} - p_{t-s})_{s\geq 1}$ when applying a memory-constrained version of Bayes' Rule:

$$f(\mu|p_t - p_{t-1}, , ..., p_{t-T+1} - p_{t-T}) \propto \prod_{s=1}^{T} f(p_{t+1-s} - p_{t-s}|\mu)^{\alpha_s^i} \tag{6}$$

where T denotes the length of the price history and μ denotes the price drift parameter, which traders learn.[23] Note that, when $\alpha_s^i = 1$ for all s, this reduces to Bayesian updating with a flat prior because past price observations are assumed to be independent (and, hence, densities in (6) are multiplied). When α_s^i fades over time, we recover fading memory from Nagel and Xu (2022). 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 C.

Lemma 1 (Posterior consistency)**.** *Say that traders calculate their posterior at time t according*

[23] An expression similar to (6) can be derived from minimization of a cost function that involves Kullback-Leibler divergence from the Bayesian posterior and Shannon entropy. See Prat-Carrabin et al. (2021).

to memory-constrained Bayes' Rule (6)*, with weights given by*

$$\alpha_s^i = \begin{cases} \gamma \frac{T}{|\mathbb{K}_t^i|} & \text{for } s \in \mathbb{K}_t^i \\ (1-\gamma)\frac{T}{T-|\mathbb{K}_t^i|} & \text{otherwise} \end{cases} \tag{7}$$

and that $|\mathbb{K}_t^i| < \infty$*. As the history length* $T \to \infty$*, their posterior mean approaches*

$$\mu_t^i = \frac{\gamma}{|\mathbb{K}_t^i|} \sum_{s \in \mathbb{K}_t^i} (p_s - p_{s-1}) + (1-\gamma)\mu$$

almost surely, assuming the price history is independent and irrespective of the price distribution.

Note that μ denotes the price drift, $0 < \gamma \leq 1$ denotes the memory strength parameter, and $\mathbb{K}_t^i$ denotes the set of recalled periods for trader i. As extreme cases, the frequentist corresponds to $\gamma = 1$, and $\gamma = 0$ implies that there are no memory concerns so traders eventually learn the true price drift. When $0 < \gamma < 1$, consider the total weight allocated to non-recalled periods in (7), $\sum_{s \notin \mathbb{K}_t^i} \alpha_s^i = (1-\gamma)T$. Adding this to the sum of weights allocated to recalled periods, $\sum_{s \in \mathbb{K}_t^i} \alpha_s^i = \gamma T$, 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 evenly across recalled periods, and the rest evenly across all other 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 (7) are *as-if* the trader not only observes an infinite number of independent non-recalled prices (this makes them "act Bayesian" over non-recalled periods), but also an infinite number of repeated observations of $(p_s - p_{s-1})$ for each $s \in \mathbb{K}_t^i$. As the history length $T \to \infty$, they become sure that their posterior mean equals the μ_t^i defined in Lemma 1 (i.e. their posterior variance tends to zero).[24] Moving forward, we employ this formula from Lemma 1, which nests the frequentist case (3).

With this setup, we can address fact three from Giglio et al. (2021): beliefs are characterized by persistent heterogeneity. Beliefs from Lemma 1 have a positive covariance over time because, mathematically,

$$\text{Cov}(\mu_t^i, \mu_{t+1}^i) = \text{Cov}\left(\frac{\gamma}{|\mathbb{K}_t^i|} \sum_{s \in \mathbb{K}_t^i} (p_s - p_{s-1}), \frac{\gamma}{|\mathbb{K}_{t+1}^i|} \sum_{s \in \mathbb{K}_{t+1}^i} (p_s - p_{s-1}) \right) > 0$$

for a trader i randomly drawn from the cross-section. This inequality holds true, even setting aside

[24] 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.

the fact that beliefs are persistent simply because traders trade infrequently. That is, even if we focus on a period $t+1$ where the trader trades (and learns), price realizations are independent and, hence, the covariance can be shown to be positive.

2.5 Zero Profit for the Market Maker

In this section, we take a small step towards solving our model. In particular, we find the implications of the zero profit condition (4). We begin with the linear price-setting rule (5). Because equilibrium requires that the past price distribution matches the current one, the past price *mean* must equal the current one. With n traders active in any given period, taking an expectation of (5), plugging in demands, and plugging in beliefs from Lemma 1, we get

$$\mu = cn\mathbb{E}[x^i(p_t)] = \frac{cn\mu}{\rho\Sigma^2 + \tau} \tag{8}$$

Important to this formula (8) is that $\mathbb{E}[\mu^i] = \mu$, where the expectation is cross-sectional. That is, while there is belief dispersion characterized in Lemma 1, traders have—on average—correct beliefs. Hence, the formula (8) does not depend on the number of past trading periods, $|\mathbb{K}_t^i|$, or the memory strength, γ. Using a similar calculation but with the equilibrium price *variance*, we can solve for Σ^2. Plugging this into (8), we show that any equilibrium price must satisfy $\mu = 0$. The intuition behind zero equilibrium drift is: positive drift generates excess demand; this, in turn, requires a positive asset supply. Appendix A contains an extension with positive exogenous asset supply.

Lemma 2 (Zero profits)**.** *Any equilibrium price distribution must satisfy $\mu = 0$. Furthermore, if $\mathbb{E}[x^i(p_t)^2] > 0$, then $c = \tau/2$.*

To understand the condition $c = \tau/2$, zero expected profits for the market maker is equivalent to zero expected profits for the average active trader by rearranging (4):

$$\mathbb{E}\left[(p_{t+1} - p_t)\frac{1}{n}\sum_{i=1}^{N} x^i(p_t) - \frac{\tau}{2}\frac{1}{n}\sum_{i=1}^{N} x^i(p_t)^2\right] = 0$$

When demands are independently drawn and $\mu = 0$, the second term can be rewritten as

$$\mathbb{E}\left[(p_{t+1} - p_t)\frac{1}{n}\sum_{i=1}^{N} x^i(p_t) - \frac{\tau}{2}\frac{1}{n}\left(\sum_{i=1}^{N} x^i(p_t)\right)^2\right] = 0$$

which is implied by

$$p_{t+1} = p_t + \frac{\tau}{2} \sum_{i=1}^{N} x^i(p_t)$$

so that this linear price-setting rule now has the zero profit condition "built in" in the following sense. Price one period ahead (LHS) must equal price in the current period, including total price impact (RHS). The only requirement for this lemma is *some* belief dispersion in the sense of $\mathbb{E}[x^i(p_t)^2] > 0$ (or, equivalently, nontrivial prices, $\Sigma^2 > 0$).

With this accounting exercise complete, we can address fact four from Giglio et al. (2021): expected GDP growth and expected returns are positively related. GDP in our model is calculated using income, which we denote w^i_{t+1}. Formally, mean-variance utility in (1) is formed over w^i_{t+1}, where it is defined

$$w^i_{t+1} = (p_{t+1} - p_t)x^i_t - \frac{\tau}{2}(x^i_t)^2 + w^i_t$$

Furthermore, we must take into account the balance sheet of the market maker:

$$\mathcal{W}_{t+1} = \frac{\tau}{2} \sum_{i=1}^{N} (x^i_t)^2 - (p_{t+1} - p_t) \sum_{i=1}^{N} x^i_t + \mathcal{W}_t$$

where $\mathcal{W}_{t+1}$ denotes market maker wealth. Then GDP growth equals $\sum_{i=1}^{N}(w^i_{t+1} - w^i_t) + (\mathcal{W}_{t+1} - \mathcal{W}_t)$, which equals zero in our baseline setting. Appendix A extends our baseline setting with positive asset supply; this generates positive GDP growth and, in this extension, we verify that expected GDP growth and expected returns are positively related in the cross-section of traders.

3 Experiential Learning

In this section, we investigate two consequences of experiential learning. First, traders have an expanded set of recalled experiences, adding another dimension of heterogeneity—beyond pessimist and optimist—to our cross-section. In other words, some traders recall more periods while others recall fewer. Second, traders' beliefs overlap because n traders observe (and learn) the same price observation each period. We begin with the former.

3.1 Expanded Recollection

For intuition, first consider the homogeneous case, where all traders recall the same number of periods (and that number is greater than one). If multiple events are recalled, good days and bad days begin to cancel each other out within that recalled set. Hence, traders with a larger set of past

experienced events hold less dispersed, or tighter, beliefs. As Theorem 1 will show, this lowers the equilibrium price variance, Σ^2. See Figure 1 for the cross-sectional distribution over μ^i for traders who recall one, two, or three periods (memory parameter γ is set to one). This graphic is consistent with fact six from Giglio et al. (2021), who find that older individuals' subjective distributions over future stock returns have lower standard deviations (than those of younger individuals).

Now consider adding heterogeneity in the number of recalled periods. When drawing a trader from the cross-section, there is some chance she is inexperienced (few recalled periods, with dispersed beliefs) and some chance she is experienced (many recalled periods, with concentrated beliefs). Hence, the traders mainly responsible for generating price volatility will be the inexperienced *youth*.

