

NBER WORKING PAPER SERIES

INVESTOR EXPERIENCES AND FINANCIAL MARKET DYNAMICS

Ulrike Malmendier
Demian Pouzo
Victoria Vanasco

Working Paper 24697
<http://www.nber.org/papers/w24697>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2018

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), SFB TR 15 (Tutzing, Germany) conferences for helpful comments. We also thank Felix Chopra, Marius Guenzel, Canyao Liu, Leslie Shen, and Jonas Sobott for excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Ulrike Malmendier, Demian Pouzo, and Victoria Vanasco. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Investor Experiences and Financial Market Dynamics
Ulrike Malmendier, Demian Pouzo, and Victoria Vanasco
NBER Working Paper No. 24697
June 2018
JEL No. E03,G02,G11,G12

ABSTRACT

How do macro-financial shocks affect investor behavior and market dynamics? Recent evidence suggests long-lasting effects 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 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 recent financial crises, which we show to be in line with the data.

Ulrike Malmendier
Department of Economics
549 Evans Hall # 3880
University of California, Berkeley
Berkeley, CA 94720-3880
and NBER
ulrike@econ.berkeley.edu

Victoria Vanasco
Stanford University
Department of Economics
vvanasco@stanford.edu

Demian Pouzo
University of California at Berkeley
530-1 Evans Hall #3880
Berkeley, CA 94720
dpouzo@econ.berkeley.edu

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 (2005)), 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, models of natural expectation formation (Fuster, Hebert, and Laibson (2011); Fuster, Laibson, and Mendel (2010)) and over-extrapolation (Barberis, Greenwood, Jin, and Shleifer (2015); Barberis, Greenwood, Jin, and Shleifer (2016)) successfully capture a wide range of the stylized facts. A core feature of these models is that agents over-weigh recent realizations of the relevant economic variables when forming beliefs.

Another set of emerging stylized facts, which focuses on the long-lasting effects of macro-financial shocks and their systematic cross-sectional differences, has been harder to capture by these approaches. As conveyed by the notion of “depression babies” or the “deep scars” of the 2008 financial crisis (Blanchard (2012)), macro-economic shocks appear to alter investment and consumption behavior for decades to come, beyond the time frame of existing models, and there is significant cross-sectional heterogeneity. Younger cohorts tend to react significantly more strongly than older cohorts. A growing empirical literature on *experience effects* documents, for example, that personal lifetime experiences in the stock-market predict future willingness to invest in the stock market (Malmendier and Nagel (2011)), and the same for IPO experiences and future IPO investment (Kaustia and Knüpfer (2008); Chiang, Hirshleifer, Qian, and Sherman (2011)). There is also evidence of experience effects in non-finance settings, e. g., on the long-term

effects of graduating in a recession on labor market outcomes (Oreopoulos, von Wachter, and Heisz (2012)) or of living in (communist) Eastern Germany to political attitudes post-reunification (Alesina and Fuchs-Schundeln (2007)).¹ In all of these applications, researchers identify a long-lasting impact of crisis experiences on individual risk-taking and illustrate their cohort-specific impact.

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), or about the outlook for durable consumption (also in the MSC).² A key difference relative to over-extrapolation and related approaches is that experience-based learning generates cohort-specific differences in beliefs and in their updating after a common shock. While more evidence on the exact process of household-level learning is needed (see the discussions in Campbell (2008) and Agarwal, Driscoll, Gabaix, and Laibson (2013)), the over-weighting of personal experiences appears to be a pervasive and robust psychological phenomenon affecting belief formation, which is 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

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 will the market 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

¹ See also Giuliano and Spilimbergo (2013), who relate the effects of growing up in a recession to redistribution preferences.

² Cf. Malmendier and Nagel (2011), Malmendier and Nagel (2016), Malmendier and Shen (2017).

³ See, for example, Weber, Böckenholt, Hilton, and Wallace (1993), Hertwig, Barron, Weber, and Erev (2004), and Simonsohn, Karlsson, Loewenstein, and Ariely (2008).

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, including the stronger reactions of the young than the old to a given macro shock. 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 CARA preferences and live for a finite number of periods.⁴ During their lifetimes, they choose portfolios of a risky and a risk-free security. We initially 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 over-weigh 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

⁴ The use of CARA preferences with Normal shocks allows us to keep our theoretical analysis tractable, and is widely used in finance for this reason (see [Vives \(2010\)](#)).

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 Section 7, we remove the myopia assumption and show that the first-order effects of experience-based learning are identical to those derived for myopic agents.

