

Investor experiences and financial market dynamics[☆]Ulrike Malmendier^{a,b,*}, Demian Pouzo^c, Victoria Vanasco^d^a UC Berkeley, NBER, and CEPR, USA^b Department of Economics and Haas School of Business, University of California, 501 Evans Hall, Berkeley, CA 94720-3880, USA^c Department of Economics, University of California, 501 Evans Hall, Berkeley, CA 94720-3880, USA^d CREI and BGSE, Ramon Trias Fargas 25–27, Barcelona 08005, Spain

ARTICLE INFO

Article history:

Received 4 June 2018

Revised 21 November 2018

Accepted 9 March 2019

Available online 15 November 2019

JEL classification:

G11

G12

G40

G41

E7

Keywords:

Experience effects

Learning

Asset prices

Portfolio choice

Demographics

ABSTRACT

How do macrofinancial shocks affect investor behavior and market dynamics? Recent evidence on experience effects suggests a long-lasting influence of personally experienced outcomes on investor beliefs and investment but also significant differences across older and younger generations. We formalize experience-based learning in an overlapping generations (OLG) model, where different cross-cohort experiences generate persistent heterogeneity in beliefs, portfolio choices, and trade. The model allows us to characterize a novel link between investor demographics and the dependence of prices on past dividends while also generating known features of asset prices, such as excess volatility and return predictability. The model produces new implications for the cross-section of asset holdings, trade volume, and investors' heterogeneous responses to crises, which we show to be in line with the data.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

Recent crises in the stock and housing markets have stimulated a new wave of macro-finance models of risk-taking. A key challenge, and motivation, has been to find tractable models of investor expectations that account not

only for asset pricing puzzles, such as return predictability (Campbell and Shiller, 1988; Fama and French, 1988) and excess volatility (LeRoy and Porter, 1981; Shiller, 1981; LeRoy, 2006), but also for micro-level stylized facts such as investors chasing past performances. As argued by Woodford (2013), the empirical evidence suggests a need for dynamic models that go beyond the rational-expectations hypothesis.

In line with Woodford's proposal, several emerging theories feature agents who over-weigh recent realizations of the relevant economic variables when forming beliefs.¹ Overweighing past realizations helps to successfully capture the above-mentioned asset pricing facts. However, as

[☆] We thank Marianne Andries, Nick Barberis, Dirk Bergemann, Julien Cujean, Xavier Gabaix, Lawrence Jin, and workshop participants at LBS, LSE, NYU, Pompeu Fabra, Stanford, UC Berkeley, as well as the ASSA, NBER EFG Behavioral Macro, NBER Behavioral Finance, SITE (psychology and economics segment), and SFB TR 15 (Tutzing, Germany) conferences for helpful comments. We also thank Felix Chopra, Marius Guenzel, Clint Hamilton, Canyao Liu, Leslie Shen, and Jonas Sobott for excellent research assistance.

* Corresponding author.

E-mail addresses: ulrike@berkeley.edu (U. Malmendier), dpouzo@berkeley.edu (D. Pouzo), vvanasco@crei.cat (V. Vanasco).

¹ Examples are models of natural expectation formation (Fuster et al., 2012; 2010) and over-extrapolation (Barberis et al., 2015; 2018).

the existing theories assume homogeneous overweighing of realizations from the last year or couple of years, they miss two important empirical patterns: first, macroeconomic shocks appear to alter investment and consumption behavior for decades to come, as conveyed by the notion of “depression babies” or of the “deep scars” left by the 2008 financial crisis (Blanchard, 2012; Malmendier and Shen, 2018). In other words, the measurable impact of past realizations goes far beyond the timeframe of existing models.

Second, there is significant and predictable cross-sectional heterogeneity. Younger cohorts tend to react significantly more strongly to a recent shock than older cohorts, both for positive and for negative realizations. For example, in the above-mentioned case of “scarred consumption” (Malmendier and Shen, 2018), periods of higher unemployment rates induces everybody to be more frugal, controlling for income, wealth, and unemployment status, but the effect is particularly strong among younger cohorts. The reverse holds after times of economic booms, when consumers continue to spend more lavishly, particularly the younger cohorts.

A growing empirical literature on experience effects shows both empirical patterns in numerous macrofinance contexts. This literature emphasizes that the long-lasting neuropsychology effects of personal “experience” is the key mechanism at work. For example, personal lifetime experiences in the stock market predict future willingness to invest in the stock market (Malmendier and Nagel, 2011), and the same holds for the bond market and bond investment, for Initial Public Offering (IPO) experiences and future IPO investment, and for inflation experiences and fixed-rate mortgage choices (Kaustia and Knüpfer, 2008; Chiang et al., 2011; Malmendier and Nagel, 2016). There is also evidence of experience effects in nonfinance settings, e. g., on the long-term effects of graduating in a recession on labor market outcomes (Oreopoulos et al., 2012), or of living in (communist) East Germany to political attitudes postreunification (Alesina and Fuchs-Schündeln, 2007),² which in turn affect beliefs about communism versus capitalism and explain east/west differences in stock market investment (Laudenbach et al., 2019a; 2019b).

The overweighing of personal experiences appears to be a pervasive and robust psychological phenomenon affecting belief formation, related to availability bias as first put forward by Tversky and Kahneman (1974) as well as the extensive evidence on the different effects of description versus experience.³ Much of the evidence on experience effects pertains directly to stated beliefs, e. g., beliefs about future stock returns (in the UBS/Gallup data), about future inflation (in the Michigan Survey of Consumers, known as the MSC), or about the outlook for durable consumption (also in the MSC).⁴

This growing empirical literature on experience effects and its strong psychological underpinning raise the question whether experience-based learning and the implied dynamic cross-cohort differences have the potential to explain aggregate dynamics. For example, which generations invest in the stock market and how much? What are the dynamics of stock market investment? How do markets react to a macro-shock?

Our paper develops an equilibrium model of asset markets that formalizes experience-based learning and the resulting belief heterogeneity across investors. The model clarifies the channels through which past realizations affect future market outcomes by pinning down the effect on investors’ own belief formation and the interaction with other generations’ belief formation. We derive the aggregate implications of learning from experience and the implied cross-sectional differences in investor behavior. To our knowledge, this model is the first to tease out the tension between experience effects and recency bias. It aims to provide a guide for testing to what extent experience-based learning can enhance our understanding of market dynamics and of the long-term effect of demographic changes.

The key model features are as follows. We consider a stylized overlapping generations (OLG) equilibrium model. Agents have constant absolute risk aversion (CARA) preferences and live for a finite number of periods. During their lifetimes, they choose portfolios of a risky and a risk-free security. We assume that agents maximize their per-period payoffs (i. e., are myopic).⁵ The risky asset is in unit net supply and pays a random dividend every period. The risk-free asset is in infinitely elastic supply and pays a fixed return. Investors do not know the true mean of dividends but learn about it by observing the history of dividends.

We begin by characterizing the benchmark economy in which agents know the true mean of dividends. In this setting, there is no heterogeneity, and thus the demands of all active market participants are equal and constant over time. Furthermore, there is a unique no-bubble equilibrium with constant prices.

We then introduce experience-based learning. The assumed belief formation process captures the two main empirical features of experience effects: first, agents overweigh their lifetime experiences. Second, their beliefs exhibit recency bias. We identify two channels through which past dividends affect market outcomes. The first channel is the belief formation process: shocks to dividends shape agents’ beliefs about future dividends. Hence, individual demands depend on personal experiences, and the equilibrium price is a function of the history of dividends observed by the oldest market participant.

The second channel is the generation of cross-sectional heterogeneity: different lifetime experiences generate per-

² See also Giuliano and Spilimbergo (2013), who relate the effects of growing up in a recession to redistribution preferences.

³ See, for example, Weber et al. (1993), Hertwig et al. (2004), and Simonsohn et al. (2008).

⁴ Cf. Malmendier and Nagel (2011), Malmendier and Nagel (2016), Malmendier and Shen (2018).

⁵ Myopic agents omit the correlation between their next period payoff and their continuation value function. This yields behavior that is analogous to the commonly used assumption of short-term traders (see Vives, 2010). In a previous version of the paper, available on <https://arxiv.org/pdf/1612.09553.pdf>, we show that, when the myopia assumption is removed, the first-order effects of experience-based learning are identical to those derived for myopic agents.

sistent differences in beliefs. Agents “agree to disagree.” Furthermore, younger cohorts react more strongly to a dividend shock than older cohorts, as it makes up a larger part of their lifetimes. A positive shock induces younger cohorts to invest relatively more in the risky asset, while a negative shock tilts the composition toward older cohorts. Thus, the model has implications for the time series of trade volume: changes in the level of disagreement between cohorts lead to higher trade volume in equilibrium.

The model captures an interesting tension between heterogeneity in personal experiences (which generates belief heterogeneity across cohorts) and recency bias (which reduces belief heterogeneity). When there is strong recency bias, all agents pay a lot of attention to the most recent dividends. Thus, their reactions to a recent shock are similar. Price volatility increases, while price autocorrelation and trade volume decrease. The opposite holds when the recency bias is weak, and agents form their beliefs using their experienced history. Hence, the reaction of prices and trade volume to changes in dividends is tightly linked to the relative extent of recency bias versus experience-based differences across cohorts in a given market, which are in turn influenced by demographics.

We explore the connection between market demographics and the dependence of prices on past dividends by analyzing the effect of a one-time change in the fraction of young agents that participate in the market. We find that the demographic composition of markets significantly influences the dependence of prices on past dividends. For example, when the market participation of the young relative to the old increases, the relative reliance of prices on more recent dividends increases. This is in line with evidence in [Cassella and Gulen \(2018\)](#) who find that the level of extrapolation in markets is positively related to the fraction of young traders in that market, and with [Collin-Dufresne et al. \(2017\)](#) who find that the price-dividend ratio is higher and more sensitive to macro-shocks when the ratio of young to old market participants is larger.

We then turn to several tests of the empirical implications of our model. First, we show that the model accommodates several key asset pricing features identified in prior literature. We follow the approach in [Campbell and Kyle \(1993\)](#) and [Barberis et al. \(2015\)](#) to contrast CARA-model moments with the data. We show the CARA-model analogues of return predictability ([Campbell and Shiller, 1988](#)) and of predictability of the dividend-price ratio. This predictability stems solely from the experience-based learning mechanism rather than, say, a built-in dependence on dividends or past returns, and it depends on the demographic structure of the market. Similarly, the model generates excess volatility in prices and price changes as established by [LeRoy and Porter \(1981\)](#), [Shiller \(1981\)](#), and [LeRoy \(2006\)](#), above and beyond the stochastic structure of the dividend process.

Experience-based learning generates new predictions for the cross-section of asset holdings and trade volume, which we test in the data. Using the representative sample of the Survey of Consumer Finance (SCF), merged with data from the Center for Research in Security Prices (CRSP) and historical data on stock market performance, we first replicate and extend the evidence in [Malmendier](#)

[and Nagel \(2011\)](#) on stock market participation. We show that cross-cohort differences in lifetime stock-market experiences predict cohort differences in stock market participation and in the fraction of liquid assets invested in the stock market. In other words, cross-cohort differences both on the extensive and on the intensive margin of stock market participation vary over time, as predicted by the time series of cross-cohort differences in lifetime experiences. We then turn to the predictions regarding trade volume and show that the detrended turnover ratio is strongly correlated with differences in lifetime market experiences across cohorts. That is, changes in the experience-based level of disagreement between cohorts predict higher abnormal trade volume, as predicted by the model.

Overall, experience-based learning offers a unifying explanation for financial market features of both prices and trade volume. It also has novel implications for the cross-sectional differences in market participation and portfolio choice, which we show are consistent with the data.

There is a wide literature on the role of learning in explaining asset pricing puzzles. Most closely related, [Cogley and Sargent \(2008\)](#) propose a model in which the representative consumer uses Bayes' theorem to update estimates of transition probabilities as realizations accrue. As in our paper, agents use less data than a “rational-expectations-without-learning econometrician” would give them. There are two important differences in our setup. First, agents are not Bayesian. Second, different cohorts have different, finite experiences. Consequently, observations during an agent's lifetime have a nonnegligible effect on beliefs and generate cross-cohort heterogeneity.

Our paper also relates to the work on extrapolation by [Barberis et al. \(2015\)](#) and [Barberis et al., 2018](#). They consider a consumption-based asset pricing model with both “rational” and “extrapolative” agents. The latter believe that positive price changes will be followed by positive changes. In contrast, the heterogeneity in extrapolation in our model is linked to the demographic structure of the market. In addition, while cross-sectional heterogeneity in their model arises from the presence of both rational and extrapolative infinitely lived agents, in our model, it results from the different experiences of different finitely lived cohorts. This allows us to generate predictions about the cross-section of asset holdings and the relation between extrapolation and demographics in line with the data.

More generally, our paper relates to the large asset pricing literature that departs from the correct beliefs paradigm. For instance, [Barsky and DeLong \(1993\)](#), [Timmermann \(1993, 1996\)](#), and [Adam et al. \(2016\)](#) study the implications of learning and [Cecchetti et al. \(2000\)](#) and [Jin, 2015](#) Working paper of distorted beliefs for stock return volatility and predictability, the equity premia, and booms and busts in markets. At the same time, our approach is different from asset pricing models with asymmetric information, as surveyed in [Brunnermeier \(2001\)](#). While in these models agents want to learn the information their counterparties hold, in our model of experience-based learning, information is available to all agents at all times.

Finally, there are contemporaneous papers that also explore the macroeconomic effects of learning from experi-

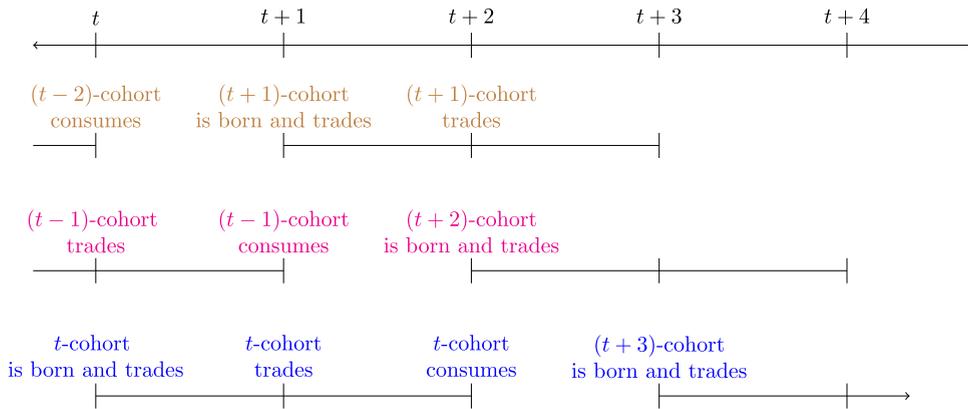


Fig. 1. A timeline for an economy with two-period lived generations, $q = 2$.

ence in OLG models. Schraeder (2015) focuses on how it impacts high-frequency trading patterns, such as overreaction and reversal, while Ehling et al. (2018) analyze the trend-chasing behavior of the young and its implications for risk premia and the risk-free rate. More closely related to our work, Collin-Dufresne et al. (2017) explore the role of demographics on asset pricing features, such as return predictability and excess volatility.

Our paper contributes to this literature in several respects. First, we allow for recency bias in the belief formation process, as both the underlying psychology literature on availability bias (Tversky and Kahneman, 1974) and the prior empirical literature on experience effects identify it as an important component of how individuals assign weights to previously experienced outcomes. Allowing for recency bias turns out to be also of interest theoretically, as the analysis reveals that an increase in recency bias reduces the cross-sectional heterogeneity driven by the experiential learning bias. Second, our agents are not Bayesian and do not update their posterior variance as they gain experience, and thus our results do not depend on heterogeneous posterior variances.⁶ Third, our CARA-normal framework allows us to obtain closed-form solutions to clearly understand the link between demographics, experience, and recency. Finally, we consider our empirical approach more comprehensive, as we test the model predictions about portfolio holdings, asset pricing features, and trade volume.

There is also a large literature that proposes other mechanisms, such as borrowing constraints or life-cycle considerations, as the link from demographics to asset prices and other equilibrium quantities. We view these other mechanisms as complementary to our paper. They are omitted for the sake of tractability of the model.

The remainder of the paper is organized as follows. First we present the model setup and the notion of experience-based learning in Section 2. We illustrate the mechanics of the model in a simplified setting in Section 3. The main

results are in Section 4. In Section 5 we extend the model to study demographic shocks, and in Section 6 we present empirical implications. Section 7 concludes. All proofs are in the Appendix.

2. Model set-up

Consider an infinite-horizon economy with overlapping generations of a continuum of risk-averse agents. At each point in time $t \in \mathbb{Z}$, a new generation is born and lives for q periods, $q \in \{1, 2, 3, \dots\}$. Hence, there are $q + 1$ generations alive at any t . The generation born at $t = n$ is called generation n . Each generation has a mass of q^{-1} identical agents.

Agents have CARA preferences with risk aversion γ . They can transfer resources across time by investing in financial markets. Trading takes place at the beginning of each period. At the end of the last period of their lives, agents consume the wealth they have accumulated. We use n_q to indicate the last time at which generation n trades; $n_q = n + q - 1$. (If the generation is denoted by t , we use t_q .) Fig. 1 illustrates the timeline of this economy for two-period lived generations ($q = 2$).

There is a risk-free asset, which is in perfectly elastic supply and has a gross return of $R > 1$ at all times. And there is a single risky asset (a Lucas tree), which is in unit net supply and pays a random dividend $d_t \sim N(\theta, \sigma^2)$ at time t . To model uncertainty about fundamentals, we assume that agents do not know the true mean of dividends θ and use past observations to estimate it. To keep the model tractable, we assume that the variance of dividends σ^2 is known at all times.

For each generation $n \in \mathbb{Z}$, the budget constraint at any time $t \in \{n, \dots, n + q\}$ is

$$W_t^n = x_t^n p_t + a_t^n, \quad (1)$$

where W_t^n denotes the wealth of generation n at time t , x_t^n is the investment in the risky asset (units of Lucas tree output), a_t^n is the amount invested in the riskless asset, and p_t is the price of one unit of the risky asset at time t . As a result, wealth next period is

$$\begin{aligned} W_{t+1}^n &= x_t^n (p_{t+1} + d_{t+1}) + a_t^n R \\ &= x_t^n (p_{t+1} + d_{t+1} - p_t R) + W_t^n R. \end{aligned} \quad (2)$$

⁶ Once we depart from the Bayesian paradigm, nothing guarantees our agents understand that more information increases the precision of their beliefs. If this was the case, for example, one should expect agents to also incorporate past data.