The steady-state cross-section of traders is characterized by a vector $(\mathcal{P}_1, ..., \mathcal{P}_{K-1})$, which represents the proportion of the population that recalls $k \in \{1, ..., K-1\}$ days. We first analyze the case where the total number of traders, N, is finite, then take limits. When $N > n$ is finite, the law of motion for $\mathcal{P}_k$ is given by

$$\mathcal{P}_k^{t+1} = \begin{cases} \mathcal{P}_k^t \left(1 - \frac{n}{N}\right) + \mathcal{P}_{K-1}^t \frac{n}{N}, \text{ if } k = 1 \\ \\ \mathcal{P}_k^t \left(1 - \frac{n}{N}\right) + \mathcal{P}_{k-1}^t \frac{n}{N}, \text{ if } 1 < k < K \end{cases}$$

For any k, the proportion decreases by a factor n/N due to active trading, learning, and traders consequently leaving group $\mathcal{P}_k^t$ for group $\mathcal{P}_{k+1}^{t+1}$. When $k > 1$, the proportion increases by $\mathcal{P}_{k-1}^t n/N$ for the same reason (these are the inflows corresponding to aforementioned outflows). When $k = 1$, the proportion increases by the amount $\mathcal{P}_{K-1}^t n/N$ due to the introduction of newborn traders (recall that this mass must match that of deceased traders). Imposing that proportions must sum to one, the steady state of such a system is uniform:

$$(\mathcal{P}_1, ..., \mathcal{P}_{K-1}) = \left(\frac{1}{K-1}, ..., \frac{1}{K-1}\right)$$

and, importantly, is independent of N.[25] While this formulation may seem overly general to establish this rather-intuitive steady state, the same formula allows us to add more realistic properties

[25] More formally, we can begin with the *number* of traders in each slice of the cross-section, $(\mathcal{N}_1, ..., \mathcal{N}_{K-1})$ where $\sum_{k=1}^{K-1} \mathcal{N}_k = N$. The law of motion applies to the *expected* number of traders,

$$\mathbb{E}[\mathcal{N}_k^{t+1}] = \begin{cases} \mathbb{E}[\mathcal{N}_k^t]\left(1 - \frac{n}{N}\right) + \mathbb{E}[\mathcal{N}_{K-1}^t]\frac{n}{N}, \text{ if } k = 1 \\ \mathbb{E}[\mathcal{N}_k^t]\left(1 - \frac{n}{N}\right) + \mathbb{E}[\mathcal{N}_{k-1}^t]\frac{n}{N}, \text{ if } 1 < k < K \end{cases}$$

Then divide the resulting steady state,

$$(\mathbb{E}[\mathcal{N}_1], ..., \mathbb{E}[\mathcal{N}_{K-1}]) = \left(\tfrac{N}{K-1}, ..., \tfrac{N}{K-1}\right),$$

of the cross-section such as a constant probability of death, irrespective of the number of recalled periods. In this case, the law of motion for $\mathcal{P}_k$ (when $N > n + d$) is given by:

$$\mathcal{P}_k^{t+1} = \begin{cases} \mathcal{P}_k^t \left(1 - \frac{d}{N} - \frac{n}{N}\right) + \mathcal{P}_{K-1}^t \frac{n}{N} + \frac{d}{N}, \text{ if } k = 1 \\ \\ \mathcal{P}_k^t \left(1 - \frac{d}{N} - \frac{n}{N}\right) + \mathcal{P}_{k-1}^t \frac{n}{N}, \text{ if } 1 < k < K \end{cases}$$

where $d > 0$ traders die each period and are consequently replaced by d new traders. The steady state of such a system is geometric:

$$(\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3, ...) = \left(C, \frac{Cn}{d+n}, \frac{Cn^2}{(d+n)^2}, ...\right)$$

where C is a constant that makes the proportions sum to one.[26] To emphasize the fact that our upcoming results do not rely on such details of the cross-section distribution, we write proofs for any cross-sectional distribution $(\mathcal{P}_1, ..., \mathcal{P}_{K-1})$.

With this setup, we can address the second part of fact three from Giglio et al. (2021): demographics do not explain optimism versus pessimism. In our model, this translates to the fact that $\mathbb{E}[\mu^i | \text{having traded } k \text{ times}] = 0$ for all $k \in \{1, ..., K-1\}$, where the expectation is cross-sectional. This relates to the exogenous nature of the forces that shape our cross section, $(\mathcal{P}_1, ..., \mathcal{P}_{K-1})$, in the discussion above. To help illustrate this point, say that, instead of the two proposed steady state cross-sections suggested in this section, traders learned more when they became optimistic or exited financial markets when they became pessimistic. Then the third fact from Giglio et al. (2021) would not hold true in our model.

3.2 Overlapping Experiences

Note that, when $n > 1$ traders recall an active trading day, their beliefs are not mutually independent. In general, this problem of overlapping beliefs can make the model highly intractable. However, when the degree of overlap is limited in a particular sense, 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 $G_t^N(m)$ converges in distribution to some stationary limit, from which trader beliefs can be drawn.

Lemma 3 (Limited dependence)**.** *Fix a zero-mean price distribution, $\overline{F}$, and assume that the*

by N and send $N \to \infty$.

[26] $C = \frac{d}{d+n}\left(1 - \left[\frac{n}{d+n}\right]^{K-1}\right)^{-1}$.

proportion of traders who recall k periods in the economy is $\mathcal{P}_k$ for $k \in \{1, ..., K-1\}$. Then, as the number of traders $N \to \infty$, the empirical distribution over beliefs $G_t^N(m)$ converges in probability to its expected value, $G(m)$, for each m.[27]

The effect of fixing the price distribution (zero-mean has already been confirmed as a necessary condition for equilibrium in Lemma 2) is that Lemma 3 becomes an out-of-equilibrium statement. Independent draws from distribution $\overline{F}$ are each observed by n traders, who learn according to (6) and (7). Lemma 3 states that the cross-section converges to some stationary distribution despite the presence of traders whose beliefs overlap, albeit in a limited way.

For intuition, fix the time t and consider whether the weak law of large numbers holds (instead of whether the empirical distribution converges). By Chebyshev's Inequality,

$$\Pr(|\overline{\mu}_N - \mu| \geq \varepsilon) \leq \frac{\text{Var}(\overline{\mu}_N)}{\varepsilon^2} \tag{9}$$

where $\overline{\mu}_N$ denotes the average belief in the cross-section of size N, and μ denotes the actual price drift. In the standard i.i.d. case, the variance on the RHS of (9) collapses into an $\mathcal{O}(1/N)$ term and so the weak law holds. However, here, posterior means μ_t^i are not independent. Although there is some overlap in beliefs, each trader's belief can only overlap with $(K-1)(n-1)$ others' because they trade on $(K-1)$ days where there were $(n-1)$ other traders on each day. Hence the variance of the average belief in (9) decomposes into N variances and $N(K-1)(n-1)$ covariances, which all disappear due to the factor $1/N^2$ from the average as $N \to \infty$. This same idea is employed in the actual proof, although the empirical distribution, $\frac{1}{N}\sum_{i=1}^{N} 1_{\mu_t^i \leq m}$, is considered.

4 Equilibrium

Before characterizing the full equilibrium price distribution, we run the following exercise. We begin with an arbitrary steady-state cross-sectional distribution from Section 3.1, $(\mathcal{P}_1, ..., \mathcal{P}_{K-1})$. Because this distribution has already reached its steady state, we turn off any learning. In other words, we assume that beliefs no longer change based on new price observations. If this were the only assumption, it would not affect equilibrium prices because we began at the steady-state.

[27] Note that

$$G(m) \equiv \sum_{k=1}^{K-1} \mathcal{P}_k \overline{F}_k(m),$$

where $\overline{F}_k(m)$ denotes the distribution of the k-fold convolution of prices multiplied by γ/k, and $\mathcal{P}_k$ denotes the proportion of traders in the population who recall k periods.

However, we additionally assume that beliefs are no longer drawn from the mixture distribution

$$G(m) \equiv \sum_{k=1}^{K-1} \mathcal{P}_k F_k(m)$$

where $F_k(m)$ denotes the k-fold convolution of prices multiplied by γ/k. Note that this formulation of $F_k(m)$ follows from the functional form for posterior means derived in Lemma 1:

$$\mu_t^i = \frac{\gamma}{k} \sum_{s \in \mathbb{K}_t^i} (p_s - p_{s-1}) \implies F_k(m) = F(km/\gamma) *_k \tag{10}$$

where $*_k$ denotes the k-fold convolution.[28] Instead, for this section only, we assume that beliefs are drawn from $F_{\overline{k}}(m)$, where

$$\overline{k} = \left\lfloor \sum_{k=1}^{K-1} \mathcal{P}_k k \right\rfloor \tag{11}$$

so that, along this dimension of age (i.e. k), there is a *representative-age trader*. Heterogeneity in the sense of pessimist versus optimist still remains, however, which guarantees nonzero trade. The representative-age trader in (11) is defined using the expected age in the population, with a floor function taken to ensure that $\overline{k}$ is an integer.

This exercise is useful for two reasons. First, we can think of this setup as an approximation of the full model, which will be investigated in upcoming Sections 4.2 and 5. The approximation is particularly good when the distribution $(\mathcal{P}_1, ..., \mathcal{P}_{K-1})$ is close to being degenerate. Second, this exercise helps to theoretically disentangle the effects of the two dimensions of heterogeneity—pessimist/optimist and age—on equilibrium returns. As we show next, the representative-age trader will lead to Normally distributed prices as the unique equilibrium outcome, so any deviations from Normality in Sections 4.2 and 5 can be attributed to age effects.

4.1 Uniqueness

Consider the representative-age trader exercise described above. Because equilibrium requires that the past price distribution equal the current one, we now apply this logic to the equilibrium price *variance*. Taking the variance of both left- and right-hand sides of the linear rule (5),

$$\Sigma^2 = c^2 n \frac{\gamma^2 \Sigma^2 / \overline{k}}{(\rho \Sigma^2 + \tau)^2} \tag{12}$$

[28] In general, $\mathcal{C}X$ has distribution $F(x/\mathcal{C})$ when $\mathcal{C}$ is a constant and X is a random variable with distribution $F(x)$.