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 persistent 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 towards 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 auto-correlation 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 in the market increases, the relative reliance of prices on more recent dividends increases. This is in line with evidence in [Cassella and Gulen \(2015\)](#) who find that the level of extrapolation in markets

is positively related to the fraction of young traders in that market.

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, Greenwood, Jin, and Shleifer \(2015\)](#) 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 of future price changes (and dividend-price ratio changes) 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 \(2005\)](#), 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 de-trended 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.

As the final step in our analysis, we argue that our qualitative results are still present when we remove the myopic-agents assumption. We consider a version of our model where agents re-balance their portfolios every period to maximize their final-period consumption.⁶ The dynamic set-up allows us to analyze how hedging concerns and lifetime-horizon effects interact with experience-based learning. Prior literature has shown that, in a rational-expectations linear equilibrium, the agents’ multi-period investment problem can be partitioned into a sequence of one-period ones (Vives (2010)). Under experience-based learning, such partitioning is no longer possible. However, exploiting the CARA-Gaussian setup, we are able to show that the demand of experience-based learners coincides with the one in a different static problem where dividends are drawn from a *modified* Gaussian distribution. That is, we can still partition the multi-period investment problem into a sequence of one-period problems, albeit with a probability distribution of dividends that differs from the original one. This latter result might also be of interest as an independent technical contribution in solving belief dependencies beyond our specific model proposed.

Related Literature. 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 non-negligible effect on beliefs and generate cross-cohort heterogeneity.

Our paper also relates to the work on extrapolation by Barberis, Greenwood, Jin, and Shleifer (2015) and Barberis, Greenwood, Jin, and Shleifer (2016). They consider a consumption-based asset pricing model with both “rational” and “extrapolative” agents.

⁶ This approach to modeling dynamic portfolio choices is again following a widely used approach in the literature, see Vives (2010).

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 \(1993\)](#), [Timmermann \(1996\)](#), [Adam, Marcet, and Nicolini \(2012\)](#) study the implications of learning and [Cecchetti, Lam, and Mark \(2000\)](#) and [Jin \(2015\)](#) 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 counter-parties hold, in our model of experienced-based learning, information is available to all agents at all times.

Finally, there are contemporaneous papers on learning in OLG models ([Collin-Dufresne, Johannes, and Lochstoer \(2016\)](#), [Schraeder \(2015\)](#)). Most closely related, [Ehling, Graniero, and Heyerdahl-Larsen \(2015\)](#) explore the role of experience in portfolio decisions and asset prices in a complete-markets setting. They focus on trend chasing and the negative relation between beliefs about expected returns and realized future returns, as shown by [Greenwood and Shleifer \(2014\)](#). In our incomplete-markets setting, instead, we study the cross-section of asset holdings and the relation between demographics, pricing, and trading dynamics. Our model further incorporates recency bias in agent’s beliefs formation.

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 first presents 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. Section 5 extends the model to study demographic shocks. Section 6 present empirical implications. We remove the myopia assumption in Section 7. Section 8 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 .) Figure 1 illustrates the time line 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

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

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. In Section 7, we show that the same determinants are at work when this assumption is removed.

For a given initial wealth level W_n^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

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

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.

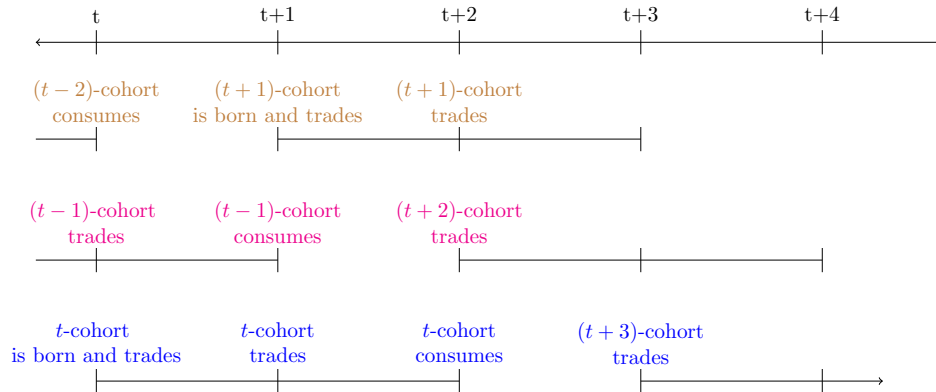


Figure 1: A time line for an economy with two-period lived generations, $q = 2$.

2.1 Experience-Based Learning

In this framework, experienced-based learning (EBL) means that agents over-weigh realization observed during their lifetimes when forecasting dividends, and that they tilt the excess weights towards 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.