We denote the excess payoff received in $t + 1$ from investing at time t in one unit of the risky asset, relative to the riskless asset, as $s_{t+1} \equiv p_{t+1} + d_{t+1} - p_t R$. This is analogous to the equity premium in our CARA model. Using this notation, $W_{t+1}^n = x_t^n s_{t+1} + W_t^n R$.

We assume that agents maximize their per-period utility (i. e., are myopic). This assumption simplifies the maximization problem considerably and highlights the main determinants of portfolio choice generated by experience-based learning.

For a given initial wealth level W_0^n , the problem of a generation n at each time $t \in \{n, \dots, n_q\}$ is to choose x_t^n to maximize $E_t^n[-\exp(-\gamma W_{t+1}^n)]$, and hence

$$x_t^n \in \arg \max_{x \in \mathbb{R}} E_t^n[-\exp(-\gamma x s_{t+1})], \tag{3}$$

where $E_t^n[\cdot]$ is the (subjective) expectation with respect to a Gaussian distribution with variance σ^2 and a mean denoted by θ_t^n . We call θ_t^n the subjective mean of dividends, and we define it below. Note that when x_t^n is negative, generation n is short selling.

2.1. Experience-based learning

In this framework, experience-based learning (EBL) means that agents overweight realizations observed during their lifetimes when forecasting dividends and they may tilt the excess weights toward the most recent observations. For simplicity, we assume that agents only use observations realized during their lifetimes.⁷ That is, even though they observe the entire history of dividends, they choose to disregard earlier observations.⁸

EBL differs from reinforcement learning-type models in two ways. First, as already discussed, EBL agents understand the model and know all the primitives except the mean of the dividend process. Hence, they do not learn about the equilibrium; they learn in equilibrium. Second, EBL is a passive learning problem in the sense that players' actions do not affect the information they receive. This would be different if we had, say, a participation decision that links an action (participate or not) to the type of data obtained for learning. We consider this to be an interesting line to explore in the future.

Let m denote the prior belief about the mean of dividends that agents are born with and where we restrict m to be Gaussian with mean θ for tractability. With this, we construct the subjective mean of dividends of generation n at time t following the empirical evidence on Malmendier and Nagel (2011) as follows:

$$\theta_t^n \equiv (1 - \omega_{age}) \cdot m + \omega_{age} \cdot \sum_{k=0}^{age} w(k, \lambda, age) d_{t-k}, \tag{4}$$

where $age = t - n$, and where, for all $k \leq age$,

$$w(k, \lambda, age) = \frac{(age + 1 - k)^\lambda}{\sum_{k'=0}^{age} (age + 1 - k')^\lambda} \tag{5}$$

⁷ We only need agents to discount prelifetime relative to lifetime observations for our results to hold.

⁸ In our full-information setting, prices do not add any additional information. While it is possible to add private information and learning from prices to our framework, these (realistic) feature would complicate matters without necessarily adding new intuition.

denotes the weight an agent aged age assigns to the dividend observed k periods earlier and $w(k, \lambda, age) \equiv 0$ for all $k > age$. The denominator in Eq. (5) is a normalizing constant that depends only on age and on the parameter that regulates the recency bias, λ . For $\lambda > 0$, more recent observations receive relatively more weight, whereas for $\lambda < 0$, the opposite holds. Finally, $\omega_{age} \equiv \frac{age+1}{\tau+age+1}$ denotes the weight that agents assign to their experience beliefs (with $1 - \omega_{age}$ being the weight they assign to their prior belief m), which increases with age and decreases with the relative importance agents assign to their prior beliefs, regulated by parameter $\tau \in [0, \infty)$. Since the presence of prior beliefs has no qualitative implications in our model, unless otherwise stated, we study the case of $\tau = 0$ (i. e. $\omega_{age} = 1$).⁹

The following are three examples of possible weighting schemes:

Example 2.1 (Linearly declining weights, $\lambda = 1$). For $\lambda = 1$, weights decay linearly as the time lag increase; i. e., for any $0 \leq k, k + j \leq age$,

$$w(k + j, 1, age) - w(k, 1, age) = -\frac{j}{\sum_{k'=0}^{age} (age + 1 - k')}.$$

Example 2.2 (Equal weights, $\lambda = 0$). For $\lambda = 0$, lifetime observations are equal weighted; i. e., for any $0 \leq k \leq age$,

$$w(k, 0, age) = \frac{1}{age+1}.$$

Example 2.3. For $\lambda \rightarrow \infty$, the weight assigned to the most recent observation converges to 1, and all other weights converge to 0; i. e., for any $0 \leq k \leq age$,

$$w(k, \lambda, age) \rightarrow 1_{\{k=0\}}.$$

Observe that by construction, $\theta_t^n \sim N(\theta, \sigma^2 \sum_{k=0}^{age} (w(k, \lambda, age))^2)$. Hence, θ_t^n does not necessarily converge to the truth as $t \rightarrow \infty$; it depends on whether $\sum_{k=0}^{age} (w(k, \lambda, age))^2 \rightarrow 0$. This in turn depends on how fast the weights for “old” observations decay to zero (i. e., how small λ is). When agents have finite lives, convergence will not occur.

We conclude this section by showing a useful property of the weights, which is used in the characterization of our results.

Lemma 2.1 (Single-crossing property). Let $age' < age$ and $\lambda > 0$. Then the function $w(\cdot, \lambda, age) - w(\cdot, \lambda, age')$ changes signs (from negative to nonnegative) exactly once over $\{0, \dots, age' + 1\}$.

2.2. Comparison to Bayesian learning

To better understand the experience-effect mechanism, we compare the subjective mean of EBL agents to the posterior mean of agents who update their beliefs using Bayes rule. We consider two cases: full Bayesian learning (FBL), wherein agents use all the available observations to form their beliefs; and Bayesian learning from experience (BLE), where agents only use data realized during their lifetimes.

⁹ We solve the model for $\tau > 0$ in Appendix B, and we discuss the quantitative implications of prior beliefs in our model in Section 6.1.

2.2.1. Full Bayesian learners

To illustrate the comparison of EBL and FBL in a common sample, just for this analysis, we start the economy at an initial time $t = 0$ since FBL use all the available observations since “the beginning of time.” Then, all generations of FBL agents consider all observations since time 0 to form their belief. We denote the prior of FBL agents as $N(m, \sigma_m^2)$. For simplicity, all generations have the same prior, though the analysis can easily be extended to heterogeneous Gaussian priors across generations.¹⁰

The posterior mean of any generation alive at time t , denoted by $\hat{\theta}_t$, is given by

$$\hat{\theta}_t = \frac{\sigma_m^{-2}}{\sigma_m^{-2} + \sigma^{-2}t} m + \frac{\sigma^{-2}t}{\sigma_m^{-2} + \sigma^{-2}t} \left(\frac{1}{t} \sum_{k=0}^t d_k \right).$$

The belief of an FBL agent is a convex combination of the prior m and the average of all observations d_k realized since time 0. The key difference to EBL agents is that differences in personal experiences do not play a role: there is no heterogeneity in beliefs, and all generations alive in any given period have the same belief about the mean of dividends. In addition, beliefs of FBL agents are nonstationary (i. e., they depend on the time period). As $t \rightarrow \infty$, the posterior mean converges (almost surely) to the true mean. That is, with FBL the implications of learning vanish as time goes to infinity. With EBL, this is not true. Since agents have finite lives and learn from their own experiences, our model generates learning dynamics even as time diverges.

2.2.2. Bayesian learners from experience

For BLE agents, the situation is different. We assume again that each generation has a prior $N(m, \sigma_m^2)$ when they are born. Here, the posterior mean of generation n at period $t = n + \text{age}$, denoted by $\hat{\theta}_t^n$, is given by

$$\hat{\theta}_t^n = \frac{\sigma_m^{-2}}{\sigma_m^{-2} + \sigma^{-2}(\text{age} + 1)} m + \frac{\sigma^{-2}(\text{age} + 1)}{\sigma_m^{-2} + \sigma^{-2}(\text{age} + 1)} \left(\frac{1}{\text{age} + 1} \sum_{k=n}^t d_k \right).$$

The belief of a BLE generation is a convex combination of the prior m and the average of (only) the lifetime observations d_k available to date. The BLE belief coincides with belief θ_t^n of EBL when there is no recency bias, $\lambda = 0$, and the importance assigned to prior belief is $\tau = \frac{\sigma_m^2}{\sigma_m^2}$.

2.3. Equilibrium definition

We now proceed to define the equilibrium of the economy with EBL agents.

Definition 2.1 (Equilibrium). An equilibrium is a demand profile for the risky asset $\{x_t^n\}$, a demand profile for the riskless asset $\{a_t^n\}$, and a price schedule $\{p_t\}$ such that

1. Given the price schedule, $\{(a_t^n, x_t^n) : t \in \{n, \dots, n_q\}\}$ solve the generation n problem.
2. The market clears in all periods: $1 = \frac{1}{q} \sum_{n=t-q+1}^t x_t^n$ for all $t \in \mathbb{Z}$.

We focus the analysis on the class of linear equilibria (i. e., equilibria with affine prices):

Definition 2.2 (Linear equilibrium). A linear equilibrium is an equilibrium wherein prices are an affine function of dividends. That is, there exists a $K \in \mathbb{N}$, $\alpha \in \mathbb{R}$, and $\beta_k \in \mathbb{R}$ for all $k \in \{0, \dots, K\}$ such that

$$p_t = \alpha + \sum_{k=0}^K \beta_k d_{t-k}. \tag{6}$$

2.3.1. Benchmark with known mean of dividends

For the sake of benchmarking our results for EBL agents, we characterize equilibria in an economy where the mean of dividends, θ , is known by all agents (i. e., $E_t^n[d_t] = \theta \forall n, t$). In this scenario, there are no disagreements across cohorts, and the demand of any cohort trading at time t is

$$x_t^n \in \arg \max_{x \in \mathbb{R}} E[-\exp(-\gamma x s_{t+1})]. \tag{7}$$

The solution to this problem is standard and given by

$$x_t^n = \frac{E[S_{t+1}]}{\gamma V[S_{t+1}]} \tag{8}$$

for all $n \in \{t - q + 1, \dots, t\}$, and zero otherwise. Since there is no heterogeneity in cohorts’ demands and there is a unit supply of the risky asset, in any equilibrium, $x_t^n = 1$ for all $n \in \{t - q + 1, \dots, t\}$, and zero otherwise. Furthermore, there exists a unique bubble-free equilibrium with constant prices $p_t = P \forall t$, where $P = \frac{\theta - \gamma \sigma^2}{R - 1}$.

3. Toy model

To illustrate the mechanics of the model, we first highlight the main results of the paper in a simple environment, namely, for $q = 2$. We will solve the model for any $q > 1$ in the next section.

In the toy model with $q = 2$, there are three cohorts alive at each point in time: a young cohort, which enters the market for the first time; a middle-aged cohort, which is participating in the market for the second time; and an old cohort, whose agents simply consume the payoffs from their lifetime investments. At time t , the problem of generations $n \in \{t, t - 1\}$ is given by Eq. (3). It is easy to show that their demands for the risky asset are

$$x_t^n = \frac{E_t^n[S_{t+1}]}{\gamma V_t^n[S_{t+1}]}.$$

As one of our first key results in Section 4, we will show that (i) prices depend on the history of dividends and (ii) this price predictability is limited to the past dividends observed (experienced) by the oldest generation trading in the market. In other words, we show that $K = q - 1$ in Eq. (6). Anticipating this result here for $q = 2$, we have

$$p_t = \alpha + \beta_0 d_t + \beta_1 d_{t-1}. \tag{9}$$

¹⁰ The assumption of Gaussianity is also not needed but simplifies the exposition greatly.

The dependence of prices on past dividends is an important feature of our model, which is shared by many models of extrapolation and learning. A distinct feature of our model is that this dependence is intrinsically linked to the demographic structure of the economy. It matters which generations are participating in the market and how much.

The cross-sectional differences in lifetime experiences, and the resulting cross-sectional differences in beliefs, determine cohorts' trading behavior. Given the functional form for prices, we can rewrite the demands of both cohorts that are actively trading as

$$x_t^t = \frac{\alpha + (1 + \beta_0)E_t^t[d_{t+1}] + \beta_1 d_t - p_t R}{\gamma(1 + \beta_0)^2 \sigma^2}$$

$$x_t^{t-1} = \frac{\alpha + (1 + \beta_0)E_t^{t-1}[d_{t+1}] + \beta_1 d_t - p_t R}{\gamma(1 + \beta_0)^2 \sigma^2}.$$

The difference between cohorts' demand arises from their different beliefs about future dividends, $E_t^t[d_{t+1}]$ and $E_t^{t-1}[d_{t+1}]$, given by

$$E_t^t[d_{t+1}] = d_t,$$

$$E_t^{t-1}[d_{t+1}] = \underbrace{\left(\frac{2^\lambda}{1 + 2^\lambda}\right)}_{w(0,\lambda,1)} d_t + \underbrace{\left(\frac{1}{1 + 2^\lambda}\right)}_{w(1,\lambda,1)} d_{t-1}.$$

These formulas illustrate the mechanics of EBL and the cause of heterogeneity among agents. In the simplified setting, the younger generation has only experienced the dividend d_t and expects the dividends to be identical in the next period. The older generation has more experience and incorporates the previous dividend in their weighing scheme. An implication of these formulas is that the younger generations react more optimistically than older generations to positive changes in recent dividends and more pessimistically to negative changes. In Section 4.2, we show that this result continues to hold in the general model. We also see that belief heterogeneity is increasing in the change in dividends, $|d_t - d_{t-1}|$, and decreasing in the recency bias, λ . In Section 4.3, we exploit this observation to link movements in the volume of trade to belief disagreements.

We now impose the market clearing condition, $\frac{1}{2}(x_t^t + x_t^{t-1}) = 1$, to derive the equilibrium price given these demands. We use the method of undetermined coefficients to solve for $\{\alpha, \beta_0, \beta_1\}$. Setting the constants and the terms that multiply d_t and d_{t-1} to zero, we obtain a system of equations whose solution determines the price constant and the loadings of prices on present and past dividends,

$$\alpha = -\frac{\gamma(1 + \beta_0)^2 \sigma^2}{R - 1}, \tag{10}$$

$$\beta_0 = \frac{2R^2}{(R - 1)(1 + 2R - \frac{2^\lambda}{1 + 2^\lambda})} - 1, \tag{11}$$

$$\beta_1 = \frac{R\left(1 - \frac{2^\lambda}{1 + 2^\lambda}\right)}{(R - 1)\left(1 + 2R - \frac{2^\lambda}{1 + 2^\lambda}\right)}. \tag{12}$$

These solutions illustrate how the price loadings on past dividends depend on the demographics of the economy and on the magnitude of the recency bias. It is easy to derive the unconditional price volatility, which is $\sigma(p_t) = (\beta_0^2 + \beta_1^2)^{\frac{1}{2}} \sigma$, and price autocorrelation, which is $\rho(p_t, p_{t+j}) = \beta_0 \beta_1$ for $j = 1$ and $\rho(p_t, p_{t+j}) = 0$ for $j > 1$. The variance of prices is increasing in the recency bias λ , while the price autocorrelation is decreasing in the recency bias. The intuition is straightforward: as the recency bias increases, prices become more responsive to the most recent dividend, $\frac{\partial \beta_0}{\partial \lambda} > 0$, increasing price volatility, and less responsive to past dividends, $\frac{\partial \beta_1}{\partial \lambda} < 0$, decreasing price autocorrelation. In Section 5, we present an enriched version of the model with demographic shocks and discuss how these price loadings vary with the demographic structure of the economy.

4. General model

We now return to the general case (i. e., allow for any $q > 1$) and characterize the portfolio choices and resulting demands for the risky asset of the different cohorts when agents exhibit EBL. We impose affine prices, use market clearing to verify the affine prices guess, and fully characterize demands and prices. Deriving the results in the general model allows us to discuss in more detail the relation between demographics, cross-sectional asset holdings, and market dynamics. We obtain testable predictions, which we bring to the data in Section 6.

4.1. Characterization of equilibrium demands and prices

For any $s, t \in \mathbb{Z}$, let $d_{s:t} = (d_s, \dots, d_t)$ denote the history of dividends from time s up to time t . For simplicity and without loss of generality, we assume that the initial wealth of all generations is zero (i. e., $W_n^n = 0$ for all $n \in \mathbb{Z}$). At time $t \in \{n, \dots, n_q\}$, an agent of generation n determines her demand for the risky asset maximizing $E_t^n[-\exp(-\gamma x s_{t+1})]$, as in Eq. (3).

The model setup allows us to derive a standard expression for risky-asset demand:

Proposition 4.1. Suppose $p_t = \alpha + \sum_{k=0}^K \beta_k d_{t-k}$ with $\beta_0 \neq -1$. Then, for any generation $n \in \mathbb{Z}$ trading in period $t \in \{n, \dots, n_q\}$, demands for the risky asset are given by

$$x_t^n = \frac{E_t^n[s_{t+1}]}{\gamma V[s_{t+1}]} = \frac{E_t^n[s_{t+1}]}{\gamma(1 + \beta_0)^2 \sigma^2}. \tag{13}$$

The expression for the risky-asset demands in Eq. (13) allows us to derive equilibrium prices. Note that Eq. (13) implies that demands at time t are affine in $d_{t-K:t}$. It is easy to see, then, that beliefs about future dividends are linear functions of the dividends observed by each generation participating in the market, and thus prices depend on the history of dividends observed by the oldest generation in the market:

Proposition 4.2. The price in any linear equilibrium is affine in the history of dividends observed by the oldest generation

participating in the market (i. e., for any $t \in \mathbb{Z}$),

$$p_t = \alpha + \sum_{k=0}^{q-1} \beta_k d_{t-k}, \text{ with} \tag{14}$$

$$\alpha = -\frac{1}{\left(1 - \sum_{j=0}^{q-1} \frac{w_j}{R^{j+1}}\right)^2} \frac{\gamma \sigma^2}{R-1} \tag{15}$$

$$\beta_k = \frac{\sum_{j=0}^{q-1-k} \frac{w_{k+j}}{R^{j+1}}}{1 - \sum_{j=0}^{q-1} \frac{w_j}{R^{j+1}}} \quad k \in \{0, \dots, q-1\}, \tag{16}$$

where $w_k \equiv \frac{1}{q} \sum_{age=0}^{q-1} w(k, \lambda, age)$.