In the numerator of the RHS of (12), we are invoking Lemma 1 with $|\mathbb{K}_t^i| = \overline{k}$. The intuition behind the Σ^2 solving (12) is: 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 (5). With this, we are ready for our first main result.

Theorem 1 (Uniqueness)**.** *The $N(0, \Sigma^2)$ distribution, where Σ^2 solves* (12)*, is the unique equilibrium price distribution when (a) traders do not learn, and (b) beliefs μ^i are drawn from $F_{\overline{k}}(m)$ as defined in* (10) *and* (11)*.*

First notice that there is a trivial no-trade equilibrium that solves (12). In this equilibrium, prices are constant, traders do not trade, and the market maker makes no profits. Given the lack of fundamental value in our main model (although we include an extension with dividends and asset supply in Appendix A), this first equilibrium should not come as a surprise. In what follows, we will focus our attention on the second non-trivial ($\Sigma^2 > 0$) equilibrium.

When $n > 4\overline{k}\gamma^{-2}$, which we will refer to as "n being large enough," a second equilibrium appears, $\Sigma^2 = (c\gamma\sqrt{n/\overline{k}} - \tau)/\rho$. First, note that the potential third root of (12) is always negative. Second, note that our "large n" inequality guarantees that this variance is strictly positive. To understand the need for this restriction on n, consider the costs and benefits of trading in the rearranged version of (12):

$$\underbrace{\Sigma^2\rho + \tau}_{\text{Costs}} = \underbrace{c\gamma\sqrt{n/\overline{k}}}_{\text{Benefits}}$$

With larger n, γ, or $1/\overline{k}$, the price history is (or appears) more volatile, leading to larger belief dispersion and perceived gains from trading. Hence, the benefits of trading increase in n, γ, and $1/\overline{k}$. And while a larger c should intuitively have similar effects, $c = \tau/2$ by the zero profit condition (4) so costs and benefits both increase in c. The net effect is that c drops out of our qualifier $n > 4\overline{k}\gamma^{-2}$. Altogether, a sufficiently large benefit (or perceived benefit) of trading—in the sense of large n, γ, or $1/\overline{k}$—is required to maintain a nonzero Σ^2. The less-obvious point made in Theorem 1 is not the calculation of such moments but the *Normality* of equilibrium prices.

The following intuition for Theorem 1 makes use of a 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 are linear

in demand because we restrict our attention to linear equilibria. Altogether,

$$p_{t+1} - p_t = c \underbrace{\sum_{i \in \mathcal{A}_t} \frac{\overbrace{\frac{\gamma}{k} \sum_{s \in \mathbb{K}_t^i} p_s - p_{s-1}}^{\text{Linear in prices}}}{\rho \Sigma^2 + \tau}}_{\text{Linear in demand}}$$

Equilibrium requires that the past price distribution matches 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.

In summary, even controlling for age effects, traders find themselves in a *bad equilibrium*: it is past volatility that leads to belief dispersion and continued price volatility. While policy and normative implications are not the focus of the paper, we outline two policies to reduce price variance in Appendix D.[29] For the remainder of the paper, we remove the approximation that only the representative-age trader trades. That is, we consider learning and the entire cross-section over ages, $(\mathcal{P}_1, ..., \mathcal{P}_{K-1})$.

4.2 Existence

Before exploring properties of equilibrium prices, we confirm that an equilibrium price density exists. We reiterate that this exercise now considers the full model without any approximation, so Theorem 1 does not apply. Several points are worth mentioning prior to the result. While one might be able to "construct" such a distribution by taking moments of both left- and right-hand sides of the market maker's rule (5), this does not guarantee the existence of a distribution consistent with those moments.[30] Also, it is not obvious that the equilibrium distribution is absolutely continuous, as guaranteed in Theorem 2. The intuition is that convolution (summing demands) and mixture (across generations), the two operations imposed on the historical price density, both preserve absolute continuity.

Theorem 2 (Existence)**.** *Say that n is large enough in the sense that $n > 4\left(\gamma^2 \sum_{k=1}^{K-1} \mathcal{P}_k/k\right)^{-1}$. Then a symmetric equilibrium price density, f, exists.*

Note that the assumed inequality is the generalized version of the previous inequality for "enough traders," $n > 4\overline{k}\gamma^{-2}$, which guarantees the existence of an equilibrium $\Sigma^2 > 0$. Without this, the theorem would fail because, although the trivial equilibrium discussed in Section 4.1 exists, it

[29] We consider the full model, as opposed to the representative-age setup, in the policy exercises.

[30] For reference, see the well-studied Hamburger moment problem.

is not absolutely continuous. In the proof, we consider the market maker's rule as a continuous operator, $\mathcal{T}$, mapping a historical price density into the current one. The trick to showing that a fixed point of $\mathcal{T}$ exists is to restrict its domain to symmetric densities with variance $\Sigma^2 > 0$. Note that these two properties are preserved under the operator $\mathcal{T}$.[31] Then, a sufficient condition to invoke the required fixed point theorems is a property called tightness.[32] To understand why the domain, restricted in this way, is tight, take any sequence $(f_j)_{j=1}^{\infty}$ in the domain. By Markov's Inequality,

$$\Pr(|Z_j| \geq R) \leq \frac{\mathbb{E}[Z_j^2]}{R^2} = \frac{\Sigma^2}{R^2}$$

where Z_j denotes the random variable with density f_j and the equality follows from the fact that $\mathbb{E}[Z_j] = 0$. This line of reasoning proves tightness. Notice also that symmetry of the equilibrium density pins down all odd price moments (to zero). An interesting extension of the model, not explored here, would be to introduce loss aversion or some asymmetric financial friction that leads to nonzero skew in equilibrium prices.

5 Fact Five

5.1 Excess Kurtosis

Next we show that mixing distributions with different variances in the cross-section results in *excess kurtosis* in equilibrium prices.

Theorem 3 (Heavy tails)**.** *Let μ^4 denote the fourth central moment of prices. Assume that the cross-section $(\mathcal{P}_1, .., \mathcal{P}_{K-1})$ is nondegenerate and that n is large enough in the sense of Theorem 2. Then $\mu^4/(\Sigma^2)^2 > 3$ in the nontrivial equilibrium. Furthermore, in the nontrivial equilibrium,*

1. *Σ^2 is increasing in $\mathbb{E}_{\mathcal{P}}[1/X]$.*
2. *$\mu^4/(\Sigma^2)^2$ is increasing in $\mathrm{Var}_{\mathcal{P}}(1/X)$.*

Mathematically, heavy tails result from the mixing of random variables with differing variances. Intuitively, heavy tails result from the small chance that all traders drawn on a given day are young. The intuition that "mixing adds tails" is correct but incomplete. This only implies 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 3 can be understood as characterizing the net effect between these two competing forces (mixing and summing).

[31] Symmetry is preserved under convolution and mixture.

[32] A family of distributions is tight if, for every $\delta > 0$, there exists N such that $\Pr(|Z_j| \geq N) \leq \delta$.

To get a sense of how the proof of Theorem 3 works, we calculate the fourth moment of the market maker's rule (5), divide by $(\Sigma^2)^2$, and simplify terms until we get

$$\frac{\mu^4}{(\Sigma^2)^2} - 3 = \frac{\left(\frac{\mu^4}{(\Sigma^2)^2} - 3\right) \sum_{k=1}^{K-1} \frac{\mathcal{P}_k}{k^3} + 3 \sum_{k=1}^{K-1} \frac{\mathcal{P}_k}{k^2} - 3 \left(\sum_{k=1}^{K-1} \frac{\mathcal{P}_k}{k}\right)^2}{n \left(\sum_{k=1}^{K-1} \frac{\mathcal{P}_k}{k}\right)^2} \tag{13}$$

As intuition might suggest, as the number of traders trading in any given period, n, tends to infinity in (13), central limit theorems apply and the price kurtosis approaches that of the Normal distribution (three). Beyond this, equation (13) provides an intuitive breakdown of the effects present in Theorem 3. Consider the first term in the numerator of the RHS of (13); we denote this as the *historical effect* because it contains the value of historical excess kurtosis. Intuitively, prices can have excess kurtosis simply because historical prices had excess kurtosis. Note, however, that this self-fulfilling effect can go in the opposite direction: prices can have negative excess kurtosis because historical prices had negative excess kurtosis. Next consider the second and third terms in the numerator of the RHS of (13); we denote this as the *mixing effect*. Factoring out the number three, this term equals the cross-sectional variance of (one over) the number of recalled periods. When this term is zero, we are back in the case of Section 4.1 and Theorem 1 applies. However, when this term is nonzero, our intuition—that mixing adds tails—applies.

For the equilibrium excess kurtosis to inherit the sign of the cross-sectional variance (positive according to the conditions of Theorem 3), we require that the weight put on the historical effect is not too strong. Mathematically, we require that

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

which just so happens to be guaranteed by the "n large enough" condition.

For the sake of completeness, the two bullet points from Theorem 3 consider comparative statics of equilibrium price moments with respect to the cross-sectional distribution $(\mathcal{P}_1, .., \mathcal{P}_{K-1})$. 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 recalled periods. Hence $1/X$ is the inverse of age, which we refer to as the trader's *youth*. When this object is larger (smaller), the trader trades more (less). Condition 1 of Theorem 3 then says that more youthful economies are characterized by more volatility. And as intuition suggests, kurtosis in Condition 2 is increasing in the variance of the cross-section. These two conditions present additional empirically testable implications of our theory, beyond those verified in Giglio et al. (2021).