We construct the subjective mean of dividends of generation n at time t following the

⁷ We only need agents to discount pre-lifetime 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.

empirical evidence Malmendier and Nagel (2011):

$$\theta_t^n \equiv \sum_{k=0}^{age} w(k, \lambda, age) d_{t-k}, \quad (4)$$

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

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

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 (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. Here 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 increases, 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 non-negative) 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.

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 $N(m, \tau^2)$. For simplicity, all generations have the same prior, though the analysis can easily be extended to heterogeneous Gaussian priors across generations.⁹

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

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

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

⁹ The assumption of Gaussianity is also not needed but simplifies the exposition greatly.

generations alive in any given period have the same belief about the mean of dividends. In addition, beliefs of FBL agents are non-stationary, 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.

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

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

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. Note that this average coincides with the belief θ_t^n of experience-based learners when $\lambda = 0$. In this case, the posterior mean of BLE agents differs from the subjective mean of EBL agents only due to the presence of the prior. As a result, if the prior of BLE agents is diffuse, i. e., $\tau \rightarrow \infty$, then $\tilde{\theta}_{n+a}^n$ coincides with the θ_{n+a}^n of EBL agents with $\lambda = 0$. The same is true as $\text{age} \rightarrow \infty$.

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*

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

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

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

The solution to this problem is standard and given by

$$x_t^n = \frac{E[s_{t+1}]}{\gamma V[s_{t+1}]} \quad (8)$$

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 Illustration: A 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, and the non-myopic case in Section 7.

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 (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 equation (6). Anticipating this result here for $q = 2$, we have

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

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 re-write the demands of both cohorts that are actively trading as

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

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, having more experience, 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 present and past dividends on prices,

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

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

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

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 auto-correlation, 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 auto-correlation 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 Results for 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, then 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 WLOG, 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 (3).

The model set-up 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*

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

The expression for the risky-asset demands in equation (13) allows us to derive equilibrium prices. Note that equation (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 $t \in \mathbb{Z}$*

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

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

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

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 .¹⁰

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.

¹⁰ 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.

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 $k > 0$.*

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 will derive and test these empirical implications in Section 6.

4.2 Cross-Section of Asset Holdings

Experience-based learning also 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 downward 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 realization 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:

$$\xi(n, k, t) = \frac{E_t^n[\theta] - E_t^{n+k}[\theta]}{\gamma(1 + \beta_0)\sigma^2} \quad \forall k = \{0, \dots, t - n\}, n = \{t - q + 1, \dots, t\} \quad (17)$$

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 $\xi(n + a, k, t + a) \approx \xi(n, k, t)$.¹¹

¹¹ This last claim follows since the inter-temporal 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 $(\sum_{j=0}^{t-n-k} \{w(j, \lambda, t - n) - w(j, \lambda, t - n - k)\}(d_{t+a-j} - d_{t-j}) + \sum_{j=t-n-k+1}^{t-n} w(j, \lambda, t - n)(d_{t+a-j} - d_{t-j}))/(\gamma(1 + \beta_0)\sigma^2)$.

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 non-decreasing 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 non-increasing, 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:

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

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 (18) can be expressed as*

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

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

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, 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.

Thought Experiment. Suppose that, for $t - t_0 > q$, $d_{t_0} = d_{t_0+1} = \dots = d_{t-1} = \bar{d}$ and that $d_t \neq \bar{d}$. 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:

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

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, $|d_t - \bar{d}|$, independently of whether the change is positive or negative, and also proportionally to a function that reflects the dispersion of the weights agents assign 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 $w(0, \lambda, age) \rightarrow 1$ for all $age \in \{0, \dots, q-1\}$. 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 and Equilibrium Prices

The results derived so far illustrate a key feature of experience-based learning: 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.¹² The goal of this exercise is to understand how a larger fraction

¹²We have also analyzed the implications of a growing market population, as opposed to a one-time market demographic shock. In Online Appendix [OA.3](#), we show that population growth generates a positive trend in prices, which is independent of experience effects: The growing mass of agents increases

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 .¹³ 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 (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 backwards 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

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

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, experience-based learning 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.

¹³ 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 scenario, other forces may be at play as well. The study of these interactions is out of the scope of this paper.