Proposition 4.2 establishes a novel link between the factors influencing asset prices and demographic composition. For each $k = \{0, 1, \dots, q-1\}$, one can interpret w_k as the average weight placed on the dividend observed at time $t-k$ by all generations trading at time t . As the formula also reveals, the relative magnitudes of the weights on past dividends, β_k , depend on the number of cohorts in the market, q , on the fraction of each cohort in the market, $\frac{1}{q}$, and on the extent of agents' recency bias, λ .¹¹

The main idea of the proposition is as follows. In a linear equilibrium, demands at time t are affine in dividends $d_{t-K:t}$. However, from these dividends, only $d_{t-q+1:t}$ matter for forming beliefs; the dividends $d_{t-K:t-q}$ only enter through the definition of linear equilibrium. The proof shows that, under market clearing, the coefficients accompanying older dividends $d_{t-K:t-q}$ are zero. The proposition also implies that we can apply the same restriction to demands and conclude that demands at time t only depend on $d_{t-q+1:t}$.

Note that $\frac{\partial \beta_k}{\partial R} < 0$ and $\frac{\partial \alpha}{\partial R} > 0$ for any λ . That is, if the interest rate is higher, the equilibrium price of the risky asset responds less strongly to past dividends. Furthermore, higher risk aversion γ decreases the equilibrium price by lowering α .

The following proposition establishes that, as long as agents exhibit any positive recency bias (i. e., $\lambda > 0$), the sensitivity of prices to past dividends is stronger the more recent the dividend realization.

Proposition 4.3. For $\lambda > 0$, more recent dividends affect prices more than less recent dividends (i. e., $0 < \beta_{q-1} < \dots < \beta_1 < \beta_0$).

This result reflects the fact that the dividends at time t are observed by all generations, whereas past dividends are only observed by older generations. At the same time, the extent to which prices depend on the most recent dividends varies with the level recency bias, as shown in the following Lemma.

Lemma 4.1. The effect of the most recent dividend realization on prices, β_0 , is increasing in λ , with $\lim_{\lambda \rightarrow \infty} \beta_0(\lambda) = 1/(R-1)$ and $\lim_{\lambda \rightarrow \infty} \beta_k(\lambda) = 0$ for $k > 0$.

¹¹ In our baseline model, cohorts are equally weighted. We remove this assumption in Section 5, where we analyze demographic shocks. In those examples, there is no link between the number of cohorts and the fraction of each cohort in the market.

As $\lambda \rightarrow \infty$, the average weights w_k (defined in Proposition 4.2) converge to $1_{\{k=0\}}$ for all $k = \{0, 1, \dots, K\}$. Therefore, $\beta_k \rightarrow 0$ for all $k > 0$ and $\beta_0 \rightarrow \frac{1}{R-1}$. In other words, under extreme recency bias ($\lambda \rightarrow \infty$), only the current dividend affects prices in equilibrium, while the weights on all past dividends vanish.

In Section 5, we show that the dependence of prices on more recent dividends is also increasing in the fraction of young agents in the market; that is, β_0 increases as the relative measure of the youngest cohort in the market increases.

These results on price sensitivity to past dividends, as well as the dampening effect of recency bias on cross-sectional heterogeneity, produce a range of asset pricing implications, from known puzzles, such as the predictability of stock returns and excess volatility, to new predictions about the link between asset prices and demographics. We derive and test these empirical implications in Section 6.

4.2. Cross-section of asset holdings

EBL has distinctive implications for the cross-section of asset holdings. We show that positive shocks (booms) induce a larger representation of younger investors in the market, while negative shocks (crashes) have the opposite effect. To illustrate this, we first show that younger investors react more optimistically than older ones to positive changes in recent dividends and more pessimistically to negative ones.

Proposition 4.4. For any $t \in \mathbb{Z}$ and any generations $n \leq m$ trading at t , there is a threshold time lag $k_0 \leq t-m-1$ such that for dividends that date back up to k_0 periods, the risky-asset demand of the younger generation (born at m) responds more strongly to changes than the demand of the older generation (born at n), while, for dividends that date back more than k_0 periods, the opposite holds. That is,

1. $\frac{\partial x_t^m}{\partial d_{t-k}} \geq \frac{\partial x_t^n}{\partial d_{t-k}}$ for $0 \leq k \leq k_0$ and
2. $\frac{\partial x_t^m}{\partial d_{t-k}} \leq \frac{\partial x_t^n}{\partial d_{t-k}}$ for $k_0 < k \leq q-1$.

Proposition 4.4 establishes that, for any two cohorts of investors, there is a threshold time lag up to which past dividends are weighted more by the younger generation and beyond which past dividend realizations are weighted more by the older generation.

In what follows, we extend this insight into predictions about relative stock market positions. We show that, as a result of the stronger impact of more recent shocks on the beliefs (and thus, demands) of younger generations, the relative positions of the young and the old in the market fluctuate. Let us denote the difference between generations n and $n+k$ in terms of their investment in the risky asset, as $\xi(n, k, t) \equiv x_t^n - x_t^{n+k}$. By Proposition 4.1, and some simple algebra, it follows that:

$$\xi(n, k, t) = \frac{E_t^n[\theta] - E_t^{n+k}[\theta]}{\gamma(1 + \beta_0)\sigma^2} \tag{17}$$

$\forall k = \{0, \dots, t-n\}, n = \{t-q+1, \dots, t\}$.

This formulation illustrates that the discrepancy between the positions of different generations is entirely explained

by the discrepancy in beliefs. For instance, if for some $a > 0$, $d_{n,t} \approx d_{n+a,t+a}$, then $\xi(n + a, k, t + a) \approx \xi(n, k, t)$.¹²

The next result shows that, among generations born and growing up in “boom times,” understood as periods of increasing dividends, the younger generations have a relatively higher demand for the risky asset than the older generations. The reverse holds for depression babies (i. e., generations born during times of contraction). In times of depression, the younger generations exhibit a particularly low willingness to invest in the risky asset, relative to older generations born during those times.

Proposition 4.5. Suppose $\lambda > 0$. Consider two points in time $t_0 \leq t_1$ such that dividends are nondecreasing from t_0 up to t_1 . Then for any two generations $n \leq n + k$ born between t_0 and t_1 , the demand of the older generation for the risky asset (x_t^n) is lower than the demand of the younger generation (x_t^{n+k}) at any point $n \leq t \leq t_1$; i. e., $\xi(n, k, t) \leq 0$. On the other hand, if dividends are nonincreasing, then $\xi(n, k, t) \geq 0$.

The proposition illustrates that, while boom times tend to make all cohorts growing up in such times more optimistic, the effect is particularly strong for the younger generations. This induces them to be overrepresented in the market for the risky asset. The opposite holds during times of downturn.

4.3. Trade volume

We now study how learning and disagreements affect the volume of trade observed in the market. We consider the following definition of the total volume of trade in the economy:

$$TV_t \equiv \left(\frac{1}{q} \sum_{n=t-q}^t (x_t^n - x_{t-1}^n)^2 \right)^{\frac{1}{2}}, \tag{18}$$

with $x_{t-1}^t = 0$. That is, trade volume is the square root of the weighted sum (squared) of the change in positions of all agents in the economy. Using this definition, we characterize the link between trade volume and belief heterogeneity.

Proposition 4.6. The trade volume defined in Eq. (18) can be expressed as

$$TV_t = \left(\frac{\chi^2}{q} \sum_{n=t-q}^t \left((\theta_t^n - \theta_{t-1}^n) - \frac{1}{q} \sum_{\tilde{n}=t-q}^t (\theta_t^{\tilde{n}} - \theta_{t-1}^{\tilde{n}}) \right)^2 + \frac{1}{q} (x_t^t)^2 + \frac{1}{q} (x_{t-1}^{t-q})^2 \right)^{\frac{1}{2}}, \tag{19}$$

where $\chi = \frac{1}{\gamma\sigma^2(1+\beta_0)}$, $\theta_{t-1}^t = \theta_t^{t-q} = 0$.

¹² This last claim follows since the intertemporal change in discrepancies between sets of generations of the same age, $\xi(n + a, k, t + a) - \xi(n, k, t)$ for $a > 0$, is given by $(\sum_{j=0}^{t-n-k} \{w(j, \lambda, t - n) - w(j, \lambda, t - n - k)\} (d_{t+a-j} - d_{t-j}) + \sum_{j=t-n-k+1}^{t-n} w(j, \lambda, t - n) (d_{t+a-j} - d_{t-j})) / (\gamma(1 + \beta_0)\sigma^2)$.

Expression (19) illustrates that the presence of EBL induces trade through changes in beliefs, which in our framework are driven by shocks to dividends. More specifically, when the change in a cohort’s beliefs is different from the average change in beliefs, trade volume increases. That is, trade volume increases in the dispersion of changes in beliefs.

To understand the drivers of trade volume, we need to understand not only the demands of agents that enter and exit the market but also, most importantly, how beliefs across cohorts change in response to a given shock. From our previous analysis, it follows that an increase (decrease) in dividends induces trade when it makes young agents more optimistic (pessimistic) than old agents. This mechanism is solely due to the presence of EBL since it is essential that each generation reacts differently to the same dividend.

We formalize this insight in the following thought experiment capturing the reaction to a dividend shock that occurs after a long period of stability.

4.3.1. Thought experiment

Suppose that for $t - t_0 > q$, $d_{t_0} = d_{t_0+1} = \dots = d_{t-1} = \bar{d}$ and $d_t \neq \bar{d}$. Hence, all generations alive at time $t - 2$ and $t - 1$ have only observed a constant stream of dividends \bar{d} over their lifetimes so far. Therefore, $E_{t-2}^n[d_{t-1}] = E_{t-1}^n[d_t] = \bar{d}$ for all $n \in \{t - 1 - q, \dots, t - 1\}$, and thus trade volume in $t - 1$ is simply given by the demand of the youngest (entering) and the oldest (exiting) agents.

What happens at time t , when a dividend $d_t \neq \bar{d}$ is observed? For each generation n trading at time t and at time $t - 1$, i. e., for $n = \{t - q + 1, \dots, t - 1\}$, beliefs are given by $E_t^n[d_{t+1}] = w(0, \lambda, t - n)(d_t - \bar{d}) + \bar{d}$ and $E_{t-1}^n[d_t] = w(0, \lambda, t - 1 - n)(d_{t-1} - \bar{d}) + \bar{d}$, which implies the following change in cohort n ’s beliefs: $E_t^n[d_{t+1}] - E_{t-1}^n[d_t] = w(0, \lambda, t - n)(d_t - \bar{d})$. Trade volume in t is therefore

$$TV_t = \left[\frac{\chi^2 (d_t - \bar{d})^2}{q} \sum_{n=t-q+1}^{t-1} \left(w(0, \lambda, t - n) - \frac{1}{q} \sum_{\tilde{n}=t-q+1}^{t-1} w(0, \lambda, t - \tilde{n}) \right)^2 + \frac{1}{q} (x_t^t)^2 + \frac{1}{q} (x_{t-1}^{t-q})^2 \right]^{\frac{1}{2}}. \tag{20}$$

This thought experiment pins down two aspects of the link between the volatility in beliefs and trade volume: first, the trade volume increases proportionally to the change in dividends, $|d_t - \bar{d}|$, independent of whether the change is positive or negative, and also proportionally to a function that reflects the dispersion of the weights agents assigned to the most recent observation in their belief formation process. Second, the increase in trade volume generated by a given change in dividends depends on the level of recency bias of the economy, which is captured by λ . For example, as $\lambda \rightarrow \infty$, the dispersion in weights decreases as $w(0, \lambda, \text{age}) \rightarrow 1$ for all $\text{age} \in \{0, \dots, q - 1\}$. Thus, our results suggest that higher recency bias, λ , should generate

lower trade volume responses for a given shock to dividends and vice versa.

5. Market participation

The results derived so far illustrate a key feature of EBL: the demographic structure of an economy, and in particular the cross-sectional composition of investors, affect equilibrium prices, demand, and trade volume in a predictable direction.

In this section, we explore the link between market demographics and financial market outcomes by considering an unexpected increase in the fraction of young market participants (e. g., due to a baby boom or a generation-specific event drawing a certain generation into the stock market).¹³ The goal of this exercise is to understand how a larger fraction of young market participants affects market dynamics.

For ease of illustration, we focus again on our $q = 2$ economy. We denote the mass of young agents at any time t by y_t and the total mass of agents at t by $m_t = y_t + y_{t-1}$. We consider a one-time unexpected (exogenous) shock to the mass of young agents in the market at time τ .¹⁴ For all $t < \tau$ and $t > \tau + 1$, instead, $y_t = y$ and thus $m_t = 2y = m$.

We know from our previous results that when the market has equal-sized cohorts, prices are given by $p_t = \alpha + \beta_0 d_t + \beta_1 d_{t-1}$, with $\{\alpha, \beta_0, \beta_1\}$ given by Eqs. (10)–(12). Here, prices follow this path for $t > \tau + 1$ and since the shock at time τ is unexpected, for $t < \tau$ as well. For these time periods, the market is as described in Section 3. We are left to characterize demands and prices for τ and $\tau + 1$, when the larger young generation enters the market and when this generation becomes old, respectively. We make the following guesses:

$$p_\tau = a_\tau + b_{0,\tau} d_\tau + b_{1,\tau} d_{\tau-1}, \quad (21)$$

$$p_{\tau+1} = a_{\tau+1} + b_{0,\tau+1} d_{\tau+1} + b_{1,\tau+1} d_\tau. \quad (22)$$

We solve the problem by backward induction. Note that the form of agents' demands remains unchanged. By imposing market clearing in $\tau + 1$, with mass y of young agents and y_τ of old agents, and using the method of undetermined coefficients, we obtain

$$a_{\tau+1} = \alpha \left[1 + \frac{R-1}{m_\tau} \right],$$

¹³ We also analyze the implications of a growing market population, as opposed to a one-time market demographic shock. In Appendix C, we show that population growth generates a positive trend in prices, which is independent of experience effects: the growing mass of agents increases the demand for the risky asset, and hence prices adjust to clear markets, since risky-asset supply is assumed to be constant. While the positive trend is independent of experience effects, EBL does affect the path of the prices fluctuating around this trend. In particular, we find that the relative reliance of prices on the most recent dividend is increasing in the population growth rate.

¹⁴ In reality, the participation of young agents in the market could also be determined endogenously (e. g., by entry costs). While the forces described in this section would still be present in such a scenario, other forces may be at play as well. The study of these interactions is out of the scope of this paper.

$$b_{0,\tau+1} = \beta_0 \left[1 + \frac{1}{R} \left(\frac{m_\tau - y_\tau}{m_\tau} + \frac{y_\tau}{m_\tau} \omega - \frac{y}{m} (1 + \omega) \right) \right] \\ + \frac{1}{R} \left(\frac{m_\tau - y_\tau}{m_\tau} + \frac{y_\tau}{m_\tau} \omega - \frac{y}{m} (1 + \omega) \right), \\ b_{1,\tau+1} = \beta_1 \frac{y_\tau}{m_\tau} \frac{m}{y},$$

where $\omega \equiv \frac{2^\lambda}{1+2^\lambda}$ and $m_\tau = y + y_\tau$. Note that for $y_\tau = y$, the coefficients are, as in the baseline model, given by Eq. (10)–(12). The above expressions show that the total mass of agents m_τ only affects the price constant, while the price loadings depend on the fraction of young agents in the market, y_τ/m_τ . We impose market clearing in τ , with mass y_τ of young agents and y of old agents. Using the method of undetermined coefficients, we obtain

$$a_\tau = \frac{1}{R} \left[a_{\tau+1} - \frac{\gamma (1 + b_{0,\tau+1})^2 \sigma^2}{m_\tau} \right], \\ b_{0,\tau} = \frac{1}{R} (1 + b_{0,\tau+1}) \left(\frac{y_\tau}{m_\tau} + \frac{m_\tau - y_\tau}{m_\tau} \omega \right) \\ + \frac{1}{R^2} (1 + \beta_0) \frac{y_\tau}{m_\tau} (1 - \omega), \\ b_{1,\tau} = \frac{1}{R} (1 + b_{0,\tau+1}) \frac{m_\tau - y_\tau}{m_\tau} (1 - \omega).$$

Fig. 2 shows how the reliance of prices on past dividends changes with the fraction of young agents in the market at time τ . We see that as the fraction of young people in the market increases ($y_\tau > 0.5$), the more current dividends matter more relative to past dividends for the determination of prices (i. e., b_0^τ increases while b_1^τ decreases). Consistent with this, when the τ -generation becomes old, prices depend less on contemporaneous dividends and more on past dividends (i. e., $b_0^{\tau+1}$ decreases while $b_1^{\tau+1}$ increases). Finally, an increase in the overall market population, and thus demand for the risky asset, generates a level increase in prices captured an increase in the price constant both at τ and $\tau + 1$. All predictions are reversed when the fraction of young agents in the market decreases ($y_\tau < 0.5$). These results are consistent with Collin-Dufresne et al. (2017), who show both theoretically and empirically that the sensitivity of the price-dividend ratio to macro-shocks increases with the relative fraction of young market participants.

6. Empirical implications

In this section, we analyze the empirical implications of our model. The analysis consists of two approaches. First, we turn to three main asset pricing puzzles established in prior empirical literature: the predictability of stocks returns, the predictability of the dividend-price ratio, and the excess volatility puzzle. We show that the EBL model can quantitatively match these empirical findings and, moreover, it generates refined predictions relating these features to the demographic composition of investors. Second, we test the additional, novel predictions generated by our model regarding the implications of the demographic composition for the predictability of the price-dividend ratio, trade volume, and the cross-section of asset holdings. We

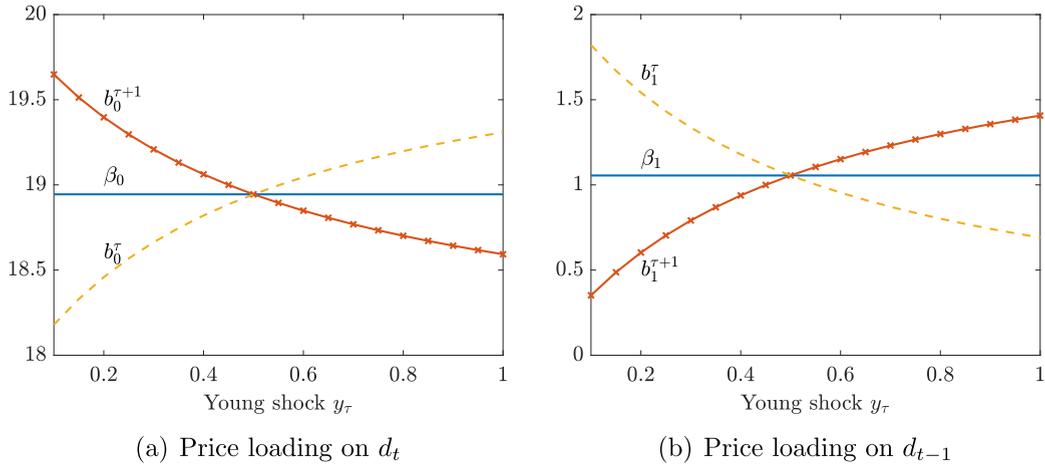


Fig. 2. Demographic shocks and price coefficients. Notes. This figure plots coefficients $\{\beta_0, b_{0,\tau}, b_{0,\tau+1}\}$ in Panel (a) and $\{\beta_1, b_{1,\tau}, b_{1,\tau+1}\}$ in Panel (b) as a function of the demographic shock y_τ . The results are for $\gamma = 0.5$, $\lambda = 3$, and $R = 1.05$.

show that these predictions are in line with evidence from micro-level data in the SCF and the CRSP.