To illustrate Theorem 3 in an example, Table 1 contrasts the first ten moments of our nontrivial

equilibrium price distribution to those of a Normal distribution. Parameters are set to $(\mathcal{P}_1, ..., \mathcal{P}_5) = (0.2, ..., 0.2)$, $\gamma = 1$, $\rho = 2$, $\tau = 0.1$, $c = 1$, and $n = 10$.

	Equilibrium	**Normal**
$\mathbb{E}[(p_{t+1} - p_t)^2 \vert p_t]$	1.0185	1.0185
$\mathbb{E}[(p_{t+1} - p_t)^4 \vert p_t]$	3.2537	3.1120
$\mathbb{E}[(p_{t+1} - p_t)^6 \vert p_t]$	18.0829	15.8474
$\mathbb{E}[(p_{t+1} - p_t)^8 \vert p_t]$	146.5919	112.9830
$\mathbb{E}[(p_{t+1} - p_t)^{10} \vert p_t]$	1588.8096	1035.6463

Table 1: Moment Comparison

Several points are worth mentioning about Table 1. First, we verify that n is large enough in the sense of Theorem 2, so a nontrivial equilibrium exists. Second, we skip odd moments because they equal zero according to Theorem 2. Lastly, the comparison is to a $N(0, 1.0185)$ distribution, so the second moment matches by construction. For higher moments, the equilibrium distribution has larger moment values than those of the Normal, consistent with the statement of Theorem 3. One way to summarize Table 1 is to say that the effect of mixing on price tails is modest, at least with this set of parameter values. Another way would be to say that the equilibrium distribution is almost Normal (with no a priori reason to expect a Normal equilibrium other than Theorem 1).

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.

5.2 Rare Disasters

Recall fact five from Giglio et al. (2021): expected returns and the subjective probability of rare disasters are negatively related. In order to investigate this point, we first define a cutoff $\mathcal{R} < 0$. To be consistent with Barro (2006) and the rare disasters literature, we say that our model is characterized by rare disasters if the probability of prices falling below $\mathcal{R}$ *exceeds* that of the Normal distribution. We use the following corollary to connect our results about excess kurtosis from Section 5.1 to tail probabilities of the price distribution. As stated in Barro (2006), "fat tails have effects analogous to rare disasters."

Corollary 1 (Rare disasters)**.** *Let $f(x)$ and $g(x)$ be symmetric around zero. Then,*

$$\mathbb{E}_f[X^4] > \mathbb{E}_g[X^4] \implies \exists \mathcal{R} < 0 \text{ s.t. } \Pr_f(X \leq \mathcal{R}) > \Pr_g(X \leq \mathcal{R})$$

First consider the Normal distribution with density $g(x)$. Our goal is to compare our equilibrium price density $f(x)$ to an *appropriate* Normal distribution, which we do by comparing to a Normal$(0, \Sigma^2)$ where Σ^2 denotes the equilibrium price variance. Next let us consider the premise of the implication in Corollary 1. Because $f(x)$ and $g(x)$ have the same mean and variance, this inequality $\mathbb{E}_f[X^4] > \mathbb{E}_g[X^4]$ precisely captures the excess kurtosis result of Theorem 3. That is, Theorem 3 implies $\mathbb{E}_f[X^4] > \mathbb{E}_g[X^4]$. Finally, the implication of Corollary 1 is that our model is characterized by rare disasters with cutoff $\mathcal{R}$. While Giglio et al. (2021) use two values of $\mathcal{R}$ in their survey, a percentage decline of -10% and -30%, here we are only able to guarantee the existence of one such negative number.

Corollary 1 guarantees that our model is characterized by rare disasters in an objective sense. We now move to fact five from Giglio et al. (2021), which requires us to discuss the subjective price distribution. Because traders form correct beliefs over all moments except the first—which is learned subject to experience effects—subjective price distributions are just shifted versions of the actual price distribution. Importantly, we assume that all traders agree on the definition of a rare disaster: the event that returns are less than $\mathcal{R}$. Then trader i's subjective probability of a rare disaster is $\Pr_f(X + \mu_t^i \leq \mathcal{R}) = F(\mathcal{R} - \mu_t^i)$, where f is the equilibrium price density and F is the equilibrium price distribution. Since distribution functions are non-decreasing, the subjective probability of a rare disaster $F(\mathcal{R} - \mu_t^i)$ is non-increasing in μ_t^i, as required by Giglio et al. (2021).

Figure 3 provides a visual summary of this section. As a point of reference, the Normal density is plotted in the dashed line. Overlaid on top of it, we plot the equilibrium price density using a solid line (centered at $\mu = 0$). Corollary 1 guarantees the existence of a point $\mathcal{R} < 0$ such that the probability of a realization below that point is *greater* for the equilibrium price distribution than that of the analogous Normal distribution. All traders agree on the definition of this rare disaster point, $\mathcal{R}$. However, they disagree on the center of the equilibrium price density; in particular, a trader with belief μ_t^i has beliefs over prices given by the solid shifted density (centered at μ_t^i). As required by Giglio et al. (2021), such a trader has a *higher* subjective probability of the rare disaster, or the event that prices fall below $\mathcal{R}$. This subjective probability of disaster, $F(\mathcal{R} - \mu_t^i)$, is given by the area under the shifted density (centered at μ_t^i) to the left of $\mathcal{R}$.

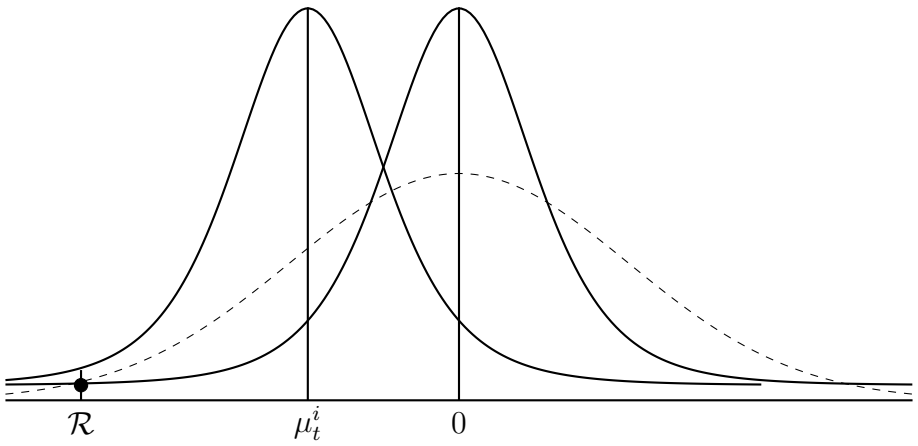

Figure 3: $F(\mathcal{R} - \mu_t^i)$

6 Conclusion

We have proposed a model consistent with, and in a large part motivated by, six facts about beliefs and portfolios from Giglio et al. (2021). First, beliefs in our model were reflected in portfolio allocations, but the elasticity of portfolios to beliefs was dampened by transaction costs. Second, beliefs affected the direction—but not timing—of trades by assumption. Third, beliefs were persistent, heterogeneous, and orthogonal to demographics. The persistence followed from two characteristics of the model: the tendency for traders to trade (and learn) in very few periods, and also the tendency for older traders to maintain past beliefs when updating. The orthogonality followed from the fact that older traders were no more likely to be an optimist versus pessimist, compared to young traders. Fourth, while our baseline setting was characterized by zero GDP growth, once we relaxed this assumption, Appendix A showed that GDP growth was positively related to returns. Fifth, the subjective probability of rare disasters decreased in subjective expected returns. Sixth, older individuals had tighter subjective distributions over future stock returns than their younger counterparts.

The starting point for our model was explicitly proposed by Giglio et al. (2021): a model where retail traders trade infrequently with some probability. We then combined this setup with one of experience effects like those documented in Andersen et al. (2019). Altogether, traders learned the equilibrium price drift, but only on those random days on which they had previously traded. When some traders recalled bear markets and others recalled bull markets, this created belief dispersion in the cross-section and, hence, a motive for trade. The first key takeaway from the paper was that past volatility may be precisely what guarantees enough belief dispersion to generate current volatility (i.e. a bad equilibrium). In the simplified version of the model where there was a representative-age trader, the Normal distribution was the unique equilibrium price distribution.

However, when we considered the full model, an interesting second dimension of heterogeneity—beyond pessimist and optimist—arose. The cross-section of traders consisted of the young (those who recalled few trading days and had dispersed beliefs) and the old (those who recalled many trading days and had concentrated beliefs). The mixture distribution over such demographics is not Normal, which led to equilibrium prices that were not Normal. The second key takeaway from the paper—that prices generated in such a setting *must* be characterized by excess kurtosis—resulted from the small chance that all traders trading on a given day were young. This result played an important role in matching the fifth fact from Giglio et al. (2021), that pessimists have higher subjective probabilities of rare disasters.