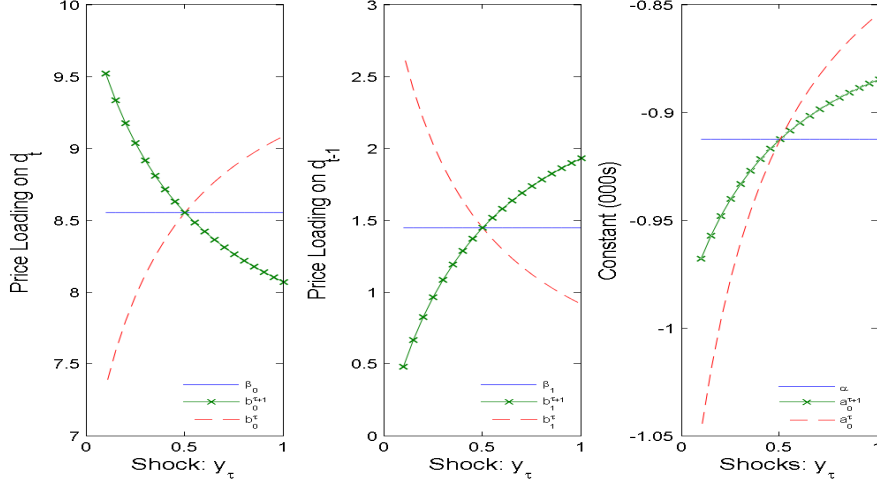


Figure 2: Demographic Shocks and Price Coefficients.

Note: This figure plots coefficients $\{\beta_0, b_{0,\tau}, b_{0,\tau+1}\}$, $\{\beta_1, b_{1,\tau}, b_{1,\tau+1}\}$, and $\{\alpha, a_\tau, a_{\tau+1}\}$, respectively, as a function of the demographic shock y_τ . The results are for $y = 0.5$, $\gamma = 1$, $\lambda = 3$, $\sigma = 1$, and $R = 1.1$.

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 (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

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

Figure 2 shows how the reliance of prices on past dividends changes with the fraction of young agents in the market at time τ . From the first two panels, we can see that as the fraction of young people in the market increases, the more current dividends matter relative to past dividends for the determination of prices. Consistent with this, when

the τ -generation becomes old, prices depend less on contemporaneous dividends and more on past dividends. Finally, the third panel shows that an increase in the overall market population generates a level increase in prices that is captured by the higher price constant. This effect stems from the higher overall demand for the risky asset since there are more people in the market. All predictions are reversed when the fraction of young agents in the market decreases, as can be seen at the left side of each graph.

6 Empirical Implications

In this section, we analyze the empirical implications of our model. The analysis consists of two approaches. On the one hand, we turn to existing evidence on asset-pricing features 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 experience-based learning model is able to quantitatively match these empirical findings, and that it generates refined predictions about their relation with the demographic composition of investors. On the other hand, we test the novel predictions generated by our model regarding the implications of the demographic composition on the predictability of the price-dividend ratio, trade volume, and the cross-section of asset holdings. We show that these predictions are in line with evidence from micro-level data in the Survey of Consumer Finances (SCF) and the Center for Research in Security Prices (CRSP).

6.1 Quantitative Implications for Asset-Pricing Moments

We first show that experience-based learning can explain several 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, Greenwood, Jin, and Shleifer \(2015\)](#), 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. For example, instead of stock returns we measure price changes ΔP , and instead of the price-dividend ratio 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. Another feature of our model is the small number of parameters to be set for generating numerical results. Following [Barberis, Greenwood, Jin, and Shleifer \(2015\)](#), 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 $q \in \{2, 5\}$.

Predictability of Excess Returns. A 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. Experience-based learning rationalizes such predictability and, at the same time, limits it to those dividend realizations experienced by the oldest cohort participating in the market.

In order 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 co-movement between the analogues of the dividend-price ratio and returns, namely, between dividend-price differences $D_t - P_t$ and price changes $P_{t+z} - P_t$ over return horizon z :

$$B_t^R(z) \equiv \frac{\text{Cov}(D_t - P_t, P_{t+z} - P_t)}{\text{Var}(D_t - P_t)}. \quad (23)$$

We compute $B_t^R(z)$ using equation (14) from [Proposition 4.2](#).

We first calculate $B_t^R(z)$ for different horizons z . Panel (a) of [Figure 3](#) plots $B_t^R(z)$ for z ranging from 1 to 7, and different levels of recency bias ($\lambda \in \{1, 3\}$), in an economy with

$q = 5$. As the panel shows, the experience-based learning model generates a positive (and strong) relation between $D - P$ and future price changes, which increases with the return horizon. These predicted patterns are consistent with the empirical findings described in (Cochrane 2011). 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 over-extrapolation model of Barberis et al. (2015) and Barberis et al. (2016), our explanation relies on agents' overweighting recent realizations.