6.1. Quantitative implications for asset pricing moments

We first show that EBL can explain key asset pricing puzzles. As the CARA-normal framework is not well suited for a thorough calibration exercise, we follow the approach of Campbell and Kyle (1993) and Barberis et al. (2015), among others, to compute the moments of interest generated by our model and contrast them with the data. As in these papers, we define quantities in terms of differences, rather than ratios, since variables in the model proxy for the log of their values in the data. We use capital letters P and D to denote prices and dividends, while small letters denote their logs, $p = \log(P)$ and $d = \log(D)$. For example, instead of stock returns, we measure price changes Δp and, instead of the price-dividend ratio P/D , we study the difference $p - d$.

A distinguishing feature of our model is that it establishes a link between the age profile of agents participating in the stock market and the factors that determine prices.¹⁵ Another feature of our model is the small number of parameters to be set for generating numerical results. Following Barberis et al. (2015), we choose the following parameter values for our numerical solutions: the gross risk-free interest rate is $R = 1.05$, the volatility of dividends is $\sigma = 0.25$, and the coefficient of risk aversion is $\gamma = 2$. We show our estimates for $\lambda \in \{1, 3\}$ and for $q \in \{2, 40\}$.

6.1.1. Predictability of excess returns

A first prominent stylized fact about stock market returns, established by Campbell and Shiller (1988), is that the dividend-price ratio predicts future returns with a positive sign. EBL rationalizes such predictability and at the same time limits it to those dividend realizations experienced by the oldest cohort participating in the market.

To relate the predictability generated in our model to the existing empirical evidence, and to show how it varies with the demographic composition of investors, we calculate the following measure of comovement between the analogs of the dividend-price ratio and returns, namely, between dividend-price differences $d_t - p_t$ Jin, 2015 and price changes $p_{t+z} - p_t$ over return horizon z :

$$B^R(z) \equiv \frac{\text{Cov}(d_t - p_t, p_{t+z} - p_t)}{\text{Var}(d_t - p_t)}. \tag{23}$$

We compute $B^R(z)$ for different horizons using Eq. (14) from Proposition 4.2. Fig. 3 plots $B^R(z)$ for z ranging from 1 to 40, for two different levels of recency bias, $\lambda \in \{1, 3\}$, in an economy with $q = 40$. Given the number of cohorts, the obtained comovements can be interpreted as annual (e. g., $z = 1$ as a one-year horizon). As the figure shows, the EBL model generates a positive (and strong) relation between the dividend-price ratio and returns, which increases with the return horizon.

The predicted patterns are consistent with the empirical findings described in Cochrane (2011). To show this, we estimate $B^R(z)$ on US stock market data from 1960 to 2013 (from Robert Shiller’s website), using the approach of Cochrane (2011). We regress the log of the dividend-price ratio on the log of returns for different horizons, $z \in \{1, 5, 10\}$. We find that the slope of these regressions are $\hat{B}^R(1) = 0.1$, $\hat{B}^R(5) = 0.3$, and $\hat{B}^R(10) = 0.6$, respectively.¹⁶ We note that the predictability of excess returns under EBL is an equilibrium phenomenon that stems solely from our learning mechanism and not from, say, a built-in dependence on dividends or past returns. Similar to prior theoretical approaches, such as the overextrapolation model of Barberis et al., 2015, 2018, our explanation relies on agents’ overweighting recent realizations.

¹⁵ The link between demographics and price features is also studied in Collin-Dufresne et al. (2017).

¹⁶ We also simulate our model, replicate the regression with simulated data, and verify that the R^2 increases with the horizon both in the regressions with US and simulated data. In particular, the R^2 of the US data regressions are $\{0.02, 0.1, 0.2\}$ respectively, while with the simulated data, $\{0.06, 0.23, 0.35\}$.

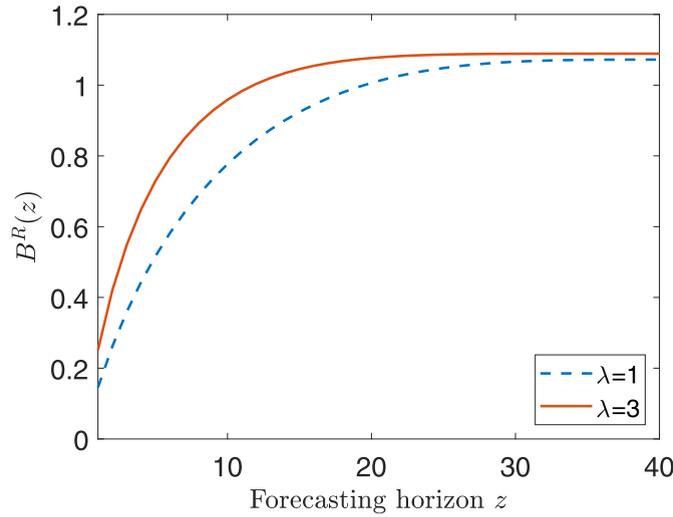


Fig. 3. Predictability of $d_t - p_t$ for $p_{t+z} - p_t$ Notes. This figure plots the coefficient $B^R(z)$ over varying horizons z for two levels of recency bias, $\lambda \in \{1, 3\}$, in a $q = 40$ economy with $R = 1.05$.

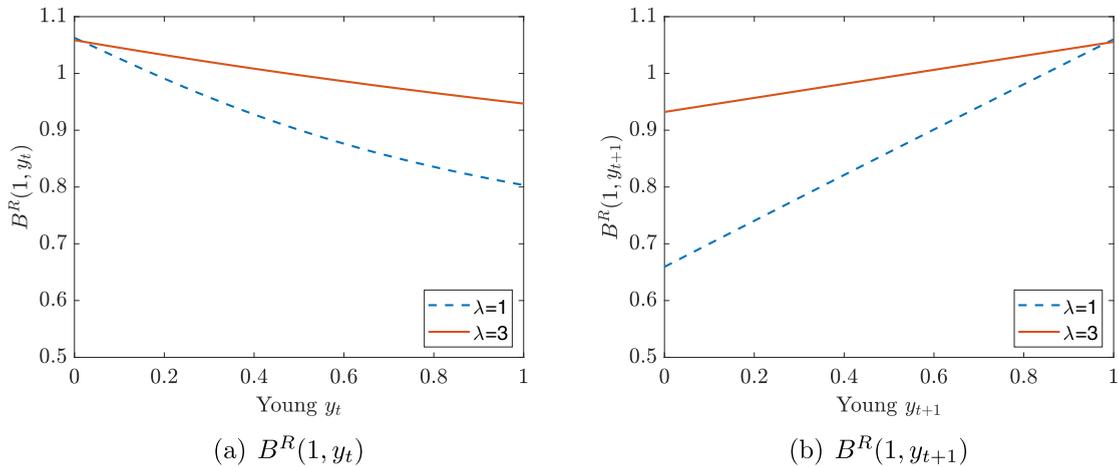


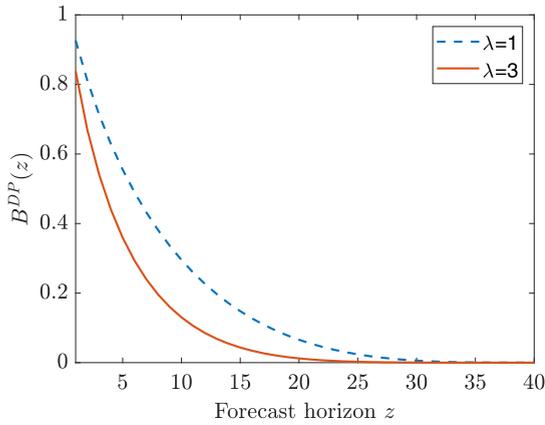
Fig. 4. Young market participants and the predictability of $D_t - P_t$ for $P_{t+1} - P_t$ Notes. This figure plots our predictability coefficients for two levels of recency bias, $\lambda \in \{1, 3\}$ in the $q = 2$ economy with $R = 1.05$. Panel (a) shows how the predictability of future returns, $B^R(1, y_t)$, varies with the fraction y_t of young agents in the market at time t , while Panel (b) shows how this predictability varies with the fraction y_{t+1} of young agents in the market at time $t + 1$. The results are calculated for $y = 0.5$.

We now turn our attention to a comparative static analysis of how our proposed measure of return predictability, given by Eq. (23), changes with the fraction of young market participants in an economy with $q = 2$.¹⁷ To do this, we use the framework of Section 5, where we study the effect of a one-time unexpected shock to the mass of young market participants.¹⁸ We first consider the case where the fraction of young agents is y at all times before and after t , and $y_t \neq y$ at time t . Therefore, the comparative statics exercise amounts to computing a version of Eq. (23), but where the moments are conditioned on the y_t shock.

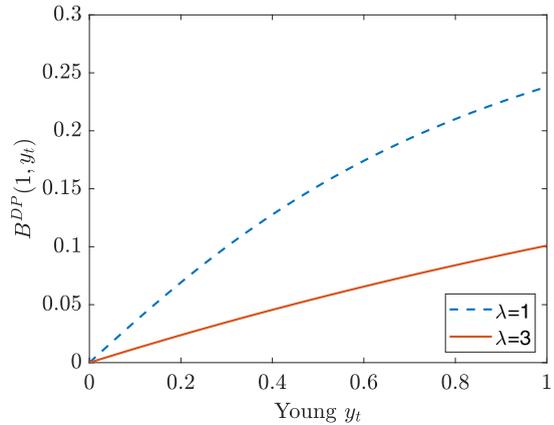
We denote this conditional predictability measure by $B^R(z, y_t)$. This demographic shock will impact $B^R(z, y_t)$ through its effect on the pricing β 's, which we analyze in detail in Section 5. Fig. 4(a) plots $B^R(1, y_t)$ as a function of the fraction of young agents at time t , y_t . As the plot shows, the predictability of next period's return decreases in the fraction of young market participants in the current period. The increase in the fraction of young market participants at time t increases the covariance between $d_t - p_t$ and returns $p_{t+1} - p_t$, as the new mass of young agents increases the sensitivity of price p_t to dividends at time t . However, this also increases the variances of $d_t - p_t$. These two effects go in opposite directions, and the latter effect dominates.

To shed more light on these two effects, we study what happens with the predictability measure when the fraction of young agents changes at time $t + 1$ as opposed to time

¹⁷ Note that with $q = 2$ the comovements cannot be interpreted as annual; in particular, $z = 1$ captures approximately a 15- to 20-year horizon.
¹⁸ We acknowledge the very stylized nature of the modeling of changes in market demographics. A more thorough analysis is beyond the scope of our baseline model and left for future work.



(a) $B^{DP}(z)$ with $q = 40$



(b) $B^{DP}(1, y_t)$ with $q = 2$

Fig. 5. $P_t - D_t$ autocorrelation Notes. This figure plots the coefficient B^{PD} for two levels of recency bias, $\lambda \in \{1, 3\}$. Panel (a) shows how $B^{PD}(z)$ varies with the investment horizon z for $q = 40$. Panel (b) shows how $B^{PD}(1)$ varies with the fraction y_t of young agents in the market at time t for $q = 2$. The results are calculated for $\gamma = 0.5$ and $R = 1.05$.

t ; that is, the fraction of young agents is y at all times before and after $t + 1$, and $y_{t+1} \neq y$ at time $t + 1$. With a slight abuse of notation we denote this predictability measure as $B^R(z, y_{t+1})$, and we plot it as a function of y_{t+1} in Fig. 4(b). In this case, the predictability measure increases with the fraction of young agents. As before, the increase in the fraction of young market participants at time $t + 1$ increases the covariance between $d_t - p_t$ and $p_{t+1} - p_t$ as the new mass of young agents increases the sensitivity of price p_{t+1} to dividends at time $t + 1$. However, the variance of $p_t - d_t$ is not affected by y_{t+1} , and thus the “covariance effect” dominates.

We thus conclude that return predictability not only is affected by the demographic composition of market participants, but the effect is sensitive to the timing of the market participation shock.

6.1.2. Predictability of price-dividend ratio

In addition to the predictability of returns, we can also compute the predictability of the price-dividend (P/D) ratio implied by the model. That is, we relate past P/D ratios to future P/D realizations and analyze the persistence of the P/D ratio. In particular, we study how this predictability of P/D ratios varies with the investment horizon and with the fraction of young people in the market. Our measure of predictability is constructed as follows:

$$B^{PD}(z) = \frac{\text{cov}(p_{t+z} - d_{t+z}, p_t - d_t)}{\text{var}(p_t - d_t)} \quad (24)$$

We first calculate how B^{PD} varies with the horizon z . Panel (a) of Fig. 5 displays how B^{PD} varies for different horizons z , and for different levels of recency bias, $\lambda \in \{1, 3\}$, in an economy with $q = 40$. The large q allows us to relate the obtained correlations to annual correlations. As in the data, we obtain that the P/D is highly autocorrelated at short lags, with the autocorrelation being zero at longer horizons. We see that as we reduce the number of cohorts in the market (see Panel (b)), or their horizon, autocorrelations are lower. Furthermore, B^{PD} decreases in the extent of recency bias present in the population for all q , as

prices become less sensitive to the most recent dividends. A direct implication is that the dividend-price ratio is positively correlated only with lagged realizations where the number of periods lagged is below the number of cohorts in the market.

We then turn to the analysis of the role of the demographic structure of the market for the $q = 2$ economy. To do so, we compute our conditional predictability measure using the results from Section 5, which we denote by $B^{DP}(1, y_t)$, and analyze how it varies with the fraction of young agents in the market at time t , y_t . As shown in Panel (b) of Fig. 5, $B^{DP}(1, y_t)$ increases with y_t , and the effect is weaker under higher recency bias.

6.1.3. Price dynamics

A third set of asset pricing implications are related to the dynamics of prices and, in particular, the excess volatility puzzle. As is standard in the literature, we analyze the volatility of the log price growth and of the log P/D ratio, and the volatility of log prices relative to that of log dividends, both in the model and in the data. To do so, we use historical price and dividend data from Robert Shiller’s website, where all log series are detrended. Our stylized model generates ample volatility relative to our benchmark economy and to the data, the main reason being that agents’ beliefs are extremely volatile when they do not put any weight on their prior belief ($\tau = 0$ in Eq. (4)), which would operate as an anchor. To highlight the quantitative effects of prior beliefs, we compute the model generated volatilities when we vary the importance that agents assign to their prior beliefs.¹⁹ Table 1 presents our results for an economy with $\lambda = 1$, $q = 40$, and for different levels of prior relevance, captured by τ .

We see that EBL generates ample excess volatility in prices, returns, and P/D ratios. This can be seen by comparing the data with the model with no prior beliefs

¹⁹ The solution to the model with prior beliefs is described in detail in Appendix B.

Table 1
Excess volatility.

Relative prior precision: τ	$\sigma(p-d)$	$\sigma(\Delta p)$	$\sigma(d)/\sigma(p)$
	$q = 40, \lambda = 1$		
0	0.96	1.77	0.04
1	0.71	1.32	0.06
5	0.41	0.76	0.19
10	0.29	0.53	0.39
20	0.22	0.34	0.93
30	0.18	0.20	2.55
Data	0.30	0.18	0.47

($\tau = 0$). However, if we allow agents to have prior beliefs, the model can generate moments more in line with the data. From these findings, we conclude that EBL has the ability to generate volatility in line with the data.

6.2. Demographics and price-dividend predictability

The predictability results in the previous section are consistent with the findings of [Cassella and Gulen \(2018\)](#), who find a positive relation between their market-wide measure of return extrapolation and the relative participation of young versus old investors in the stock market. Our model of EBL goes beyond a rationalization of the evidence on agents extrapolating from past dividends (cf. also [Greenwood and Shleifer, 2014](#)). It puts structure on the extent of such extrapolation exhibited by different market participants and links it to market demographics. We now bring this prediction to the data and show that it is aligned with empirical observations.

We want to test whether the predictive power of the lagged P/D ratios for the current one depends on the relative representation of younger versus older generations in the market in the manner predicted by the model. EBL predicts that the correlation between future and current lags is higher when the current share of young market participants is large. Moreover, the model generates the heuristic that young people put little weight on observations of the “distant” past (cf. [Proposition 4.4](#)).

To test these predictions, we regress the log of the P/D ratio onto lags of itself interacted with a dummy variable that indicates a larger presence of young people in the market. To model the dynamics of the P/D process, we depart from the standard linear AR models and postulate a Markov-switching regime (MSR) model, which allows us to capture richer nonlinear dynamics in a tractable way.²⁰ The regression model is thus given by

$$p_{t+1} - d_{t+1} = \mu(S_{t+1}) + \sum_{j=1}^3 (p_{t+1-j} - d_{t+1-j})(\beta_j + \delta_j \times Y_{t+1}) + \sigma \epsilon_{t+1}, \quad (25)$$

where p_t and d_t denote the log of dividends and prices at time t , respectively; $S_{t+1} \in \{0, 1\}$ is an unobserved state that evolves according to a Markov transition kernel Q ; Y_t is a dummy variable that takes value one if the share of young generations participating in the market at time t is

Table 2
Markov-switching regime (MSR) model.

This table shows estimation results for model specification (25), where $p_t - d_t$ is the log of the price-to-dividend ratio and is regressed onto lags of itself interacted with a demographic dummy variable. Y_t is the fraction of young people, which we define as an indicator equal to one when the fraction of investors under 50 is larger than 0.5 (in column 1) or as an indicator equal to one when the fraction of wealth of investors below 50 is larger than their 1960–2013 sample average (in column 2). The demographic data including age and wealth (liquid assets) of stock market participants are from the SCF and the stock data are from Robert Shiller's website. Standard errors are in parentheses. * and ** denote significance at the 10% and 5% levels, respectively.