A Asset Supply and Dividends

In this appendix, we generalize the problem from Section 2 with positive asset supply $S > 0$ and positive dividends $\delta > 0$. Both values are exogenous and deterministic. The exact setting from Section 2 is recovered when $S = \delta = 0$. Mean-variance utility becomes

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

where, like before, τ denotes the transaction cost and ρ denotes the risk aversion parameter. Demand that maximizes such utility is given by

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

where Σ^2 denotes the equilibrium price variance. Nothing changes from the baseline setting with respect to the trader's recollection, and hence the trader's belief μ_t^i is

$$\mu_t^i = \frac{\gamma}{|\mathbb{K}_t^i|}\sum_{s\in\mathbb{K}_t^i}(p_s - p_{s-1}) + (1-\gamma)\mu$$

where $\mathbb{K}_t^i$ denotes the recalled periods for trader i at time t and γ denotes the memory parameter. The steady-state cross-section over μ^i has mean μ and variance $\gamma^2\Sigma^2\sum_{k=1}^{K-1}\mathcal{P}_k/k$, where $\mathcal{P}_k$ denotes the proportion of the population that recalls k periods.

With respect to the market maker, the linear rule becomes

$$p_{t+1} = p_t + c\left(\sum_{i=1}^N x^i(p_t) - S\right) \tag{14}$$

where the term in parentheses is excess demand. Zero profits imposes

$$\mathbb{E}\left[\frac{\tau}{2}\sum_{i=1}^N x^i(p_t)^2 - (\delta + p_{t+1} - p_t)\left(\sum_{i=1}^N x^i(p_t) - S\right)\right] = 0$$

because the market maker takes the position $-\left(\sum_{i=1}^N x^i(p_t) - S\right)$. To show that all of our results in the main setting apply to this extension, we first solve for the equilibrium drift μ by taking expectations of both left- and right-hand sides of (14),

$$\mu = cn\frac{\delta + \mu}{\rho\Sigma^2 + \tau} - cS$$

Subtracting this expression from both left- and right-hand sides of (14),

$$\begin{aligned} p_{t+1} - p_t - \mu =& c\sum_{i\in\mathcal{A}_t}\frac{\delta + \frac{\gamma}{|\mathbb{K}_t^i|}\sum_{s\in\mathbb{K}_t^i}(p_s - p_{s-1}) + (1-\gamma)\mu}{\rho\Sigma^2 + \tau} - cn\frac{\delta + \mu}{\rho\Sigma^2 + \tau} \\ =& c\sum_{i\in\mathcal{A}_t}\frac{\frac{\gamma}{|\mathbb{K}_t^i|}\sum_{s\in\mathbb{K}_t^i}(p_s - p_{s-1} - \mu)}{\rho\Sigma^2 + \tau} \end{aligned}$$

where the term cS cancels out in the first equality and the second equality follows from algebra. Now

redefine the random variable $q_{t+1} - q_t \equiv p_{t+1} - p_t - \mu$, so that $\mathbb{E}[q_{t+1} - q_t] = 0$ and almost all of our theoretical results from the main paper apply. The only result that does not hold is Lemma 2, because $\mu \neq 0$ and c may equal a complicated expression to ensure zero profits.

This extension allows us to address fact four from Giglio et al. (2021): expected GDP growth and expected returns are positively related. GDP in our model is calculated using income, which we denote w^i_{t+1}. Formally, mean-variance utility in (1) is formed over w^i_{t+1}, where it is defined

$$w^i_{t+1} = (\delta + p_{t+1} - p_t)x^i_t - \frac{\tau}{2}(x^i_t)^2 + w^i_t$$

Furthermore, we must take into account the balance sheet of the market maker:

$$\mathcal{W}_{t+1} = \frac{\tau}{2}\sum_{i=1}^{N}(x^i_t)^2 - (\delta + p_{t+1} - p_t)\left(\sum_{i=1}^{N} x^i_t - S\right) + \mathcal{W}_t$$

where $\mathcal{W}_{t+1}$ denotes market maker wealth. Then GDP growth equals $\sum_{i=1}^{N}(w^i_{t+1} - w^i_t) + (\mathcal{W}_{t+1} - \mathcal{W}_t)$. Because the fact from Giglio et al. (2021) cites perceptions of each variable,

$$\begin{aligned}
&\text{Cov}\left(\mathbb{E}^i\left[\sum_{i=1}^{N}(w^i_{t+1} - w^i_t) + (\mathcal{W}_{t+1} - \mathcal{W}_t)\middle| p_t\right], \mathbb{E}^i[p_{t+1} - p_t|p_t]\right) \\
=&\text{Cov}\left(\mathbb{E}^i\left[S(\delta + p_{t+1} - p_t)|p_t\right], \mathbb{E}^i[p_{t+1} - p_t|p_t]\right) \\
=&S\text{Var}\left(\mathbb{E}^i[p_{t+1} - p_t|p_t]\right) > 0,
\end{aligned}$$

where the covariance is cross-sectional. This discussion emphasizes the requirement that $S > 0$, else our economy is characterized by zero GDP growth.

B Limit Order Model

Here we undertake a similar exercise to that of Section 2, but the market microstructure is changed. Instead of traders submitting market orders and the market maker making zero profits, say that traders submit limit orders (their entire demand schedules) and, consequently, markets clear. Traders recall past market clearing prices, p_{t-s} for $s \geq 1$, and use this knowledge to form expectations over future market clearing prices, $\mathbb{E}^i[p_{t+1}]$. Today's market clearing price, p_t, is taken as given.[33] Demand functions become

$$x^i(p_t) = \frac{\frac{\gamma}{|\mathbb{K}^i_t|}\sum_{s\in\mathbb{K}^i_t} p_s + (1-\gamma)\mu - p_t}{\rho\Sigma^2 + \tau}$$

where γ denotes the memory parameter, $\mathbb{K}^i_t$ denotes the set of recalled periods for trader i, μ denotes the average price, ρ denotes the risk aversion parameter, Σ^2 denotes the equilibrium price variance, and τ denotes the transaction cost. Market clearing requires that

$$\frac{\sum_{i\in\mathcal{A}_t}\frac{\gamma}{|\mathbb{K}^i_t|}\sum_{s\in\mathbb{K}^i_t} p_s + n(1-\gamma)\mu - np_t}{\rho\Sigma^2 + \tau} = S \tag{15}$$

[33] This market clearing setup emphasizes that, while traders are short-lived, the asset is long-lived. Prices p_t clear markets today, then short-lived traders unwind their positions at market clearing prices tomorrow, p_{t+1}, which are determined by demand tomorrow.

where S denotes asset supply, which is included here to emphasize the fact that its presence does not affect our argument. The equilibrium requirement—that the past price distribution matches the current price distribution—now applies to (15). In particular, taking variances of both LHS and RHS,

$$\gamma^2 n\Sigma^2 \sum_{k=1}^{K-1} \mathcal{P}_k/k - n^2\Sigma^2 = 0$$

so the only equilibrium is trivial, $\Sigma^2 = 0$.[34] In the equality above, the first Σ^2 on the LHS denotes past price volatility and the second Σ^2 on the LHS denotes current price volatility. Together, this formula says that current prices must have a lower variance than past prices, with a trivial fixed point. Consider the *elasticity* of current price variance to past variance, which is $\gamma^2 \sum_{k=1}^{K-1} \mathcal{P}_k/(nk) < 1$ here. In our baseline setting (market orders from Section 2), that elasticity is

$$\mathcal{E} \equiv \frac{c^2 n\gamma^2 \sum_{k=1}^{K-1} \mathcal{P}_k/k}{(\rho\Sigma^2 + \tau)^2}$$

Even controlling for the fact that c is endogenous due to the zero profit condition (that is, fixing c), this elasticity is endogenous in our model. In our equilibrium, we find a Σ^2 such that this elasticity $\mathcal{E}$ equals the value one.

C Proofs

C.1 Proof of Lemma 1

The posterior according to memory-constrained Bayes' Rule (6) and weights (7) is given by

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

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

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

and divide by the log of the product of $(1-\gamma)T$ densities evaluated at independent prices.[35] We claim that, as $T \to \infty$, this ratio satisfies

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

[34] This is because $\frac{\gamma^2 \sum_{k=1}^{K-1} \mathcal{P}_k/k}{n} < 1$.

[35] 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 recalled prices, $\frac{\gamma T}{|\mathbb{K}_t^i|}$.

where the equality follows from algebra and the convergence follows from the strong law of large numbers, assuming $|\mathbb{K}_t^i| < \infty$. This line of reasoning justifies replacing non-recalled terms in the posterior with the product of $(1-\gamma)T$ densities evaluated at independent prices. Next consider the log of each recalled term,

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

which (trivially) equals the log of the product of $\frac{\gamma T}{|\mathbb{K}_t^i|}$ densities evaluated at the same recalled price, $(p_s - p_{s-1})$ for $s \in \mathbb{K}_t^i$:

$$\ln\left([f(p_s - p_{s-1}|\mu)]^{\gamma T/|\mathbb{K}_t^i|}\right) = \frac{\gamma T}{|\mathbb{K}_t^i|} \ln[f(p_s - p_{s-1}|\mu)]$$

Putting arguments altogether, the frequencies $(1-\gamma)$ and $\gamma/|\mathbb{K}_t^i|$ observed in the posterior precisely match those of the proposed mixture H_t. Mixture distributions have an expected value that is equal to the mixture of the expected values:

$$\mathbb{E}_{H_t}[\mu] = \frac{\gamma}{|\mathbb{K}_t^i|} \sum_{s \in \mathbb{K}_t^i} (p_s - p_{s-1}) + (1-\gamma)\mu$$