Our model has the additional implication that different demographic structures generate different β 's, which directly determine the level of predictability (or extrapolation) of stock returns. We show this by studying how the coefficient $B_t^R(1)$ varies with the fraction of young market participants. Here, we use the results from Section 5 on the effect of a one-time unexpected shock to the mass of young agents (in an economy with $q = 2$). We assume that the fraction of young agents equals y at all times before and after t , but there is a one-time exogenous shock in t , resulting in $y_t \neq y$. The resulting variation in next period's return predictability (based on this period's dividend-price ratio), $B_t^R(1)$, and in the current period's return predictability (based on last period's dividend-price ratio), $B_{t-1}^R(1)$, are shown in panels (b) and (c) of Figure 3. For both graphs we fix y at 0.5, and we plot y_t over a range $[0, 1]$.

As the plots show, the predictability of next period's return, $B_t^R(1)$, decreases in the number of young market participants (panel b), while the predictability of this period's return, $B_{t-1}^R(1)$, increases in their number (panel c). The key channel is the differential impact on the variance of the dividend-price ratio. In both cases the covariance between future returns and the dividend-price ratio increases in the fraction of young agents; but only the variance of the dividend-price ratio at time t increases such that it off-sets the increase in covariance. Thus, return predictability is affected by the demographic

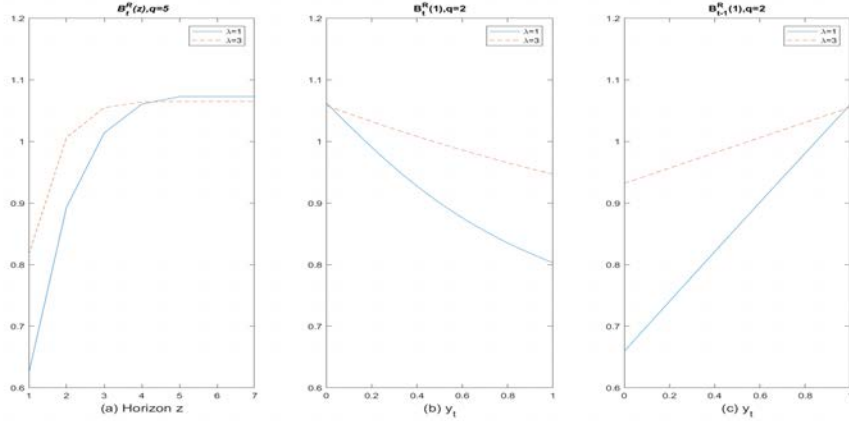


Figure 3: Predictive Power of $D_t - P_t$ for $P_{t+z} - P_t$.

Note: This figure plots the coefficient B_t^R for two levels of recency bias, $\lambda \in \{1, 3\}$. Panel (a) shows how $B_t^R(z)$ varies with the return horizon z (for $q = 5$). Panel (b) shows how $B_t^R(1)$ varies with the fraction y_t of young agents in the market at time t , and panel (c) with the fraction y_{t+1} of young agents in the market at time $t + 1$, in both cases for $q = 2$. The results are calculated for $y = 0.5$ and $R = 1.05$.

composition of market participants, and the effect is sensitive to the timing of the participation shock. With a larger generation of market participants coming in at t , the return experienced in that period is more predictable.

Predictability of Price-Dividend Ratio. In addition to the predictability of returns, we can also compute the predictability of the price-dividend ratio implied by the model. That is, we relate past P/D ratios to future P/D realizations, and analyze the persistence of the price-dividend 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:

$$B_t^{PD}(z) = \frac{\text{cov}(P_{t+z} - D_{t+z}, P_t - D_t)}{\text{var}(P_t - D_t)} \quad (24)$$

We first calculate how B_t^{PD} varies with the horizon z . Panel (a) of Figure 4 displays how B_t^{PD} varies for z ranging from 1 to 7, and for different levels of recency bias ($\lambda \in \{1, 3\}$),

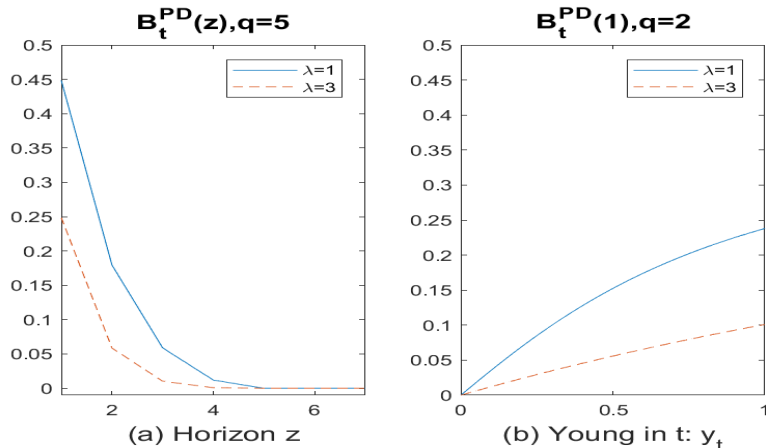


Figure 4: $P_t - D_t$ autocorrelation.