	Dependent variable: $p_t - d_t$	
	(1) Y_t age-based	(2) Y_t age/wealth-based
δ_1	0.701** (0.154)	0.475* (0.252)
δ_2	-0.013 (0.146)	-0.115 (0.366)
δ_3	-0.745** (0.115)	-0.329 (0.232)
β_1	0.377** (0.120)	0.622** (0.159)
β_2	-0.216** (0.088)	-0.074 (0.136)
β_3	0.714** (0.093)	0.249** (0.099)
$\mu(S_1)$	5.089** (1.554)	5.741** (1.812)
$\mu(S_2)$	19.450** (3.070)	18.350** (4.768)
σ	3.812 (0.388)	4.343 (0.600)
Q_{11}	0.956 (0.026)	0.978 (0.017)
Q_{21}	0.365 (0.204)	0.206 (0.154)
N	51	51

large relative to the participation of older generations, and zero otherwise; and we assume $\epsilon_{t+1} \sim N(0, 1)$. The parameters, $(\{\mu(s)\}_{s \in \{0,1\}}, \sigma, Q, \{\beta_j, \delta_j\}_{j=1}^3)$, are jointly estimated using maximum likelihood (see, e. g., [Hamilton, 1994](#) for details).

We consider two dummies for the relative representation of younger generations in the market. First, we compute the ratio of investors who are less than 50 years old in the total population and construct an indicator that equals one if their share is bigger than 50% (or, for robustness, bigger than 55% or 60%). Second, we calculate young investors' share of liquid wealth and use an indicator that equals one if their liquid wealth share is above its sample average (or, for robustness above 90% or 110% of the sample average). Details on the variable construction and robustness checks are in [Appendix D](#).

The theoretical prediction of our model is that the correlation between future and current lags should be higher when the current share of young market participants is large. This translates into the hypothesis that $\delta_1 > 0$ in the estimation model in [Eq. \(25\)](#).

The estimated values are reported in [Table 2](#). In column (1), we use the fraction of young people in the population and, in column (2), the fraction of their wealth

²⁰ For a more thorough discussion of MSR, see [Hamilton \(1989\)](#).

to proxy for the relative representation of younger people in the market. In both cases, the estimates provide evidence in favor of the model hypothesis. We estimate a positive δ_1 coefficient, which is either significant at the 5% or at the 10% level. Moreover, considering all three coefficients $(\delta_i)_{i=1}^3$ jointly, a roughly “decreasing” pattern emerges: δ_1 is typically positive, δ_2 is typically nonsignificant, and δ_3 is negative or insignificantly negative, consistent with the heuristics that young people put little weight on observations of the distant past. Thus, in periods when their share is relatively large, the correlation between future and distant past values is weakened.

6.3. Cross-section of asset holdings and trade volume

We now turn to the novel empirical predictions of the EBL model about the cross-section of equity holdings and stock turnover. We investigate two sets of predictions that are directly testable and jointly hard to generate by alternative models.

The first prediction is that cross-sectional differences in the demand for risky securities reflect cross-sectional differences in lifetime experiences of risky payoffs. That is, cohorts with more positive lifetime experiences are predicted to invest more in the risky asset than cohorts with less positive experiences (Proposition 4.1). We test this both in terms of stock market participation (extensive margin) and in terms of the amount of liquid assets invested in the stock market (intensive margin). The second prediction is that changes in the cross-section of experience-based beliefs generate trade (Proposition 4.6).

To test these model predictions, we combine historical stock returns data from Robert Shiller’s website with SCF data on stock holdings and CRSP data on stock turnover. The key explanatory variable is a measure of cohorts’ lifetime experiences of risky-asset payoffs. Theoretically, dividends in the Lucas tree economy capture the performance of the risky asset or the stock market. Empirically, dividend payments do not necessarily reflect how well firms are doing. For example, firms have incentives to smooth dividends and also to retain earnings rather than distribute them. In other words, dividends in our model do not translate one-to-one to the dividend payments recorded in CRSP. We therefore use an array of empirical measures to capture the performance of the risky asset in our model: (1) annual stock market returns, (2) real dividends, (3) real earnings, and (4) US GDP. We obtain the first three series from Robert Shiller’s website and the nominal GDP data from the Federal Reserve Bank of St. Louis (for 1929–2016) and Historical Statistics of the United States Millennial Edition Online (for 1871–1928). We convert nominal GDP into real GDP using Shiller’s consumer price index variable.

Dividends in our model are best interpreted as the performance of the risky asset at medium frequencies. Therefore, we use the Christiano and Fitzgerald (2003) band-pass filter and remove stochastic cycles at frequencies lower than two years and higher than eight years,²¹ for all nonstationary series (dividend, earnings, and GDP).

²¹ These are the default frequencies for the CF-filter. We also remove a linear trend of the series before applying the filter and, in addition, work

To construct the experienced returns, dividends, earnings, and GDP of different generations over the course of their lives, we apply the formulas from Eqs. (4) and (5) for $\omega_{age} = 1$. We calculate generation-specific weighted averages, employing both linearly declining weights ($\lambda = 1$), and a steeper weighting function ($\lambda = 3$), corresponding to the range of empirical estimates in Malmendier and Nagel (2011).

6.3.1. Stock market participation

We test the first prediction relating the differences in lifetime experiences between older and younger cohorts (i. e., those above 60 and those below 40 years of age) to the differences in their stock market investment. Our source of household-level micro data is the cross-sectional data on asset holdings and various household background characteristics in the SCF. We use all waves of the modern, triannual SCF, available from the Board of Governors of the Federal Reserve System since 1983. We follow the variable construction of Malmendier and Nagel (2011) and extend their analysis to the most recently released data. In addition, we employ some waves of the precursor survey, available from the Inter-university Consortium for Political and Social Research at the University of Michigan since 1947. We use all survey waves that include age and stock market participation.²²

For the extensive margin of stock holdings, we construct an indicator of stock market participation. It equals one when a household holds more than zero dollars worth of stocks. We define stock holdings as the sum of directly held stocks (including stock held through investment clubs) and the equity portion of mutual fund holdings, including stocks held in retirement accounts (e. g., IRA, Keogh, and 401(k) plans).²³

For the intensive margin of stock holdings, we calculate the fraction of liquid assets invested in stocks as the share of directly held stocks plus the equity share of mutual funds, using all surveys from 1960–2013 other than 1971. Liquid assets are defined as the sum of stock holdings, bonds, cash, and short-term instruments (checking and savings accounts, money market mutual funds, certificates of deposit). In these analyses of the intensive margin, we drop all households that have no money in stocks.

For both the young and old age group, we calculate their experience and their stock market investment as a weighted average across cohorts, with the weight variable provided in the SCF. The weighted estimates are representative of the US population.²⁴

with the natural logarithm of earnings and GDP to remove nonlinearities in these series. In unreported analyses, we also use the natural logarithm of dividends and obtain very similar results.

²² Those are 1960, 1962, 1963, 1964, 1967, 1968, 1969, 1970, 1971, and 1977.

²³ For 1983 and 1986, we need to impute the stock component of retirement assets from the type of the account or the institution at which they are held and allocation information from 1989. From 1989 to 2004, the SCF offers only coarse information on retirement assets (e. g., mostly stocks, mostly interest bearing, or split), and we follow a refined version of the Federal Reserve Board’s conventions in assigning portfolio shares. See Malmendier and Nagel (2011) for more details.

²⁴ The 1983–2013 SCF waves oversample high-income households with significant stock holdings. The oversampling is helpful for our analysis of

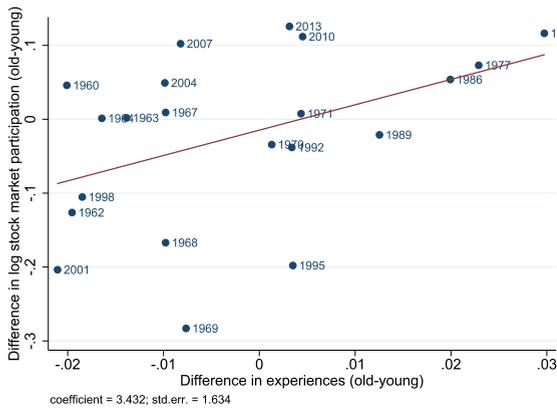
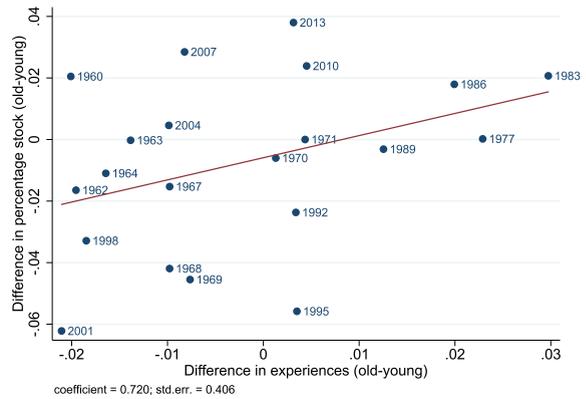
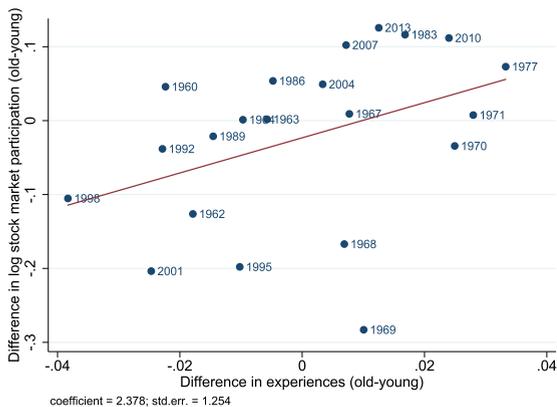
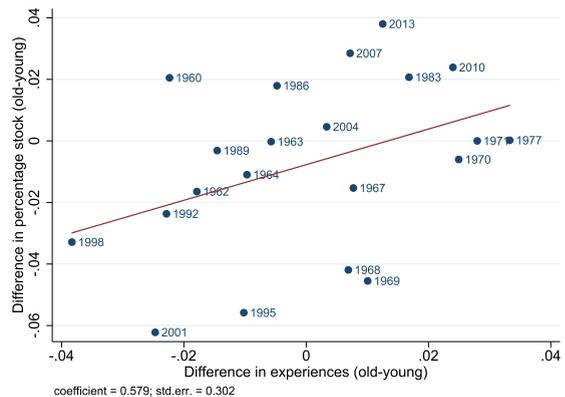
(a) Stock market participation ($\lambda = 1$)(b) Fraction invested in stock ($\lambda = 1$)(c) Stock market participation ($\lambda = 3$)(d) Fraction invested in stock ($\lambda = 3$)

Fig. 6. Experienced returns and stock holdings Notes. *Difference in experienced returns* is calculated as the lifetime average experienced returns of the S&P500 Index as given on Robert Shiller's website, using declining weights with either $\lambda = 1$ or $\lambda = 3$ as in Eq. (5). *Stock market participation* is measured as the fraction of households in the respective age groups that hold at least \$1 of stock ownership, either as directly held stock or indirectly (e. g. via mutuals or retirement accounts). *Fraction invested in stock* is the fraction of liquid assets stock market participants invest in the stock market. We classify households whose head is above 60 years of age as "old" and households whose head is below 40 years of age as "young." Difference in stock holdings, they-axis in graphs (a) and (c), is calculated as the difference between the logs of the fractions of stock holders among the old and among the young age group. Percentage stock, the y-axis in graphs (b) and (d), is the difference in the fraction of liquid assets invested in stock. The red line depicts the linear fit. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

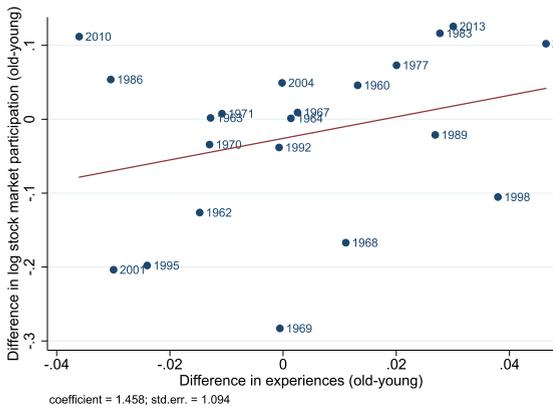
We present the results graphically. We plot the relation between stock holdings (extensive and intensive margin) and experienced returns (Fig. 6), dividends (Fig. 7), earnings (Fig. 8), and GDP (Fig. 9). (Graphs 6(a) and 6(c) update the evidence on the extensive margin and returns presented in Malmendier and Nagel, 2011).

The results for all four performance measures and both for the extensive and intensive margin are in line with the predictions of our model. Starting from experienced returns with $\lambda = 1$ in Panel (a) of Fig. 6, we see that the older age group is more likely to hold stock, compared to the younger age group, when they have experienced

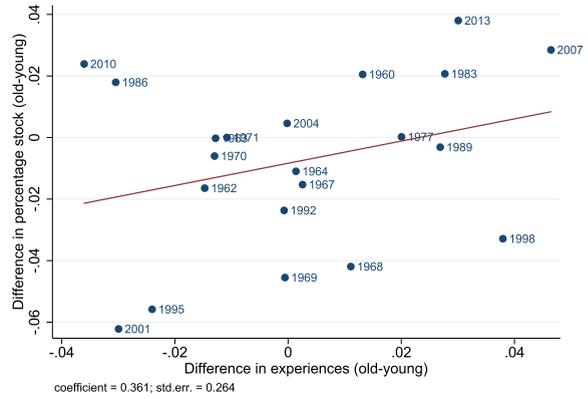
higher stock-market returns in their lives. The opposite holds when the returns experienced by the younger generations are higher than those of the older generations. The slope coefficient of the linear line of fit is significant at 5%. The steepness of the weighting function, and hence the extent of imposed weight on recent data points, makes little difference, as the comparison with graph (b) for $\lambda = 3$ reveals.

The analysis of the intensive margin of stock market investment yields the same conclusion. Both graph (c) and graph (d) indicate that older generations invest a higher share of their liquid assets in stock, compared to the younger generations, when their experienced returns have been higher than those of the younger age group over their respective lifespans so far, and vice versa when they have experienced lower returns than the younger cohorts. Here, the slope coefficient is significant at 10%.

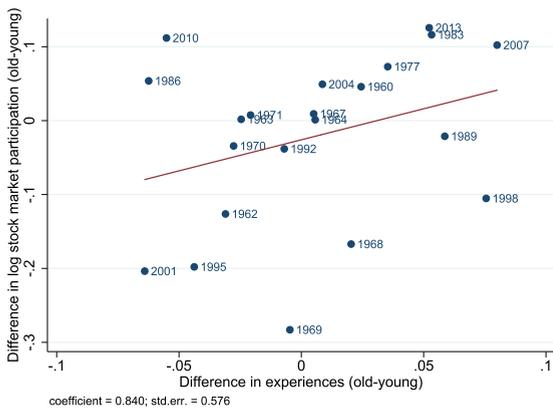
asset allocation but could induce selection bias. By applying SCF sample weights, we undo the overweighting of high-income households and also adjust for non-response bias.



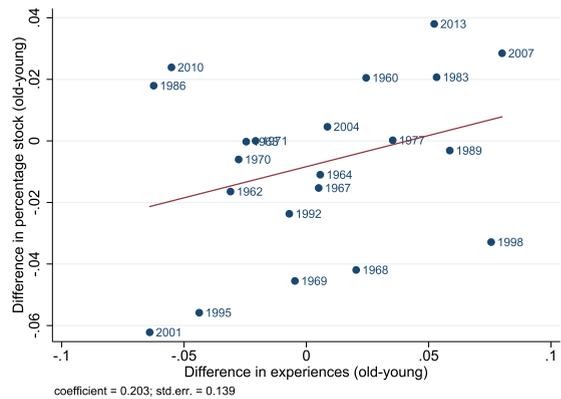
(a) Stock market participation ($\lambda = 1$)



(b) Fraction invested in stock ($\lambda = 1$)



(c) Stock market participation ($\lambda = 3$)



(d) Fraction invested in stock ($\lambda = 3$)

Fig. 7. Experienced dividends and stock holdings Notes. *Difference in experienced dividends* is calculated as the lifetime average experienced real dividends as given on Robert Shiller’s website, using declining weights with either $\lambda = 1$ or $\lambda = 3$ as in Eq. (5). *Stock market participation* is measured as the fraction of households in the respective age groups that hold at least \$1 of stock ownership, either as directly held stock or indirectly (e. g. via mutuals or retirement accounts). *Fraction invested in stock* is the fraction of liquid assets stock market participants invest in the stock market. We classify households whose head is above 60 years of age as “old” and households whose head is below 40 years of age as “young.” Difference in stock holdings, they-axis in graphs (a) and (c), is calculated as the difference between the logs of the fractions of stock holders among the old and among the young age group. Percentage stock, the y-axis in graphs (b) and (d), is the difference in the fraction of liquid assets invested in stock. The red line depicts the linear fit. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Figs. 7 to 9 present the corresponding results for experienced dividends, earnings, and GDP. For all measures, we observe a positive relation of differences in experienced performance and stock investments between the young and the old. The fact that we obtain very similar findings for a wide array of performance measures lends support to the link between our theoretical model and the empirical facts and ameliorates concerns about dividends not translating one-to-one into an empirical performance measure.

6.3.2. Trade volume

We now turn to the second prediction, which relates trade volume to the dispersion of changes in disagreement among investors. We calculate changes in the level of disagreement as the cross-cohort standard deviation of the change in experienced performance between the current

year and the previous year. We weight the cohorts by their sizes when computing the standard deviation.²⁵

As a measure of abnormal trade volume, we calculate the deviation of the turnover ratio from its trend. Following prior literature (Statman et al., 2006; Lo and Wang, 2000), we first compute firm-level turnover ratio (i. e., the number of shares traded over the number of shares outstanding, on a monthly basis). We require that firms be listed on the NYSE or Amex. We exclude Nasdaq-listed firms because the dealer market has volume measurement conventions that differ from exchange-traded securities (Atkins and Dyl, 1997; Statman et al., 2006). Then, we aggregate these numbers into a market-wide turnover

²⁵ For this, we obtain data on US population by age between 1985 and 2015 from the US Census Bureau.