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

C.2 Proof of Lemma 2

We begin with the equilibrium condition for μ, (8), and the equilibrium condition for Σ^2, which is found by taking the variance of both LHS and RHS of market maker (5),

$$\Sigma^2 = c^2 n \sum_{k=1}^{K-1} \frac{\mathcal{P}_k}{k} \frac{\gamma^2 \Sigma^2}{(\rho\Sigma^2 + \tau)^2} \tag{16}$$

In the trivial equilibrium from (16) where $\Sigma^2 = 0$, say that $\mu \neq 0$. Then (8) requires that $\tau = cn$. Zero profits (4) becomes

$$\begin{aligned}\mathbb{E}\left[\frac{\tau}{2}\sum_{i=1}^{N} x^i(p_t)^2 - \frac{\tau}{n}\left(\sum_{i=1}^{N} x^i(p_t)\right)^2\right] =& \frac{\tau}{2} n \frac{\mu^2}{\tau^2} - \frac{\tau}{n} n(n-1)\left(\frac{\mu}{\tau}\right)^2 - \frac{\tau}{n} n \frac{\mu^2}{\tau^2} \\ =& -\frac{n}{2}\frac{\mu^2}{\tau} \\ \neq& 0\end{aligned}$$

where the first equality follows from the fact that[36]

$$\begin{aligned}
\mathbb{E}[x^i(p_t)^2] =& \mathbb{E}\left[\left(\frac{\frac{\gamma}{|\mathbb{K}_t^i|}\sum_{s\in\mathbb{K}_t^i}(p_s - p_{s-1}) + (1-\gamma)\mu}{\tau}\right)^2\right] \\
=& \mathbb{E}\left[\left(\frac{\frac{\gamma}{|\mathbb{K}_t^i|}\sum_{s\in\mathbb{K}_t^i}(p_s - p_{s-1} - \mu) + \mu}{\tau}\right)^2\right] \\
=& \mathbb{E}\left[\left(\frac{\frac{\gamma}{|\mathbb{K}_t^i|}\sum_{s\in\mathbb{K}_t^i}(p_s - p_{s-1} - \mu)}{\tau}\right)^2 + 2\frac{\frac{\gamma\mu}{|\mathbb{K}_t^i|}\sum_{s\in\mathbb{K}_t^i}(p_s - p_{s-1} - \mu)}{\tau^2} + \frac{\mu^2}{\tau^2}\right] \\
=& \frac{\mu^2}{\tau^2}
\end{aligned}$$

where the last equality directly above follows from $\Sigma^2 = 0$ and $\mathbb{E}[p_s - p_{s-1} - \mu] = 0$. Altogether, this contradicts zero profits. Now consider the nontrivial equilibrium from (16) satisfying $\rho\Sigma^2 + \tau = c\gamma(n\sum_{k=1}^{K-1}\mathcal{P}_k/k)^{1/2}$. Plugging into (8),

$$\mu = \frac{\mu\sqrt{n}}{\gamma\sqrt{\sum_{k=1}^{K-1}\mathcal{P}_k/k}}$$

But note that

$$\frac{\sqrt{n}}{\gamma\sqrt{\sum_{k=1}^{K-1}\mathcal{P}_k/k}} > 1$$

which implies that $\mu = 0$. Finally, consider the zero profit condition (4) with (5) plugged in:

$$0 = \mathbb{E}\left[\frac{\tau}{2}\sum_{i=1}^{N}x^i(p_t)^2 - c\left(\sum_{i=1}^{N}x^i(p_t)\right)^2\right] = \left(\frac{\tau}{2} - c\right)\mathbb{E}\left[\sum_{i=1}^{N}x^i(p_t)^2\right]$$

where the second equality follows from independently drawn demands and $\mu = 0$. If $\mathbb{E}\left[x^i(p_t)^2\right]$ is not zero, then $c = \tau/2$. □

C.3 Proof of Lemma 3

By Chebyshev's Inequality, for each m

$$\Pr\left(\left|\frac{1}{N}\sum_{i=1}^{N}1_{\mu_t^i\leq m} - G(m)\right| \geq \varepsilon\right) \leq \frac{\mathrm{Var}(\sum_{i=1}^{N}1_{\mu_t^i\leq m})}{N^2\varepsilon^2} \tag{17}$$

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

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

[36] By a similar calculation, $\mathbb{E}[x^i(p_t)] = \mu/\tau$.

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

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

With these two bounds, we continue the line of reasoning in (17):

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

as $N \to \infty$. The inequality above first uses the previously-derived variance and covariance bounds. Second, it uses an upper bound for the number of covariance terms, which is attained when all traders recall $(K-1)$ periods. That is, when traders recall $(K-1)$ periods, their beliefs overlap with up to $(n-1)(K-1)$ other traders. □

C.4 Proof of Theorem 1

To keep notation clear, we index traders by j in this proof. We begin with the market maker rule (5), multiply by $i \times s$ where $i = \sqrt{-1}$ and s is a real number, take exponentials of both LHS and RHS, then take expectations:

$$\mathbb{E}[\exp(isX)] = \mathbb{E}\left[\exp\left(is\frac{c\gamma}{\rho\Sigma^2 + \tau}\sum_{j=1}^{n}\frac{1}{\bar{k}}\sum_{t=1}^{\bar{k}} X_{jt}\right)\right]$$

where $X = p_{t+1} - p_t$, $X_{jt} = p_{jt} - p_{jt-1}$, and $(p_{jt} - p_{jt-1})$ denotes the t-th recalled price of the $j-$th independently drawn trader. Due to independence (across traders and past prices) and the equilibrium requirement that current and recalled prices share the same distribution,

$$\mathbb{E}[\exp(isX)] = \mathbb{E}\left[\exp\left(is\frac{c\gamma/\bar{k}}{\rho\Sigma^2 + \tau}X\right)\right]^{n\bar{k}}$$

Plugging in the nontrivial equilibrium variance that satisfies $\Sigma^2\rho + \tau = c\gamma(n/\bar{k})^{1/2}$,

$$\mathbb{E}[\exp(isX)] = \mathbb{E}\left[\exp\left(\frac{is}{\sqrt{n\bar{k}}}X\right)\right]^{n\bar{k}} \iff \varphi(s) = \varphi\left(\frac{s}{\sqrt{n\bar{k}}}\right)^{n\bar{k}}$$

where $\varphi(s)$ denotes the characteristic function. Letting $\bar{n} = n\bar{k}$, this last equality can be rewritten

$$\varphi(s\sqrt{\bar{n}})^{1/\bar{n}} = \varphi(s) \tag{18}$$

Iterating forward,

$$\varphi(s) = \varphi\left(s\sqrt{\bar{n}}^{\ell}\right)^{1/\bar{n}^{\ell}}$$

Define the sequence $s_\ell = 1/\sqrt{\bar{n}}^{\ell} \to 0$ as $k \to \infty$ (because $\bar{n} \equiv n\bar{k} > 1$), so that

$$\varphi(s_\ell) = \varphi(1)^{(s_\ell)^2} \tag{19}$$

Because prices have moments up to (at least) the second order, the characteristic function is (at least) twice continuously differentiable on the real line. Hence (19) implies that

$$\varphi(s) = \varphi(1)^{s^2}$$

when s non-negative and small enough. To solve for the constant $\varphi(1)$, we take two derivatives evaluated at zero:

$$\Sigma^2 = -\frac{\partial^2 \varphi}{\partial^2 s}(0) = -2\ln(\varphi(1))$$

or $\varphi(1) = \exp(-\Sigma^2/2)$. Note that the first equality above follows from properties of characteristic functions and the second equality follows from direct computation. Similarly, we can evaluate along the sequence $s'_\ell = -1/\sqrt{n}^\ell$ so that

$$\varphi(s'_\ell) = \varphi(-1)^{(s'_\ell)^2}, \text{ or } \varphi(s) = \varphi(-1)^{s^2}$$

when s is non-positive and large enough. Taking two derivatives evaluated at zero again yields $\varphi(-1) = \exp(-\Sigma^2/2)$. To conclude, $\varphi(s) = \exp(-s^2\Sigma^2/2)$, which is the characteristic function of the Normal$(0, \Sigma^2)$ distribution, in a neighborhood around zero. Then (18) uniquely extends this function to all $s \in \mathbb{R}$. □

C.5 Proof of Theorem 2

First, we set up the continuous operator $\mathcal{T}$ induced by the market maker's rule (5), mapping historical price densities into the current one:

$$\mathcal{T}(f) = \left(\sum_{k=1}^{K-1} \mathcal{P}_k [\hat{\Sigma}^2 k f(x\hat{\Sigma}^2 k)] *_k \right) *_n \tag{20}$$

where $\hat{\Sigma}^2 \equiv (n \sum_{k=1}^{K-1} \mathcal{P}_k/k)^{1/2}$ and $*_k$ denotes the k-fold convolution. Starting from the outside of the formula (20) and moving inward, prices from the market maker's rule (5) are the n-fold convolution of the mixture of

$$\frac{\frac{c\gamma}{k} \sum_{s \in \mathbb{K}_t^i} (p_s - p_{s-1})}{\rho\Sigma^2 + \tau} = \frac{\sum_{s \in \mathbb{K}_t^i} (p_s - p_{s-1})}{\hat{\Sigma}^2 k}$$

assuming that the trader recalls k periods. Note that we have imposed zero drift here from Lemma 2. The equality above follows from plugging in the nontrivial equilibrium $\rho\Sigma^2 + \tau = c\gamma(n \sum_{k=1}^{K-1} \mathcal{P}_k/k)^{1/2}$. In general, note that $\mathcal{C}X$ has distribution $F(x/\mathcal{C})$ and density $f(x/\mathcal{C})/\mathcal{C}$ when $\mathcal{C}$ is a generic constant and X is a random variable with distribution $F(x)$. This is reflected inside of the brackets in (20), with $\mathcal{C} = 1/(\hat{\Sigma}^2 k)$. Finally, the inner k-fold convolution in (20) reflects a trader averaging over k recalled periods. The goal is to apply the Schauder fixed point theorem to $\mathcal{T}$.