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

in an economy with $q = 5$. As in the data, $P - D$ is highly autocorrelated at short lags, with the autocorrelation being zero at longer horizons. Furthermore, B_t^{PD} decreases in the extent of recency bias present in the population.

We then turn to the demographic structure, using again the results from Section 5. As shown in panel (b) of Figure 4, B_t^{PD} increases with the fraction of young agents in the market at time t , and the effect is weaker under higher recency bias. Furthermore, B_t^{PD} increases with the number of cohorts in the market. A direct implication is that $P - D$ is positively correlated only with lagged realizations where the number of periods lagged is below the number of cohorts in the market.

Price Dynamics. A third set of asset pricing implications are related to the dynamics of prices, and in particular the excess volatility puzzle. Campbell and Kyle (1993) estimate the volatility of dividends to be 0.032 that of the change in prices. We calculate the corresponding volatility of prices relative to that of dividends implied by our model.

Table 1 presents our results for $\sigma(P - D)$, $\sigma(P)$, and the ratio $\sigma(D)/\sigma(P)$. We see

Table 1: Excess Volatility.

	$\sigma(\mathbf{P-D})$	$\sigma(\mathbf{P})$	$\sigma(\mathbf{D}) / \sigma(\mathbf{P})$
q=5			
$\lambda=1$	3.07	3.24	0.08
$\lambda=3$	3.71	4.10	0.06
$\lambda=10$	4.58	5.08	0.05
q=10			
$\lambda=1$	2.31	2.24	0.11
$\lambda=3$	2.90	3.08	0.08
$\lambda=10$	4.06	4.52	0.06

that the experience-based learning model generates ample excess volatility in prices and in the price-dividend differences, and this volatility is increasing in the recency bias and decreasing with the number of cohorts in the market. Furthermore, in our calibration, the volatility of dividends is between 0.03 and 0.05 that of the change in prices, consistent with the 0.032 estimate of Campbell and Kyle.

6.2 Demographics on Price-Dividend Predictability

The predictability results in the previous subsection are consistent with the findings of [Cassella and Gulen \(2015\)](#), 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 experience-based learning 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 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. Experience-based learning 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).

In order to test these predictions, we regress P/D onto lags of itself interacted with a dummy variable that indicates a larger presence of young people in the market. As the estimation model has to account for non-stationarity 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, non-linear dynamics in a tractable way.¹⁴ The regression model is thus given by

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

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 1 if the share of young generations participating in the market at time t is large relative to the participation of older generations, and 0 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 1 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 1 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 Online-Appendix OA.4.

The theoretical prediction of our model is that the correlation between future and

¹⁴ For a more thorough discussion of MSR see Hamilton (1989).

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 (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 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 non-significant, 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 experience-based learning 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.

Table 2: Markov-Switching Regime (MSR) Model

Estimation results for model specification (25), where $P_t - D_t$ is the log of the price-to-dividend ratio, and 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 1 when the fraction of investors under 50 is larger than 0.5 (in column 1), or as an indicator equal to 1 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 is from the SCF, stock data from CRSP.

	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

Robust standard errors in parentheses. * significant at 10%; ** significant at 5%.

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 dividends 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) U.S. 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 2 years and higher than 8 years,¹⁵ for all non-stationary series (dividend, earnings, and GDP).

In order to construct the experienced returns, dividends, earnings, and GDP of different generations over the course of their lives, we apply the formula from equation (5). 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 \(2011\)](#).

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

¹⁵ 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 with the natural logarithm of earnings and GDP to remove non-linearities in these series. In unreported analyses, we also use the natural logarithm of dividends and obtain very similar results.

of household-level micro data is the cross-sectional data on asset holdings and various household background characteristics in the Survey of Consumer Finances (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.¹⁶

For the extensive margin of stock holdings, we construct an indicator of stock-market participation. It equals 1 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).¹⁷

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 U.S. population.¹⁸

¹⁶ Those are 1960, 1962, 1963, 1964, 1967, 1968, 1969, 1970, 1971, and 1977.

¹⁷ 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.