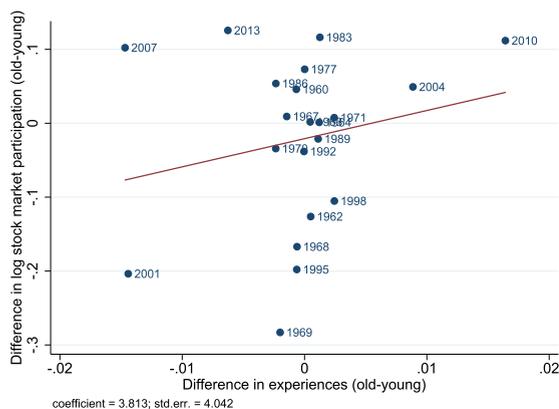
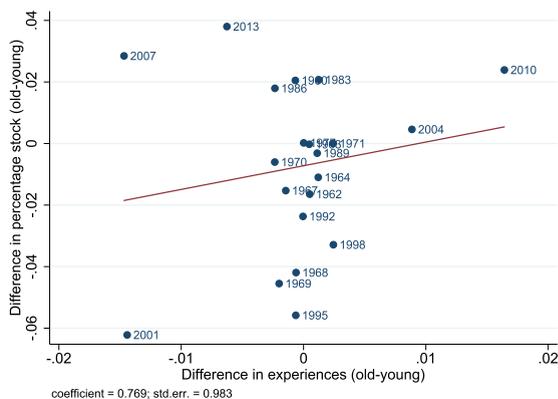
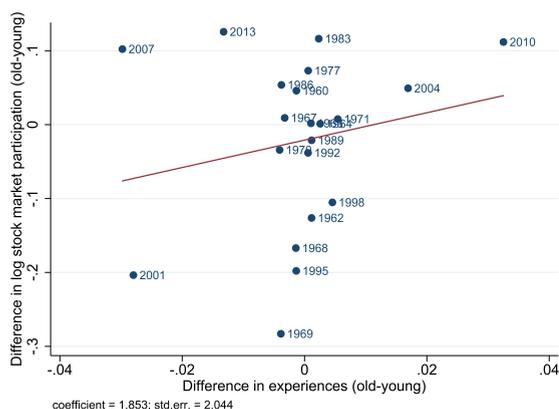
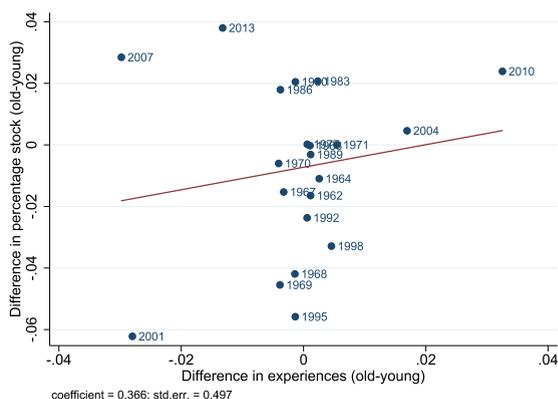
(a) Stock market participation ($\lambda = 1$)(b) Fraction invested in stock ($\lambda = 1$)(c) Stock market participation ($\lambda = 3$)(d) Fraction invested in stock ($\lambda = 3$)

Fig. 8. Experienced earnings and stock holdings Notes. *Difference in experienced earnings* is calculated as the lifetime average experienced log real earnings as given on Robert Shiller's website, using declining weights with either $\lambda = 1$ or $\lambda = 3$ as in Eq. (5). *Stock market participation* is measured as the fraction of households in the respective age groups that hold at least \$1 of stock ownership, either as directly held stock or indirectly (e. g. via mutuals or retirement accounts). *Fraction invested in stock* is the fraction of liquid assets stock market participants invest in the stock market. We classify households whose head is above 60 years of age as "old" and households whose head is below 40 years of age as "young." Difference in stock holdings, the y-axis in graphs (a) and (c), is calculated as the difference between the logs of the fractions of stock holders among the old and among the young age group. Percentage stock, the y-axis in graphs (b) and (d), is the difference in the fraction of liquid assets invested in stock. The red line depicts the linear fit. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

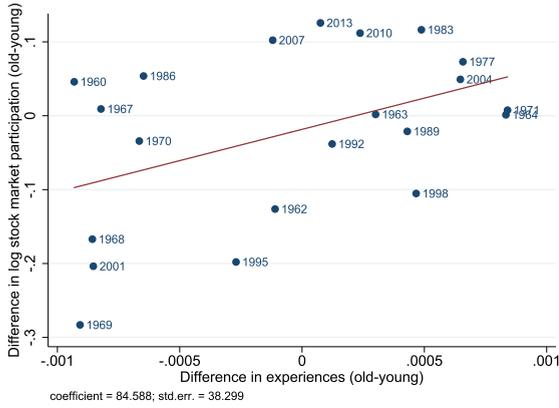
ratio, weighting firms by their market capitalization.²⁶ Since the turnover ratio is nonstationary, we proceed in the same way as above and apply the Christiano and Fitzgerald (2003) to the logarithm of the turnover ratio series so that we keep frequencies between two and eight years. We examine the comovement between the aforementioned measure of disagreement (i. e., the standard deviation of the change in experienced stock returns) and the above measure of (abnormal) trade volume.

Fig. 10 displays the trade volume in dashed, dark (blue) color and changes in the experience-based disagreement about returns between cohorts in light (orange) color over time. Graph (a) shows the results when we apply linear weights for the calculation of experienced returns, and

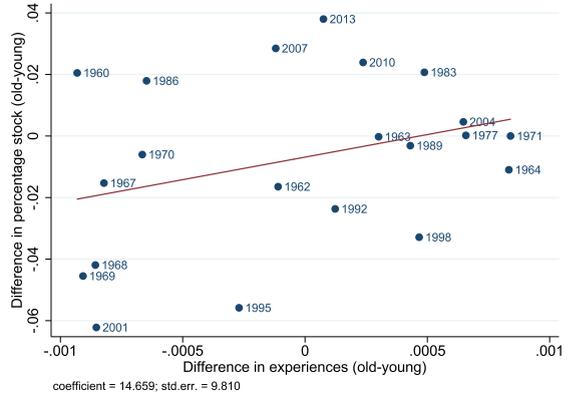
graph (b) displays the case with superlinear weights ($\lambda = 3$). Since we work with annual data for our disagreement variable, we choose the average of the turnover ratio in December of a given year and in January of the following year as our measure for trading volume of the given year. That is, Fig. 10 compares the variation (standard deviation) in changes in experienced returns in a given year to trading volume in December of that year and January of the following year. We choose 1985 as the starting year for this analysis since individual investors were trading substantially less frequently when trading cost were significantly higher up to the mid-1980s, making it less likely that (individual) investors traded repeatedly based on experienced performance.

Consistent with the predictions of our model, we observe a clear comovement between disagreement among cohorts and trading volume. Table 3 reveals that the

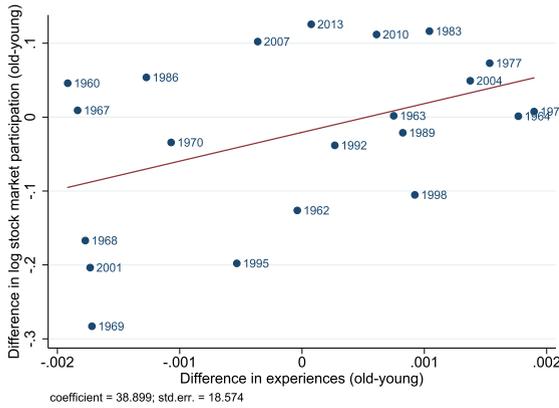
²⁶ This measure is equivalent to dollar turnover ratio (i. e., the ratio of the dollar value of all shares traded and the dollar value of the market).



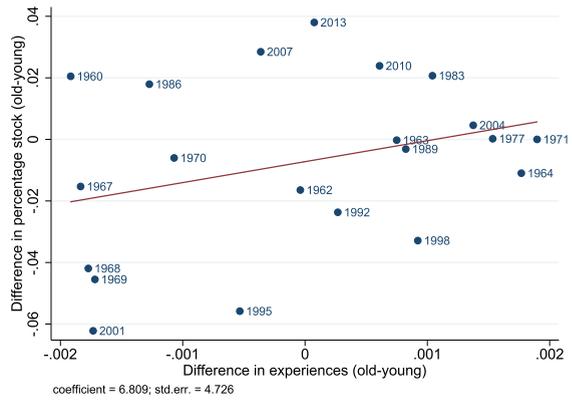
(a) Stock market participation ($\lambda = 1$)



(b) Fraction invested in stock ($\lambda = 1$)



(c) Stock market participation ($\lambda = 3$)



(d) Fraction invested in stock ($\lambda = 3$)

Fig. 9. Experienced Log GDP and Stock Holdings Notes. *Difference in experienced GDP* is calculated as the lifetime average experienced log real GDP, using declining weights with either $\lambda = 1$ or $\lambda = 3$ as in Eq. (5). *Stock market participation* is measured as the fraction of households in the respective age groups that hold at least \$1 of stock ownership, either as directly held stock or indirectly (e. g. via mutuals or retirement accounts). *Fraction invested in stock* is the fraction of liquid assets stock market participants invest in the stock market. We classify households whose head is above 60 years of age as “old” and households whose head is below 40 years of age as “young.” Difference in stock holdings, the y-axis in graphs (a) and (c), is calculated as the difference between the logs of the fractions of stock holders among the old and among the young age group. Percentage stock, the y-axis in graphs (b) and (d), is the difference in the fraction of liquid assets invested in stock. The red line depicts the linear fit. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3

Trading volume and changes in experience-based disagreement.

Notes. The table displays the pairwise correlations (and corresponding *p*-values in parentheses) of trading volume and eight measures of the change in experience-based disagreement. Trading volume is calculated using the market-capitalization weighted average turnover ratio (shares traded divided by shares outstanding) across all firms for January of the current year and December of the preceding year (averaged). We log, linearly detrend, and CF-filter the yearly variable. Experience-based disagreement is calculated separately for returns, dividends, earnings, and GDP, where returns are defined as the inflation-adjusted change in price from the prior year divided by inflation-adjusted price in the prior year and dividends, earnings, and GDP are inflation adjusted. Returns, dividends, log earnings, and log GDP are linearly detrended and CF-filtered, and experience is calculated both with linear weights ($\lambda=1$) and with superlinear weights ($\lambda = 3$). For each measure, we calculate the change in experience for individuals of a given age from the experience of the same individuals in the prior year. We then calculate the current-year age-cohort population-weighted standard deviation of the changes in experiences.

Experiences constructed using:		Returns	Dividends	Log Earnings	Log GDP
$\lambda = 1$	Correlation	0.5976	0.1788	0.3225	0.1780
	(<i>p</i> -value)	(0.0004)	(0.3358)	(0.0768)	(0.3379)
$\lambda = 3$	Correlation	0.4904	0.1489	0.3099	0.1886
	(<i>p</i> -value)	(0.0051)	(0.4240)	(0.0898)	(0.3096)

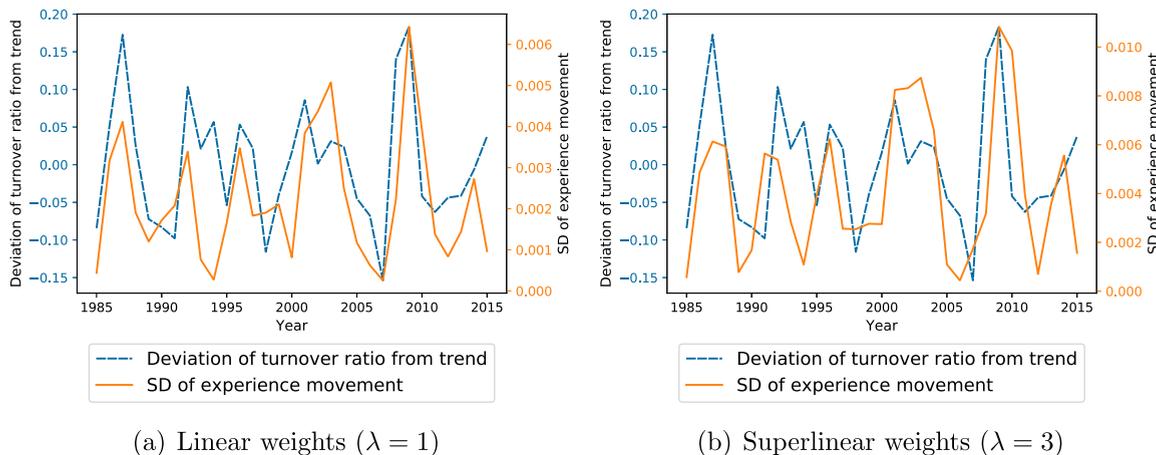


Fig. 10. Trading volume and standard deviation of changes in experienced returns Notes. Trading volume, shown in dashed dark (blue), is calculated as the market-capitalization weighted average monthly turnover ratio (shares traded divided by shares outstanding) across all firms in January and in December of the preceding year. We log, linearly detrend, and CF-filter the yearly variable to obtain the deviation of turnover ratio from the trend. Returns are defined as the inflation-adjusted change in price from the prior year divided by inflation-adjusted price in the prior year. Returns are linearly detrended and CF-filtered. After creating the experience variables for returns, we take the change of the experience variable for individuals of a given age from the experience of those individuals in the prior year. We then calculate the current-year age-cohort population-weighted standard deviation of this difference variable for each year as our measure of experience-based disagreement.

comovement is statistically significant at 1%. The table presents the correlation between trading volume and our measures of changes in return disagreement, as well as the correlations when disagreement is measured using our alternative performance measures (i. e., using again dividends, earnings, or GDP). In each case, the correlation coefficient is again positive, albeit (marginally) significant only for changes in disagreement in experienced earnings.

The relation between trade volume and changes in disagreement in experience-based beliefs about future returns in Table 3 and Fig. 10, as well as the directionally similar correlations with the disagreement about other proxies for returns, corroborate the empirical relevance of our model for a better understanding of investor behavior. The pattern is consistent with EBL and suggests that our novel explanation is worth considering. Moreover, as long as we assume that people trade based on their beliefs, it is unlikely that our channel is spurious. At the same time, other variables might also affect both the change in beliefs and the fluctuations in the trade volume. For example, if fluctuation in trade volume is caused both by variability in change in the beliefs (our model) and by another business cycle macro variable, and both factors are positively correlated, we might still obtain a graph similar to Fig. 10. The claim in this section is not that there are no such factors, nor even that we can attribute most or all of the correlation depicted in Fig. 10 to belief-based learning. Instead, the conclusion is that all empirical findings in this section are consistent with EBL and suggest EBL as a novel and relevant factor that helps explain these empirical regularities jointly.

For a more detailed and careful empirical analysis, it will be useful to analyze long-term individual-level panel data, which allows to link cumulative experiences and new experiences to trading decisions in the corresponding year.

7. Conclusion

In this paper, we propose an OLG equilibrium framework to study the effect of personal experiences on market dynamics and how the demographic composition of an economy can have important implications for the extent to which prices depend on fundamentals. We incorporate the two main empirical features of experience effects, the overweighing of lifetime experiences and recency bias, into the belief formation process of agents. By doing so, we generate what we think are two important channels through which shocks have long-lasting effects on market outcomes. The first is the belief formation process: all agents update their beliefs about the future after experiencing a given shock. The second is the cross-sectional heterogeneity in the population: different experiences generate belief heterogeneity.

We show that EBL not only generates several well-known asset pricing puzzles that have been observed in the data, but it also produces new testable predictions about the relation between demographics, prices, trading behavior, and the cross-section of asset holdings, which are in line with the data.

Appendix A. Proofs

A.1. Proofs for results in Section 2

Proof of Lemma 2.1. Let $\Delta(k) \equiv w(k, \lambda, age) - w(k, \lambda, age')$ for all $k \in \{0, \dots, age\}$. We need to show that $\exists k_0 \in \{0, \dots, age'\}$ such that $\Delta(k) < 0$ for all $k \leq k_0$, and $\Delta(k) \geq 0$ for all $k > k_0$, with the last inequality holding strictly for some k .

For $k > age'$, $\Delta(k) > 0$ since $w(k, \lambda, age') \equiv 0$, and hence $\Delta(k) = w(k, \lambda, age) > 0$, for all $k \in \{age' + 1, \dots, age\}$.

For $k \leq \text{age}'$, we note that $\Delta(k) > 0 \iff Q(k) := \frac{w(k, \lambda, \text{age})}{w(k, \lambda, \text{age}')} > 1$. Hence, it remains to be shown that $\exists k_0 \in \{0, \dots, \text{age}'\}$ such that $Q(k) < 1$ for all $k \leq k_0$, and $Q(k) \geq 1$ for all $k > k_0$. Since the normalizing constants used in the weights $w(k, \lambda, \text{age})$ are independent of k (see the definition in Eq. (5)), we absorb them in a constant $c \in \mathbb{R}^+$ and rewrite

$$Q(k) = c \cdot \frac{(\text{age} + 1 - k)^\lambda}{(\text{age}' + 1 - k)^\lambda} = c \cdot \left[\frac{\text{age} + 1 - k}{\text{age}' + 1 - k} \right]^\lambda$$

$$= c \cdot \alpha(k)^\lambda \quad \forall k \in \{0, \dots, \text{age}'\}. \quad (26)$$

The function $x \mapsto \alpha(x) = \frac{\text{age}+1-x}{\text{age}'+1-x}$ has derivative $\alpha'(x) = \frac{\text{age}-\text{age}'}{(\text{age}'+1-x)^2} > 0$ for $x \in [0, \text{age}' + 1)$, and hence $Q(\cdot)$ is strictly increasing over $\{0, \dots, \text{age}'\}$. Thus, to complete the proof, we only have to show that $Q(k) < 1$ or, equivalently, $\Delta(k) < 0$ for some $k \in \{0, \dots, \text{age}'\}$. We know that $\sum_{k=0}^{\text{age}'} \Delta(k) = 0$ because $\sum_{k=0}^{\text{age}'} w(k, \lambda, \text{age}) = \sum_{k=0}^{\text{age}'} w(k, \lambda, \text{age}') = 1$, and we also know that $\sum_{k=\text{age}'+1}^{\text{age}} \Delta(k) > 0$ since $\Delta(k) = w(k, \lambda, \text{age}) > 0$ for all $k \in \{\text{age}' + 1, \dots, \text{age}\}$. Hence, it must be that $\Delta(k) < 0$ for some $k < \text{age}'$. \square

A.2. Proofs for results in Section 4

Proposition 4.1 directly follows from the following Lemma.