In order for us to do so, we restrict the domain of $\mathcal{T}$ to symmetric densities (in particular, with mean zero) with variance Σ^2. Call this space of densities $\mathcal{G}$. By construction of these moments, $\mathcal{T}(f)$ has mean zero and variance Σ^2 when f has mean zero and variance Σ^2. To show that $\mathcal{T} : \mathcal{G} \to \mathcal{G}$, we must show that $\mathcal{T}(f)$ is symmetric. We state without proof that the mixture operation in (20) preserves symmetry.

Convolution in (20) preserves symmetry because

$$\begin{aligned}(h*g)(x) &= \int_{-\infty}^{\infty} h(\tau)g(x-\tau)\mathrm{d}\tau \\ &= \int_{-\infty}^{\infty} h(-\tau)g(-x+\tau)\mathrm{d}\tau \\ &= \int_{-\infty}^{\infty} h(-y-x)g(y)\mathrm{d}y \\ &= (h*g)(-x)\end{aligned}$$

where the second equality follows symmetry of h and g and the third equality follows from the change of variable $y = \tau - x$.[37] The Schauder fixed point theorem requires that $\mathcal{G}$ be nonempty, convex, and closed. These properties can be explicitly verified. The last two required properties of the theorem are that $\mathcal{T}$ be continuous and $\mathcal{T}(\mathcal{G})$ be contained in a compact subset of $\mathcal{G}$.

We first tackle continuity, taking as given that multiplication by constants and finite mixtures in (20) are continuous operations. To show that k-fold convolution, which we denote T_k, is continuous, begin with the base case $k = 2$. Let $\lim_{j\to\infty} \int_{-\infty}^{\infty} |f_j(x) - f(x)|\mathrm{d}x = 0$ for some sequence of densities $(f_j)_{j=1}^{\infty}$. Then

$$\begin{aligned}&\lim_{j\to\infty} \int_{-\infty}^{\infty} |T_2 \circ f_j(x) - T_2 \circ f(x)|\mathrm{d}x \\ =& \lim_{j\to\infty} \int_{-\infty}^{\infty} |f_j(x) * f_j(x) - f(x) * f(x)|\mathrm{d}x \\ =& \lim_{j\to\infty} \int_{-\infty}^{\infty} |f_j(x) * f_j(x) - f_j(x) * f(x) + f_j(x) * f(x) - f(x) * f(x)|\mathrm{d}x \\ \leq& \lim_{j\to\infty} \int_{-\infty}^{\infty} |f_j(x) * f_j(x) - f_j(x) * f(x)|\mathrm{d}x + \lim_{j\to\infty} \int_{-\infty}^{\infty} |f_j(x) * f(x) - f(x) * f(x)|\mathrm{d}x \\ =& \lim_{j\to\infty} \int_{-\infty}^{\infty} |f_j(x) * (f_j(x) - f(x))|\mathrm{d}x + \lim_{j\to\infty} \int_{-\infty}^{\infty} |f(x) * (f_j(x) - f(x))|\mathrm{d}x \\ \leq& \lim_{j\to\infty} \int_{-\infty}^{\infty} f_j(x)\mathrm{d}x \int_{-\infty}^{\infty} |f_j(x) - f(x)|\mathrm{d}x + \lim_{j\to\infty} \int_{-\infty}^{\infty} f(x)\mathrm{d}x \int_{-\infty}^{\infty} |f_j(x) - f(x)|\mathrm{d}x \\ =& \lim_{j\to\infty} 2 \int_{-\infty}^{\infty} |f_j(x) - f(x)|\mathrm{d}x \\ =& 0\end{aligned}$$

where the first inequality is due to the triangle inequality and the second inequality is due to Young's convolution inequality. In the inductive step, assume that T_k is continuous. Then, using the same line of

[37] Note that we write the proof for any symmetric h and g, as opposed to $(f * f)$, because it applies now to $h \equiv f*_k$ and $g \equiv f*_\ell$.

reasoning as above,

$$
\begin{aligned}
&\lim_{j\to\infty}\int_{-\infty}^{\infty}|T_{k+1}\circ f_j(x)-T_{k+1}\circ f(x)|\mathrm{d}x\\
=&\lim_{j\to\infty}\int_{-\infty}^{\infty}|T_k\circ f_j(x)*f_j(x)-T_k\circ f(x)*f(x)|\mathrm{d}x\\
=&\lim_{j\to\infty}\int_{-\infty}^{\infty}|T_k\circ f_j(x)*f_j(x)-T_k\circ f_j(x)*f(x)+T_k\circ f_j(x)*f(x)-T_k\circ f(x)*f(x)|\mathrm{d}x\\
\leq&\lim_{j\to\infty}\int_{-\infty}^{\infty}|T_k\circ f_j(x)*f_j(x)-T_k\circ f_j(x)*f(x)|\mathrm{d}x\\
&+\lim_{j\to\infty}\int_{-\infty}^{\infty}|T_k\circ f_j(x)*f(x)-T_k\circ f(x)*f(x)|\mathrm{d}x\\
=&\lim_{j\to\infty}\int_{-\infty}^{\infty}|T_k\circ f_j(x)*(f_j(x)-f(x))|\mathrm{d}x+\lim_{j\to\infty}\int_{-\infty}^{\infty}|f(x)*(T_k\circ f_j(x)-T_k\circ f(x))|\mathrm{d}x\\
\leq&\lim_{j\to\infty}\int_{-\infty}^{\infty}T_k\circ f_j(x)\mathrm{d}x\int_{-\infty}^{\infty}|f_j(x)-f(x)|\mathrm{d}x+\lim_{j\to\infty}\int_{-\infty}^{\infty}f(x)\mathrm{d}x\int_{-\infty}^{\infty}|T_k\circ f_j(x)-T_k\circ f(x)|\mathrm{d}x\\
=&\lim_{j\to\infty}\int_{-\infty}^{\infty}|f_j(x)-f(x)|\mathrm{d}x+\lim_{j\to\infty}\int_{-\infty}^{\infty}|T_k\circ f_j(x)-T_k\circ f(x)|\mathrm{d}x\\
=&0
\end{aligned}
$$

where the last equality follows from continuity of T_k. Finally, we tackle compactness. Our strategy is to show that $\mathcal{G}$ is tight. Then, by Prokhorov's theorem, the closure of $\mathcal{G}$ (which is just $\mathcal{G}$) is compact when equipped with the topology of weak convergence. Take any sequence $(f_j)_{j=1}^{\infty}$ with $f_j\in\mathcal{G}$. By Markov's Inequality,

$$\Pr(|Z_j|\geq R)\leq\frac{\mathbb{E}[Z_j^2]}{R^2}=\frac{\Sigma^2}{R^2}$$

where Z_j denotes the random variable with density f_j and the equality follows from the fact that $\mathbb{E}[Z_j]=0$. This line of reasoning proves tightness. □

C.6 Proof of Theorem 3

If the trader drawn from the cross section recalls k periods, the fourth moment of demand would be

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

because the price history is independent and prices have mean zero. Since the cross section is distributed according to $(\mathcal{P}_k)_{k=1}^{K-1}$, each demand has fourth moment

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

The fourth price moment is given by the expression

$$\mu^4 = c^4(n\mathbb{E}[x^i(p_t)^4] + 3n(n-1)\mathbb{E}[x^i(p_t)^2]^2) \tag{21}$$

because demands are independently drawn from the cross-section and have mean zero. We divide (21) by $(\Sigma^2)^2$ and plug in two terms: first, the fourth moment of demand from above and, second, the demand variance,

$$\mathbb{E}[x^i(p_t)^2] = \sum_{k=1}^{K-1} \mathcal{P}_k \mathbb{E}\left[\left(\frac{\frac{\gamma}{k}\sum_{s\in\mathbb{K}_t^i}(p_s - p_{s-1})}{\rho\Sigma^2+\tau}\right)^2\right]$$

so that

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

where

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

Canceling terms and plugging in the nontrivial equilibrium value $(\rho\Sigma^2+\tau)^2 = c^2\gamma^2 n\sum_{k=1}^{K-1}\frac{\mathcal{P}_k}{k}$,

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

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

$$\frac{\mu^4}{(\Sigma^2)^2} = 3 + \frac{3\left[\sum_{k=1}^{K-1}\frac{\mathcal{P}_k}{k^2} - \left(\sum_{k=1}^{K-1}\frac{\mathcal{P}_k}{k}\right)^2\right]}{n\left(\sum_{k=1}^{K-1}\frac{\mathcal{P}_k}{k}\right)^2 - \sum_{k=1}^{K-1}\frac{\mathcal{P}_k}{k^3}} \tag{22}$$

Now consider the assumptions of the theorem:

$$n > 4\left(\gamma^2\sum_{k=1}^{K-1}\frac{\mathcal{P}_k}{k}\right)^{-1}$$

It turns out that this will guarantee that the denominator on the RHS of (22) is positive. To see why, rearrange our assumption:

$$n\sum_{k=1}^{K-1}\frac{\mathcal{P}_k}{k} > \frac{4}{\gamma^2} \geq 4$$

where the weak inequality follows from $\gamma \leq 1$. Multiply both sides by $\sum_{k=1}^{K-1}\mathcal{P}_k/k$:

$$n\left(\sum_{k=1}^{K-1}\frac{\mathcal{P}_k}{k}\right)^2 > 4\sum_{k=1}^{K-1}\frac{\mathcal{P}_k}{k} \geq 4\sum_{k=1}^{K-1}\frac{\mathcal{P}_k}{k^3} > \sum_{k=1}^{K-1}\frac{\mathcal{P}_k}{k^3}$$

where the weak inequality follows from the fact that both LHS and RHS of the inequality are expected values with the same probabilities and weakly smaller values (on the RHS). The last strict inequality is algebraic. Then for the desired result, it suffices to show that the numerator of the RHS of (22) is also positive,

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

But this inequality can be rewritten $\mathbb{E}[X^2] - \mathbb{E}[X]^2 > 0$, where probabilities are given by $\mathcal{P}_k$ and the random variable takes on values $1/k$. The variance is, indeed, strictly positive when the distribution is nondegenerate.