¹⁸ The 1983-2013 SCF waves oversample high-income households with significant stock holdings. The oversampling is helpful for our analysis of asset allocation, but could induce selection bias. By applying

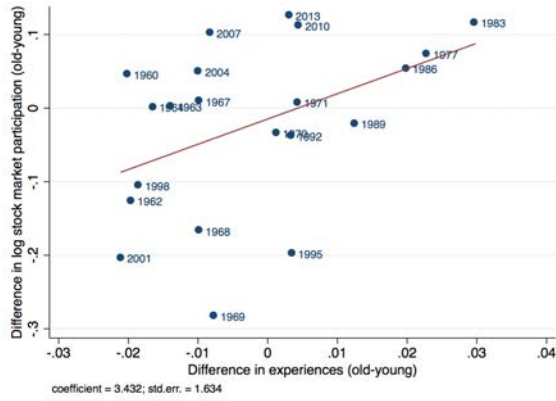
We present the results graphically. We plot the relation between stock holdings (extensive and intensive margin) and experienced returns (Figure 5), dividends (Figure 6), earnings (Figure 7), and GDP (Figure 8). Graphs 5.(a) and 5.(c) update the evidence on the extensive margin and returns presented 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 Figure 5, 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.

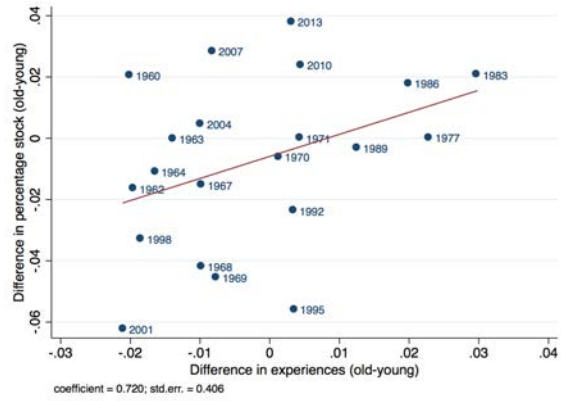
The analysis of the intensive margin of stock-market investment yields the same conclusion. Both graph (c) and graph (d) indicate that older generations invest a higher share of their liquid assets in stock, compared to the younger generations, when their experienced returns have been higher than those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

Figures 6 to 8 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.

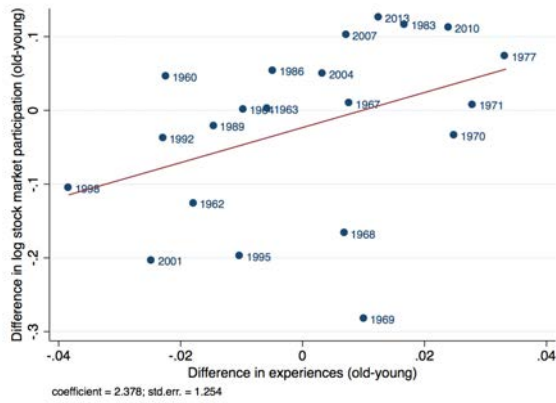
SCF sample weights, we undo the over-weighting of high-income households and also adjust for non-response bias.



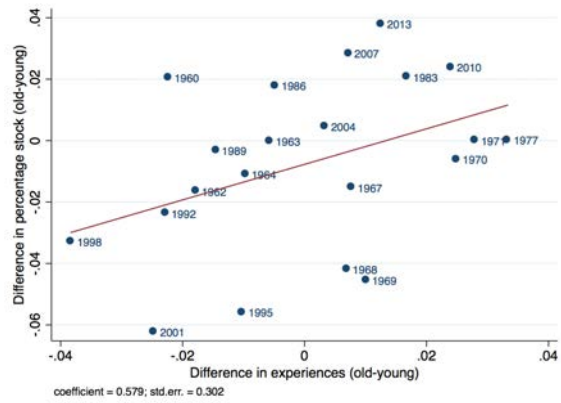
(a) Stock-market participation ($\lambda = 1$)



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



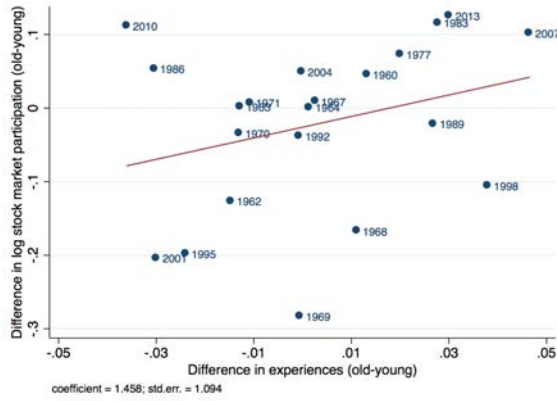
(c) Stock-market participation ($\lambda = 3$)