Lemma A.1. Let $z \sim N(\mu, \sigma^2)$, then for any $a > 0$,

$$x^* = \arg \max_x E[-\exp\{-axz\}] = \frac{\mu}{a\sigma^2}$$

and

$$\max_x E[-\exp\{-axz\}] = -\exp\left\{-\frac{1}{2}(\sigma ax^*)^2\right\}$$

$$= -\exp\left(-\frac{1}{2} \frac{\mu^2}{\sigma^2}\right).$$

Proof of Lemma A.1. Since $z \sim N(\mu, \sigma^2)$, we can rewrite the problem as follows:

$$x^* = \arg \max_x -\exp\left(-axE[z] + \frac{1}{2}a^2x^2V[z]\right)$$

$$= \arg \max_x ax\mu - \frac{1}{2}a^2x^2\sigma^2.$$

From FOC, $x^* = \frac{\mu}{a\sigma^2}$. Plugging x^* into $-\exp(-ax^*\mu + \frac{1}{2}a^2(x^*)^2\sigma^2)$ the second result follows. \square

Proof of Proposition 4.2. We show the result for the guess $p_t = \alpha + \beta_0 d_t + \dots + \beta_K d_{t-K}$ with $K = q$. This case shows the logic of the proof; the proof for the case starting from an arbitrary lag $K \geq q$ is analogous but more involved and is omitted for simplicity.

From Lemma A.1, agents' demand for the risky asset is given by $x_t^n = \frac{E_t^n[s_{t+1}]}{\gamma V[s_{t+1}]}$. Plugging in our guess for prices, and for $\beta_0 \neq -1$, we obtain:

$$x_t^n = \frac{(1 + \beta_0)\theta_t^n + \alpha + \beta_1 d_t + \dots + \beta_q d_{t-q+1} - p_t R}{\gamma(1 + \beta_0)^2 \sigma^2}. \quad (27)$$

By market clearing, $\frac{1}{q} \sum_{n=t-q+1}^t x_t^n = 1$, which implies that

$$\frac{(1 + \beta_0) \frac{1}{q} \sum_{n=t-q+1}^t \theta_t^n}{\gamma(1 + \beta_0)^2 \sigma^2} + \frac{\alpha + \beta_1 d_t + \dots + \beta_q d_{t-q+1} - p_t R}{\gamma(1 + \beta_0)^2 \sigma^2} = 1.$$

By straightforward algebra and the definition of θ_t^n , it follows that

$$(1 + \beta_0) \frac{1}{q} \sum_{n=t-q+1}^t \left[\sum_{k=0}^{t-n} w(k, \lambda, t-n) d_{t-k} \right] + [\alpha - \gamma(1 + \beta_0)^2 \sigma^2] + \beta_1 d_t + \dots + \beta_q d_{t-q+1} = p_t R.$$

Plugging in (again) our guess for p_t and using the method of undetermined coefficients, we find the expressions for α and the β 's:

$$-\frac{\gamma(1 + \beta_0)^2 \sigma^2}{R - 1} = \alpha \quad (28)$$

$$(1 + \beta_0) \frac{1}{q} \sum_{n=t-q+1}^{t-k} w(k, \lambda, t-n) + \beta_{k+1} = \beta_k R$$

$$\forall k \in \{0, 1, \dots, q-1\} \quad (29)$$

$$0 = \beta_q R. \quad (30)$$

Let w_k be the average of the weights assigned to dividend d_{t-k} by each generation in the market at time t ; i. e., $w_k = \frac{1}{q} \sum_{n=t-q+1}^t w(k, \lambda, t-n)$. Given that a weight of zero is assigned to dividends that a generation did not observe, i. e., for $k > t-n$, we can rewrite $w_k = \frac{1}{q} \sum_{n=t-q+1}^{t-k} w(k, \lambda, t-n)$. Also using $\beta_q = 0$ from Eq. (45) we obtain

$$(1 + \beta_0)w_k + \beta_{k+1} = \beta_k R \quad \forall k \in \{0, 1, \dots, q-2\} \quad (31)$$

$$(1 + \beta_0)w_{q-1} = \beta_{q-1} R \quad (32)$$

By solving this system of equations, we obtain the expressions in the proposition. In particular, $(1 + \beta_0)(w_{q-2} + w_{q-1}/R) = \beta_{q-2} R$ for $k = q-2$, $(1 + \beta_0)(w_{q-3} + w_{q-2}/R + w_{q-1}/R^2) = \beta_{q-3} R$ for $k = q-3$, and so on. This allow us to express (31) and (32) as

$$(1 + \beta_0) \sum_{j=0}^{k-1} w_{q-(k-j)}/R^j = \beta_{q-k} R \quad \text{for } k = 1, \dots, q. \quad (33)$$

The last expression (33) implies $\beta_0 = \frac{\sum_{j=0}^{q-1} w_j/R^j}{R - \sum_{j=0}^{q-1} w_j/R^j} =$

$$\frac{\sum_{j=0}^{q-1} w_j/R^{j+1}}{1 - \sum_{j=0}^{q-1} w_j/R^{j+1}}$$

(from plugging in $k = q$), which in turn, plugged into (43), allows us to obtain the expression for α from (15) in Proposition 4.2. And expression (33) implies $\beta_k = \frac{\sum_{j=0}^{q-1-k} w_{k+j}/R^{j+1}}{1 - \sum_{j=0}^{q-1} w_j/R^{j+1}}$ (from substituting k with $q-k$

and using the expression for β_0) as expressed in Eq. (16) of the proposition. The latter also subsumes Eq. (32), solved

for β_{q-1} , and the above formula for β_0 , and hence holds for $k = 0, \dots, q - 1$. \square

Proof of Proposition. 4.3. For this proof, we use Eq. (31) and (32). In addition, note that by construction, $w_k < w_{k-1}$ for $\lambda > 0$ since for all generations, $w(k, \lambda, \text{age})$ is decreasing in k and more agents observe the realization of $d_{t-(k-1)}$ than d_{t-k} . Given this, it follows that since $\beta_0 > 0$ then $\beta_{q-1} > 0$ and

$$\beta_{q-1} = \frac{1}{R}(1 + \beta_0)w_{q-1} < \frac{1}{R}[(1 + \beta_0)w_{q-2} + \beta_{q-1}] = \beta_{q-2}. \tag{34}$$

In addition, if $\beta_k < \beta_{k-1}$, then

$$\begin{aligned} \beta_{k-1} &= \frac{1}{R}[(1 + \beta_0)w_{k-1} + \beta_k] \\ &< \frac{1}{R}[(1 + \beta_0)w_{k-2} + \beta_{k-1}] = \beta_{k-2}. \end{aligned} \tag{35}$$

Thus, the proof that $\beta_k < \beta_{k-1}$ for all $k \in \{1, \dots, q - 1\}$ follows by induction. \square

Proof of Lemma. 4.1. To show that β_0 is increasing in λ , let $G_q(\lambda) = \sum_{k=0}^{q-1} w_k/R^{k+1}$. We thus have $\beta_0 = \frac{G_q(\lambda)}{1-G_q(\lambda)}$, and it suffices to show that $G'_q(\lambda) > 0 \forall q > 0$ and $\forall \lambda > 0$. After some algebra, the terms in $G_q(\cdot)$ can be reorganized as follows:

$$G_q(\lambda) = \sum_{\text{age}=0}^{q-1} \frac{1}{q} \sum_{k=0}^{\text{age}} w(k, \lambda, \text{age})/R^{k+1}. \tag{36}$$

Note that for any $\text{age} \in \{0, \dots, q - 1\}$, (i) $\sum_{k=0}^{\text{age}} w(k, \lambda, \text{age}) = 1$, and (ii) for any λ_1, λ_2 such that $\lambda_1 > \lambda_2 > 0$, $\sum_{k=j}^{\text{age}} w(k, \lambda_1, \text{age}) < \sum_{k=j}^{\text{age}} w(k, \lambda_2, \text{age})$. Thus, the weight distribution given by λ_2 first-order stochastically dominates the weight distribution given by λ_1 . Since $1/R > 1/R^2 > 1/R^3 > \dots > 1/R^{q-1}$, stochastic dominance implies that for all $\text{age} \in \{0, \dots, q - 1\}$, $\sum_{k=0}^{\text{age}} c^{k+1}w(k, \lambda_1, \text{age}) > \sum_{k=0}^{\text{age}} c^{k+1}w(k, \lambda_2, \text{age})$, and thus $G_q(\lambda_1) > G_q(\lambda_2)$.

To show the limit results, note that $\lim_{\lambda \rightarrow \infty} w(0, \lambda, \text{age}) = 1$, while $\lim_{\lambda \rightarrow \infty} w(k, \lambda, \text{age}) = 0$ for all $k > 0$. \square

Proof of Proposition 4.4. From Propositions 4.1 and 4.2, we know that, for any t , any generations $m \geq n$ both in $\{t - q + 1, \dots, t\}$ and any $k \in \{0, \dots, q - 1\}$,

$$\begin{aligned} &\xi(n, k, t) \\ &= \frac{\sum_{m=0}^{t-n} (F(m, t - n) - F(m - 1, t - n))d_{t-m} - \sum_{m=0}^{t-n} (F(m, t - n - k) - F(m - 1, t - n - k))d_{t-m}}{\gamma(1 + \beta_0)\sigma^2} \\ &= \frac{\sum_{j=0}^{t-n-1} (d_{t-j} - d_{t-j-1})(F(j, t - n) - F(j, t - n - k))}{\gamma(1 + \beta_0)\sigma^2}. \end{aligned}$$

$$\frac{\partial(x_t^n - x_t^m)}{\partial d_{t-k}} = \frac{(1 + \beta_0)}{\gamma V[s_{t+1}]} \frac{\partial(\theta_t^n - \theta_t^m)}{\partial d_{t-k}}.$$

We note that, for any $n \in \{t - q + 1, \dots, t\}$, $\frac{\partial \theta_t^n}{\partial d_{t-k}} = w(k, \lambda, n - t)$ if $k \in \{0, \dots, t - n\}$, and $\frac{\partial \theta_t^n}{\partial d_{t-k}} = 0$ if $k \in \{t - n + 1, \dots, q - 1\}$. (Observe that $t - n \leq q - 1$.) Hence, it suffices to compare $w(k, \lambda, t - n)$ with $w(k, \lambda, t - m)$ for any $k \in \{0, \dots, q - 1\}$. (As usual, here we adopt the convention that for any age , $w(k, \lambda, \text{age}) = 0$ for all $k \geq \text{age}$.) From Lemma 2.1, there exists a k_0 such that $w(k, \lambda, t - n) < w(k, \lambda, t - m)$ for all $k \in \{0, \dots, k_0\}$ and $w(k, \lambda, t - n) \geq w(k, \lambda, t - m)$ for the rest of the k 's, $k \in \{k_0 + 1, \dots, q - 1\}$. \square

The proof of Proposition 4.5 relies on the following first-order stochastic dominance result:

Lemma A.2. For any $a \in \{0, 1, \dots\}$, $a' < a$ and any $m \in \{0, \dots, a\}$, let $F(m, a) \equiv \sum_{j=0}^m w(j, \lambda, a)$. Suppose the conditions of Lemma 2.1 hold; then $F(m, a) \leq F(m, a')$ for all $m \in \{0, \dots, a\}$.

Proof of Lemma A.2. From Lemma 2.1, we know that there exists a unique j_0 where $w(j_0, \lambda, a') - w(j_0, \lambda, a)$ “crosses” zero. Thus, for $m \leq j_0$, the result is true because $w(j, \lambda, a') > w(j, \lambda, a)$ for all $j \in \{0, \dots, m\}$. For $m > j_0$, the result follows from the fact that $w(j, \lambda, a') < w(j, \lambda, a)$ for all $j \in \{m, \dots, a\}$ and $F(a, a) = F(a', a') = 1$. \square

Proof of Proposition 4.5. We first introduce some notation. For any $j \in \{t - n - k + 1, \dots, t - n\}$, let $w(j, \lambda, t - n - k) = 0$; i. e., we define the weights of generation $n + k$ for time periods before they were born to be zero. Thus, $\sum_{j=0}^{t-n-k} w(j, \lambda, t - n - k)d_{t-j} = \sum_{j=0}^{t-n} w(j, \lambda, t - n - k)d_{t-j}$. In addition, we note that $(w(j, \lambda, t - n - k))_{j=0}^{t-n}$ and $(w(j, \lambda, t - n))_{j=0}^{t-n}$ are sequences of positive weights that add up to one.

Let for any $m \in \{0, \dots, t - n\}$,

$$\begin{aligned} F(m, t - n - k) &= \sum_{j=0}^m w(j, \lambda, t - n - k) \text{ and} \\ F(m, t - n) &= \sum_{j=0}^m w(j, \lambda, t - n). \end{aligned}$$

These quantities, as functions of m , are nondecreasing and $F(t - n, t - n - k) = F(t - n, t - n) = 1$. Moreover, $F(m + 1, t - n - k) - F(m, t - n - k) = w(m + 1, \lambda, t - n - k)$, and $F(m + 1, t - n) - F(m, t - n) = w(m + 1, \lambda, t - n)$. Finally, we set $F(-1, t - n) = F(-1, t - n - k) = 0$.

By these observations, by the definition of $\xi(n, k, t)$, and by straightforward algebra, it follows that,

If the weights are nondecreasing, then $d_{t-j} - d_{t-j-1} \geq 0$ for all $j = 0, \dots, t - n - 1$, and it suffices to show that $F(j, t - n) \leq F(j, t - n - k)$ for all $j = 0, \dots, t - n - 1$. This

follows from applying Lemma A.2 with $a = t - n > t - n - k = a'$.

If the weights are nonincreasing, then $d_{t-j} - d_{t-j-1} \leq 0$, and the sign of $\xi(n, k, t)$ changes accordingly. \square

Proof of Proposition. 4.6. By Propositions 4.1 and 4.2, it follows that for any t and $n \leq t$,

$$x_t^n = \frac{1}{\gamma \sigma^2 (1 + \beta_0)^2} \left(\alpha_0 (1 - R) + (1 + \beta_0) \theta_t^n - R \beta_0 d_t + \sum_{k=1}^{q-1} \beta_k (d_{t+1-k} - R d_{t-k}) \right). \quad (37)$$

Thus, for $n \in \{t - q + 1, \dots, t - 1\}$,

$$x_t^n - x_{t-1}^n = \frac{(1 + \beta_0)(\theta_t^n - \theta_{t-1}^n) + \mathcal{T}(d_{t:t-q})}{\gamma \sigma^2 (1 + \beta_0)^2}, \quad (38)$$

where $\mathcal{T}(d_{t:t-q}) \equiv \sum_{k=1}^{q-1} \beta_k (d_{t+1-k} - d_{t-k} - R(d_{t-k} - d_{t-1-k})) - R \beta_0 (d_t - d_{t-1})$. Note that $\mathcal{T}(d_{t:t-q})$ is not cohort specific (i. e., does not depend on n).

The fact that $x_t^t - x_{t-1}^t = x_t^t$ and $x_t^{t-q} - x_{t-1}^{t-q} = -x_{t-1}^{t-q}$, and market clearing imply

$$q^{-1} \left(\sum_{n=t-q}^t x_t^n - x_{t-1}^n \right) = 0. \quad (39)$$

This expression and the expression in (38) imply that

$$\begin{aligned} & \frac{1}{q} \left(\sum_{n=t-q+1}^{t-1} \frac{(1 + \beta_0)(\theta_t^n - \theta_{t-1}^n)}{\gamma \sigma^2 (1 + \beta_0)^2} + x_t^t - x_{t-1}^{t-q} \right) \\ &= -\frac{1}{q} \sum_{n=t-q}^t \frac{\mathcal{T}(d_{t:t-q})}{\gamma \sigma^2 (1 + \beta_0)^2} = -\frac{\mathcal{T}(d_{t:t-q})}{\gamma \sigma^2 (1 + \beta_0)^2}. \end{aligned}$$

Letting $\theta_{t-1}^t = \theta_t^{t-q} = 0$, it follows that

$$\frac{1}{q} \left(\sum_{n=t-q}^t (1 + \beta_0)(\theta_t^n - \theta_{t-1}^n) \right) = -\mathcal{T}(d_{t:t-q}).$$

Thus, we can express the change in individual demands for those agents with $n = \{t - q + 1, \dots, t - 1\}$ in expression (38) as follows:

$$x_t^n - x_{t-1}^n = \chi \left[(\theta_t^n - \theta_{t-1}^n) - \frac{1}{q} \sum_{n=t-q}^t (\theta_t^n - \theta_{t-1}^n) \right], \quad \forall n \in \{t, \dots, t - q\}, \quad (40)$$

where $\chi \equiv \frac{1}{\gamma \sigma^2 (1 + \beta_0)}$. By squaring and summing at both sides and including the demands on the youngest ($n = t$) and oldest ($n = t - q$) market participants, the desired result follows. \square

Appendix B. Incorporating prior beliefs

In this section, we show how the model can be extended to allow agents to have prior beliefs; that is, $\tau > 0$ in (4). We will prove results analogous to those in Proposition 4.2.

As a reminder, we now suppose that all cohorts are born with prior belief $N(m, \sigma_m^2)$ and update their beliefs during their lifetime as follows:

$$\theta_t^n = (1 - \omega_{t-n})m + \omega_{t-n} \left[\sum_{k=0}^{t-n} w(k, \lambda, t - n) d_{t-k} \right], \quad (41)$$

where ω_{t-n} is given by

$$\omega_{t-n} = \frac{t - n + 1}{\tau + (t - n + 1)},$$

and where τ captures the relative importance of prior beliefs to experience-based beliefs. For example, if agents are Bayesian from experience as described in Section 2.1, then $\tau = \frac{\sigma^2}{\sigma_m^2}$. For the purpose of our analysis, however, all that is important is how results vary with τ .