With respect to Condition 1 (that Σ^2 is increasing in $\mathbb{E}_{\mathcal{P}}[1/X]$), this follows from the closed-form solution for the equilibrium variance:

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

Condition 2 (that $\mu^4/(\Sigma^2)^2$ is increasing in $\mathrm{Var}_{\mathcal{P}}(1/X)$) follows from the closed form solution (22). □

C.7 Proof of Corollary 1

By the Corollary on page 47 of Serfling (1980), we have that

$$\mathbb{E}_f[X^4] = 4\int_0^\infty t^3 \mathrm{Pr}_f(|X| \geq t)\mathrm{d}t$$

Based on the premise of Corollary 1, we then have that

$$\int_0^\infty t^3 \mathrm{Pr}_f(|X| \geq t)\mathrm{d}t > \int_0^\infty t^3 \mathrm{Pr}_g(|X| \geq t)\mathrm{d}t \tag{23}$$

Next we claim that there exists $s > 0$ such that $\mathrm{Pr}_f(|X| \geq s) > \mathrm{Pr}_g(|X| \geq s)$. Say otherwise,

$$\mathrm{Pr}_f(|X| \geq s) \leq \mathrm{Pr}_g(|X| \geq s) \text{ for all } s > 0$$

Multiplying by $s^3 > 0$,

$$s^3\mathrm{Pr}_f(|X| \geq s) \leq s^3\mathrm{Pr}_g(|X| \geq s) \text{ for all } s > 0$$

Then, integrating,

$$\int_0^\infty s^3 \mathrm{Pr}_f(|X| \geq s)\mathrm{d}s \leq \int_0^\infty s^3 \mathrm{Pr}_g(|X| \geq s)\mathrm{d}s$$

which contradicts (23). Since f and g are both symmetric by assumption,

$$\mathrm{Pr}_f(|X| \geq s) = 2\mathrm{Pr}_f(-X \geq s) = 2\mathrm{Pr}_f(X \leq -s)$$

and the same logic applies for g. Then our result that $\mathrm{Pr}_f(|X| \geq s) > \mathrm{Pr}_g(|X| \geq s)$ for some $s > 0$ can be rewritten

$$\mathrm{Pr}_f(X \leq -s) > \mathrm{Pr}_g(X \leq -s)$$

which completes the proof. □

D Policy

Our theory allows us to examine traders' responses to policy. We consider two policies to reduce price variance in the general model (no representative-age trader): a transaction tax and a capital gains tax. So as to sidestep issues of trader welfare, we take a reduced form approach and take price variance reduction to be welfare improving. Note that all upcoming policy results describe 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.

D.1 Transaction Tax

In this section we assume that, on top of the transaction cost $\frac{\tau}{2}x_t^2$ paid to the market maker, traders pay an additional transaction tax $\frac{r}{2}x_t^2$, where $r > 0$. Note that linear costs (instead of quadratic)—either in τ or r—are technically problematic because they create an inaction region. That is, some traders would neither buy at the high price nor sell at the low, and this mass of traders trading zero destroys absolute continuity from Theorem 2. We assume that this tax is not paid to the market maker, but instead to a government or planner, so the zero profit condition from Lemma 2 remains unaffected.

Taking the variance of both left- and right-hand sides of the market maker (5),

$$\Sigma^2 = c^2 n \sum_{k=1}^{K-1} \frac{\mathcal{P}_k}{k} \frac{\gamma^2 \Sigma^2}{(\rho\Sigma^2 + \tau + r)^2}$$

Because the tax has no effect on the trivial equilibrium, we focus on the nontrivial one. Taking this nontrivial equilibrium variance and plugging into an equilibrium drift condition analogous to (8),

$$\mu = cn \frac{\mu}{\rho\Sigma^2 + \tau + r} \iff \mu = cn \frac{\mu}{c\gamma\sqrt{n\sum_{k=1}^{K-1} \mathcal{P}_k/k}}$$

which implies that $\mu = 0$, like in the baseline setting.[38] This ensures that Lemma 2 still applies, and that $c = \frac{\tau}{2}$. Then, as intuition would suggest, the tax r reduces the equilibrium price variance from that of the baseline setting, with the caveat that the parameter values which ensure a nontrivial equilibrium exists would need slight adjustment.

D.2 Capital Gains Tax

We model the capital gains tax with an adjusted budget set:

$$w_{t+1} = (1 - \mathcal{T})(p_{t+1} - p_t)x_t - \frac{\tau}{2}x_t^2 + w_t$$

where w_{t+1} denotes wealth and the tax is in the range $0 < \mathcal{T} < 1$. The equilibrium variance must solve the variance of both left- and right-hand sides of market maker (5):

$$\Sigma^2 = c^2 n \sum_{k=1}^{K-1} \mathcal{P}_k \frac{(1 - \mathcal{T})^2 \gamma^2 \Sigma^2 / k}{[\rho(1 - \mathcal{T})^2 \Sigma^2 + \tau]^2} \tag{24}$$

[38] This follows from the fact that $n\frac{1}{\gamma\sqrt{n\sum_{k=1}^{K-1} \mathcal{P}_k/k}} > 1$.

The tax enters into expression (24) in two places: first, dampened historical volatility in the numerator of the RHS of (24) and, second, reduced conjectured wealth volatility in the denominator of the RHS of (24). The first effect makes traders trade less, and the second effect makes traders trade more. It turns out that whether the first dampening effect or the second amplifying effect dominates depends on parameter values. That is, $\partial\Sigma^2/\partial\mathcal{T}$ cannot be signed and we conclude that the tax may be counterproductive.

To be clear, the problem is not with potential changes in c. The nontrivial equilibrium variance that solves (24), when plugged into a drift condition analogous to (8),

$$\mu = cn\frac{(1-\mathcal{T})\mu}{\rho(1-\mathcal{T})^2\Sigma^2+\tau} \iff \mu = cn\frac{(1-\mathcal{T})\mu}{c\gamma(1-\mathcal{T})\sqrt{n\sum_{k=1}^{K-1}\mathcal{P}_k/k}}$$

implies that $\mu = 0$ like in the baseline setting. This ensures that Lemma 2 still applies, and that $c = \frac{\tau}{2}$.

To conclude the discussion about policy, while the transaction tax is a relatively effective policy tool, the capital gains tax need not be. The transaction tax targets trade volume reduction (working through x_t only), without perverse wealth volatility effects present in the case of capital gains taxes (that, instead, work through $p_{t+1} - p_t$).

References

A. Abel, J. Eberly, and S. Panageas. Optimal inattention to the stock market. *American Economic Review*, 97(2):244–249, 2007.

A. Admati and P. Pfleiderer. A theory of intraday patterns: Volume and price variability. *Review of Financial Studies*, 1(1):3–40, 1988.

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.

R. Azeredo da Silveira, Y. Sung, and M. Woodford. Optimally imprecise memory and biased forecasts. *American Economic Review*, 114(10):3075–3118, 2024.

R. Barro. Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics*, 121(3):823–866, 2006.

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.

B. De Jong, A. Shleifer, L. Summers, and R. Waldmann. Positive feedback investment strategies and destabilizing rational speculation. *Journal of Finance*, 45(2):379–395, 1990a.

B. De Long, A. Shleifer, L. Summers, and R. Waldmann. Noise trader risk in financial markets. *Journal of Political Economy*, 98(4):703–738, 1990b.

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

D. Duffie and T.-s. Sun. Transaction costs and portfolio choice in a discrete-continuous-time setting. *Journal of Economic Dynamics and Control*, 14(1):35–51, 1990.

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.

Z. Jiang, H. Liu, C. Peng, and H. Yan. Investor memory and biased beliefs: Evidence from the field. *Working Ppaer*, 2023.

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

F. Palomino. Noise trading in small markets. *Journal of Finance*, 51(4):1537–1550, 1996.

J. Peress and D. Schmidt. Noise traders incarnate: Describing a realistic noise trading process. *Journal of Financial Markets*, 54:1–26, 2021.

A. Prat-Carrabin, F. Meyniel, M. Tsodyks, and R. Azeredo da Silveira. Biases and variability from costly bayesian inference. *Entropy*, 23(603):1–15, 2021.

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. Serfling. *Approximation Theorems of Mathematical Statistics*. John Wiley and Sons, 1980.

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. The optimal shape and horizon of the price impact function. *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.

A. Wang. Informed arbitrage with speculative noise trading. *Journal of Banking and Finance*, 34(2): 304–313, 2010.