We continue to guess that prices are affine in past dividends,

$$p_t = \alpha + \beta_0 d_t + \dots + \beta_q d_{t-q}$$

with $K = q$, as in the baseline model. From Lemma A.1, agents' demand for the risky asset is given by $x_t^n = \frac{E_t^n [s_{t+1}]}{\gamma V [s_{t+1}]}$. Plugging in our guess for prices, and for $\beta_0 \neq -1$, we obtain

$$x_t^n = \frac{(1 + \beta_0)\theta_t^n + \alpha + \beta_1 d_t + \dots + \beta_q d_{t-q+1} - p_t R}{\gamma (1 + \beta_0)^2 \sigma^2}. \quad (42)$$

By market clearing, $\frac{1}{q} \sum_{n=t-q+1}^t x_t^n = 1$, which implies that

$$\begin{aligned} & \frac{(1 + \beta_0) \frac{1}{q} \sum_{n=t-q+1}^t \theta_t^n}{\gamma (1 + \beta_0)^2 \sigma^2} \\ &+ \frac{\alpha + \beta_1 d_t + \dots + \beta_q d_{t-q+1} - p_t R}{\gamma (1 + \beta_0)^2 \sigma^2} = 1. \end{aligned}$$

By straightforward algebra and the definition of θ_t^n , it follows that

$$\begin{aligned} & \left(\frac{1}{\gamma (1 + \beta_0)^2 \sigma^2} \right) \left[(1 + \beta_0) \frac{1}{q} \sum_{n=t-q+1}^t \theta_t^n + \alpha + \beta_1 d_t \right. \\ & \quad \left. + \dots + \beta_q d_{t-q+1} - p_t R \right] = 1 \\ & (1 + \beta_0) \frac{1}{q} \sum_{n=t-q+1}^t \\ & \quad \times \left[(1 - \omega_{t-n})m + \omega_{t-n} \sum_{k=0}^{t-n} w(k, \lambda, t - n) d_{t-k} \right] \\ & \quad + \alpha - \gamma (1 + \beta_0)^2 \sigma^2 \dots \\ & \quad + \beta_1 d_t + \dots + \beta_q d_{t-q+1} = p_t R. \\ & \left[(1 + \beta_0) \frac{1}{q} \sum_{n=t-q+1}^t (1 - \omega_{t-n})m + \alpha - \gamma (1 + \beta_0)^2 \sigma^2 \right] \\ & \quad + \dots (1 + \beta_0) \frac{1}{q} \sum_{n=t-q+1}^t \sum_{k=0}^{t-n} \omega_{t-n} w(k, \lambda, t - n) d_{t-k} \\ & \quad + \beta_1 d_t + \dots + \beta_q d_{t-q+1} = p_t R. \end{aligned}$$

Plugging in (again) our guess for p_t and using the method of undetermined coefficients, we find the expressions for α and the β 's:

$$\frac{\gamma(1+\beta_0)^2\sigma^2 + (1+\beta_0)\frac{1}{q}\sum_{n=t-q+1}^t(1-\omega_{t-n})m}{1-R} = \alpha \quad (43)$$

$$(1+\beta_0)\frac{1}{q}\sum_{n=t-q+1}^{t-k}\omega_{t-n}w(k,\lambda,t-n) + \beta_{k+1} = \beta_k R$$

$$\forall k \in \{0, 1, \dots, q-1\} \quad (44)$$

$$0 = \beta_q R, \quad (45)$$

where w_k is now the average of the weights assigned to dividend d_{t-k} by each generation in the market at time t ; i. e., $w_k = \frac{1}{q}\sum_{n=t-q+1}^t\omega_{t-n}w(k,\lambda,t-n)$.

Introducing prior beliefs requires two adjustments. First, the constant in prices, α , now increases to incorporate the demand driven by prior belief, m . Second, all the weights that an agent with age $t-n$ gives to past dividends are now adjusted by $\omega(t-n)$, which keeps track of the importance that these agents assign to their EBL. Such adjustment affects the β 's in the pricing equation. Given these adjustments, the model is isomorphic to the baseline model.

Appendix C. Population growth

In addition to considering the effects of a one-time shock to population structure, we also explore the implications of population growth.

In this section of the Appendix, we consider an OLG model two-period lived agents where the mass of young agents born every period grows at rate g . For this growth setting, we need to set an initial date for the economy, which we define to be $t=0$. Let y_t denote the mass of young agents born at time t ; then $y_{t+1} = (1+g)y_t = y_0(1+g)^t$. We further denote the total mass of people at any point in time $t > 0$ as n_t , and hence $n_t = y_t + y_{t-1} = (2+g)y_{t-1}$. It is easy to check that $n_t = (1+g)n_{t-1}$; that is, total population grows at rate g .

The framework is otherwise as in the “toy model” in Section 3 of the main paper. The main difference is that now population is growing over time. As a result, we make a different guess for the price function:

$$p_t = \alpha_0(1+g)^{-t} + \beta_0 d_t + \beta_1 d_{t-1}.$$

We verify this guess using our market clearing condition, which requires the demand of the young and the old to add up to total supply of the asset, one:

$$1 = y_t \frac{E_t^i[p_{t+1} + d_{t+1}] - R p_t}{\gamma V[p_{t+1} + d_{t+1}]} + y_{t-1} \frac{E_t^{t-1}[p_{t+1} + d_{t+1}] - R p_t}{\gamma V[p_{t+1} + d_{t+1}]} \iff$$

$$1 = \frac{y_0(1+g)^{t-1}}{\gamma(1+\beta_0)^2\sigma^2} [(1+\beta_0)[(1+g)E_t^i[d_{t+1}] + E_t^{t-1}[d_{t+1}]] + (2+g)[\alpha_0(1+g)^{-(t+1)} + \beta_1 d_t - R p_t].$$

and after simple algebra,

$$R p_t = (1+\beta_0) \left\{ \frac{1+g}{2+g} d_t + \frac{1}{2+g} [(1-\omega)d_{t-1} + \omega d_t] \right\} + \frac{\alpha_0}{(1+g)^{t+1}} + \beta_1 d_t - \frac{\gamma\sigma^2(1+\beta_0)^2}{y_0(2+g)(1+g)^{t-1}}.$$

We plug in $p_t = \alpha_0(1+g)^{-t} + \beta_0 d_t + \beta_1 d_{t-1}$, and we use the method of undetermined coefficients to obtain:

$$\alpha_0 = -\frac{\gamma(1+\beta_0)^2\sigma^2}{R - \frac{1}{1+g}} \frac{(1+g)}{y_0(2+g)}$$

$$R\beta_0 = (1+\beta_0) \left(\frac{1+g}{2+g} + \frac{1}{2+g}\omega \right) + \beta_1$$

$$R\beta_1 = (1+\beta_0) \frac{1-\omega}{2+g}.$$

Let $\alpha_t \equiv \alpha_0(1+g)^{-t}$ and $\gamma \equiv \frac{y_t}{n_t}$ denote the fraction of young agents, which is easy to verify is constant over time. Then, we can rewrite the above equations as

$$\alpha_t = -\frac{\gamma(1+\beta_0)^2\sigma^2}{R - \frac{1}{1+g}} \frac{1+g}{n_t}$$

$$R\beta_0 = (1+\beta_0)(\gamma + (1-\gamma)\omega) + \beta_1$$

$$R\beta_1 = (1+\beta_0)(1-\gamma)(1-\omega).$$

The latter expressions reveal that the total mass of agents in the market is reflected only in the price constant, while the fraction of young people in the market determines the dividend loadings β_0 and β_1 . Overall, we see that adding population growth generates to our model a positive trend in prices. The relative reliance of prices on the most recent experiences (dividends) is increasing in the population growth rate.

Appendix D. Empirical analysis

We use two alternative approaches to measure the fraction of younger agents (below 50 years of age) in the market. First, we compute an indicator variable that equals one when the fraction of young agents in the market is above 0.5 and zero otherwise, $\mathbb{I}\{\text{fraction of young investors}_t > 0.5\}$. Here, the fraction of young investors is based on their relative cohort sizes, with

$$\text{fraction of young investors}_t = \frac{\sum_j \mathbb{I}(\text{age}_{j,t} < 50) \cdot w_{j,t}^{\text{scf}}}{\sum_{j=1} w_{j,t}^{\text{scf}}},$$

where $\text{age}_{j,t}$ is the age of household head j in year t , and $w_{j,t}^{\text{scf}}$ is the weight given to household head j in year t in the SCF to compensate for unequal probabilities of household selection in the original design and for unit nonresponse (failure to obtain an interview).

Our second proxy captures the wealth of younger generations of investors. We construct an indicator variable that equals one when the fraction of liquid wealth owned by agents below 50 is above the 1960–2013 sample average of their liquid wealth and zero otherwise, $\mathbb{I}(\text{fraction of young investors' wealth}_t >$

Table D.1

Markov-switching regime (MSR) model.

Robustness checks of the estimation results in Table 2. $p_t - d_t$ is the log of the price-to-dividend ratio, and regressed on its lagged values interacted with the demographic indicator variable Y_t for the fraction of young investors. We use different thresholds to construct Y_t . In column (1), Y_t equals one when the fraction of investors below 50 is larger than 0.55, and in column (2), the threshold is 0.60. In column (3), Y_t equals one when the fraction of wealth of investors below 50 is larger than 90% of their 1960–2013 sample average, and in column (4), the threshold is 110% of the sample average. As in Table 2, the demographic data including age and wealth (liquid assets) of stock market participants are from the SCF and the stock data are from Robert Shiller’s website. * and ** denote significance at the 10% and 5% levels, respectively.

	Dependent variable: $p_t - d_t$			
	Y_t age-based		Y_t age/wealth-based	
	(1)	(2)	(3)	(4)
δ_1	0.430 (0.288)	1.134** (0.188)	0.681** (0.209)	-0.339 (0.268)
δ_2	0.016 (0.304)	-0.194 (0.322)	-0.082 (0.210)	-0.137 (0.422)
δ_3	-0.460** (0.225)	-0.778** (0.282)	-0.629** (0.161)	0.428* (0.220)
β_1	0.623** (0.259)	0.297** (0.134)	0.417** (0.177)	1.084** (0.081)
β_2	-0.099 (0.219)	-0.132 (0.140)	-0.168 (0.166)	-0.223** (0.100)
β_3	0.298** (0.151)	0.452** (0.106)	0.580** (0.140)	-0.044 (0.064)
$\mu(S_1)$	4.925** (1.541)	5.780** (2.915)	5.734** (1.513)	5.430** (1.683)
$\mu(S_2)$	19.590** (3.399)	14.100** (2.959)	20.310** (3.144)	19.670** (3.102)
σ	3.948 (0.392)	3.259 (0.393)	3.792 (0.397)	3.737 (0.375)
Q_{11}	0.953 (0.030)	0.779 (0.152)	0.956 (0.026)	0.931 (0.035)
Q_{21}	0.374 (0.214)	0.115 (0.068)	0.365 (0.210)	0.473 (0.220)
N	51	51	51	51

sample average)²⁷; i. e.,

fraction of young investors’ wealth,

$$= \frac{\sum_j \mathbb{I}(\text{age}_{j,t} < 50) \cdot w_{j,t}^{scf} \cdot \text{wealth}_{j,t}}{\sum_j w_{j,t}^{scf} \cdot \text{wealth}_{j,t}}$$

For robustness, we also consider thresholds 0.55 and 0.60 for the age-based first proxy and $0.9 \times$ Sample average and $1.1 \times$ Sample average for the age- and wealth-based second proxy. Results are presented in Appendix Table D.1. We estimate a positive δ_1 coefficient, which is significant when requiring a fraction of 0.6 for the age-based coefficient and when requiring wealth above 0.9 of the sample average for the age- and wealth-based coefficient.

²⁷ The SCF shows the age and wealth (liquid assets) information of each respondent in 1960, 1962, 1963, 1964, 1967, 1968, 1969, 1970, 1971, 1977, 1983, 1986, 1989, 1992, 1995, 1998, 2001, 2004, 2007, 2010, and 2013. We use linear interpolation to fill the missing years and construct a yearly sample from 1960 to 2013. The liquid assets variable is defined to be the sum of assets in an investor’s checking, savings, and money market accounts, as well as any call accounts at brokerages and prepaid cards.

References

Adam, K., Marcet, A., Nicolini, J.P., 2016. Stock market volatility and learning. *Journal of Finance* 71 (1), 33–82.

Alesina, A., Fuchs-Schündeln, N., 2007. Good-bye Lenin (or not?): the effect of communism on people’s preferences. *American Economic Review* 97, 1507–1528.

Atkins, A.B., Dyl, E.A., 1997. Market structure and reported trading volume: Nasdaq versus the NYSE. *Journal of Financial Research* 20 (3), 291–304.

Barberis, N., Greenwood, R., Jin, L., Shleifer, A., 2015. X-CAPM: an extrapolative capital asset pricing model. *Journal of Financial Economics* 115 (1), 1–24.

Barberis, N., Greenwood, R., Jin, L., Shleifer, A., 2018. Extrapolation and bubbles. *Journal of Financial Economics* 129 (2), 203–227.

Barsky, R., DeLong, J.B., 1993. Why does the stock market fluctuate? *Quarterly Journal of Economics* 108 (2), 291–311.

Blanchard, O., 2012. Sustaining a global recovery. *Finance and Development* 46 (3), 9–12.

Brunnermeier, M.K., 2001. *Asset Pricing under Asymmetric Information: Bubbles, Crashes, Technical Analysis, and Herding*. Oxford University Press, Oxford.

Campbell, J.Y., Kyle, A.S., 1993. Smart money, noise trading, and stock price behavior. *Review of Economic Studies* 60, 1–34.

Campbell, J.Y., Shiller, R.J., 1988. Stock prices, earnings, and expected dividends. *Journal of Finance* 43 (3), 661–676.

Cassella, S., Gulen, H., 2018. Extrapolation bias and the predictability of stock returns by price-scaled variables. *Review of Financial Studies* 31 (11), 4345–4397.

Cecchetti, S.G., Lam, P., Mark, N.C., 2000. Asset pricing with distorted beliefs: Are equity returns too good to be true? *American Economic Review* 90 (4), 787–805.

Chiang, Y.-M., Hirshleifer, D., Qian, Y., Sherman, A.E., 2011. Do investors learn from experience? Evidence from frequent IPO investors. *Review of Financial Studies* 24, 1560–1589.

Christiano, L.J., Fitzgerald, T.J., 2003. The band pass filter. *International Economic Review* 44 (2), 435–465.

Cochrane, J.H., 2011. Presidential address: discount rates. *Journal of Finance* 66 (4), 1047–1108.

Cogley, T., Sargent, T.J., 2008. The market price of risk and the equity premium: a legacy of the great depression? *Journal of Monetary Economics* 55 (3), 454–476.

Collin-Dufresne, P., Johannes, M., Lochstoer, L.A., 2017. Asset pricing when this time is different. *Review of Financial Studies* 30 (2), 505–535.

Ehling, P., Graniero, A., Heyerdahl-Larsen, C., 2018. Asset prices and portfolio choice with learning from experience. *Review of Economic Studies* 85 (1), 1752–1780.

Fama, E.F., French, K.R., 1988. Dividend yields and expected stock returns. *Journal of Financial Economics* 22 (1), 3–25.

Fuster, A., Hebert, B., Laibson, D., 2012. Natural expectations, macroeconomic dynamics and asset pricing. In: Acemoglu, D., Woodford, M. (Eds.), *NBER Macroeconomics Annual 2011*. University of Chicago Press, Chicago, pp. 1–48.

Fuster, A., Laibson, D., Mendel, B., 2010. Natural expectations and macroeconomic fluctuations. *Journal of Economic Perspectives* 24, 67–84.

Giuliano, P., Spilimbergo, A., 2013. Growing up in a recession. *Review of Economic Studies* 81 (2), 787–817.

Greenwood, R., Shleifer, A., 2014. Expectations of returns and expected returns. *Review of Financial Studies* 27, 714–746.

Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57, 357–384.

Hamilton, J.D., 1994. Chapter 50 state-space models. In: Engle, R.F., McFadden, D.L. (Eds.), *Handbook of Econometrics*, 4th Edition. Elsevier, Amsterdam, pp. 3039–3080.

Hertwig, R., Barron, G., Weber, E.U., Erev, I., 2004. Decisions from experience and the effect of rare events in risky choice. *Psychological Science* 15, 534–539.

Kaustia, M., Knüpfer, S., 2008. Do investors overweight personal experience? Evidence from IPO subscriptions. *Journal of Finance* 63, 2679–2702.

Laudenbach, C., Malmendier, U., Niessen-Ruenzi, A., 2019a. Emotional tagging and belief formation: the long-lasting effects of experiencing communism. In *AEA Papers and Proceedings* 109, 567–571. (Accessed 29 November 2019).

Jin, L.J., 2015. A speculative asset pricing model of financial instability Available at SSRN 2524762.

Laudenbach, C., Malmendier, U., Niessen-Ruenzi, A., 2019b. The long-lasting effects of experiencing communism on financial risk-taking.

- LeRoy, S.F., 2006. Excess volatility. In *The New Palgrave Dictionary of Economics*. Palgrave Macmillan.
- LeRoy, S.F., Porter, R.D., 1981. The present value relation: tests based on implied variance bounds. *Econometrica* 49, 555–574.
- Lo, A.W., Wang, J., 2000. Trading volume: definitions, data analysis, and implications of portfolio theory. *Review of Financial Studies* 13 (2), 257–300.
- Malmendier, U., Nagel, S., 2011. Depression babies: do macroeconomic experiences affect risk-taking? *Quarterly Journal of Economics* 126, 373–416.
- Malmendier, U., Nagel, S., 2016. Learning from inflation experiences. *Quarterly Journal of Economics* 131, 53–87.
- Malmendier, U., Shen, L.S., 2018. Scarred consumption (No. w24696). National Bureau of Economic Research.
- Oreopoulos, P., von Wachter, T., Heisz, A., 2012. The short- and long-term career effects of graduating in a recession: hysteresis and heterogeneity in the market for college graduates. *American Economic Journal: Applied Economics* (4) 1–29.
- Schraeder, S., 2015. Information processing and non-bayesian learning in financial markets. *Review of Finance* 4, 1–31.
- Shiller, R.J., 1981. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review* 71, 421–436.
- Simonsohn, U., Karlsson, N., Loewenstein, G., Ariely, D., 2008. The tree of experience in the forest of information: overweighing experienced relative to observed information. *Games and Economic Behavior* 62 (1), 263–286.
- Statman, M., Thorley, S., Vorkink, K., 2006. Investor overconfidence and trading volume. *Review of Financial Studies* 19 (4), 1531–1565.
- Timmermann, A.G., 1993. How learning in financial markets generates excess volatility and predictability in stock prices. *Quarterly Journal of Economics* 108 (4), 1135–1145.
- Timmermann, A.G., 1996. Excess volatility and predictability of stock prices in autoregressive dividend models with learning. *Review of Economic Studies* 63 (4), 523–557.
- Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: heuristics and biases. *Science* 185, 1124–1131.
- Vives, X., 2010. *Information and Learning in Markets: the Impact of Market Microstructure*. Princeton University Press, Princeton.
- Weber, E.U., Böckenholt, U., Hilton, D.J., Wallace, B., 1993. Determinants of diagnostic hypothesis generation: effects of information, base rates, and experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19 (5), 1151–1164.
- Woodford, M., 2013. Macroeconomic analysis without the rational expectations hypothesis. *Annual Review of Economics* 5 (1), 303–346.