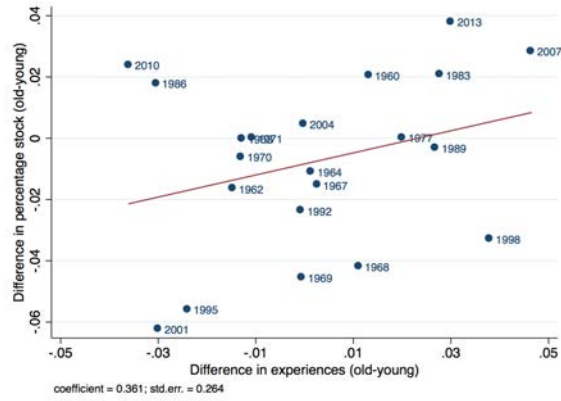
(d) Fraction invested in stock ($\lambda = 3$)

Figure 5: Experienced Returns and Stock Holdings

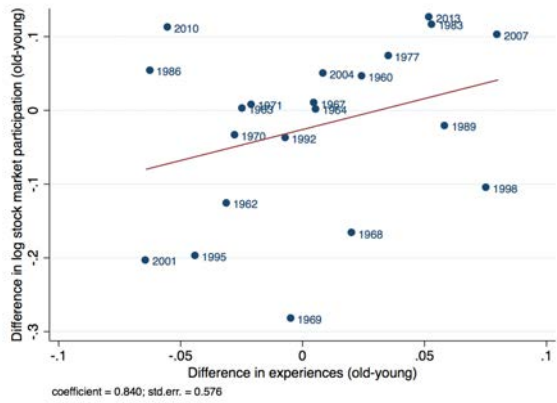
Difference in experienced returns is calculated as the lifetime average experienced returns of the S&P500 Index as given on Robert Shiller’s website, using declining weights with either $\lambda = 1$ or $\lambda = 3$ as in equation (5). *Stock-market participation* is measured as the fraction of households 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.



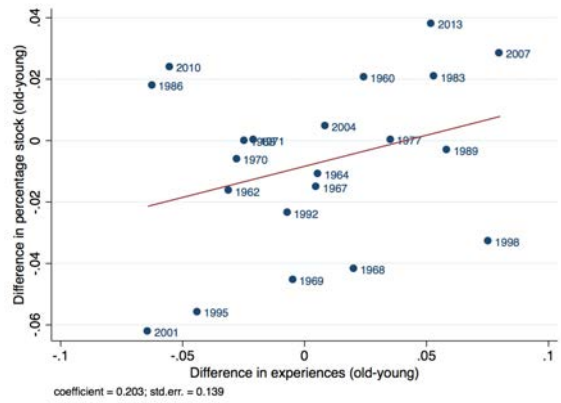
(a) Stock-market participation ($\lambda = 1$)



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



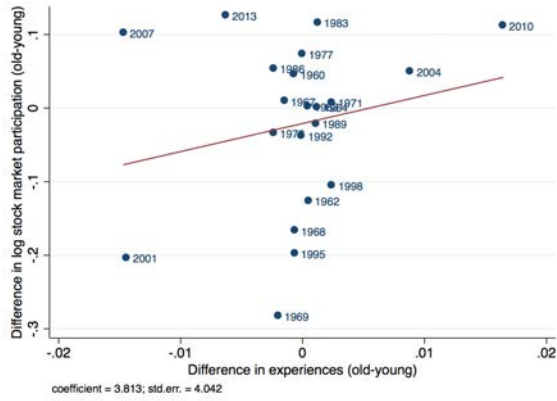
(c) Stock-market participation ($\lambda = 3$)



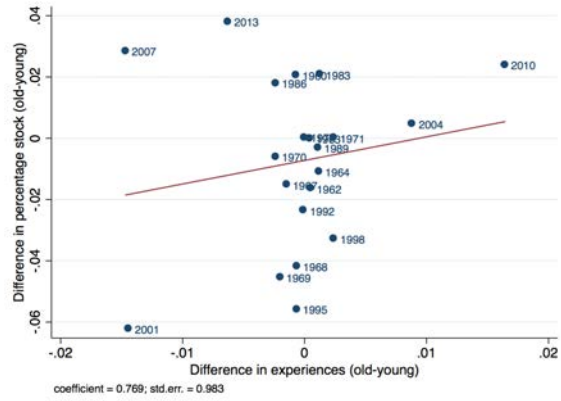
(d) Fraction invested in stock ($\lambda = 3$)

Figure 6: Experienced Dividends and Stock Holdings

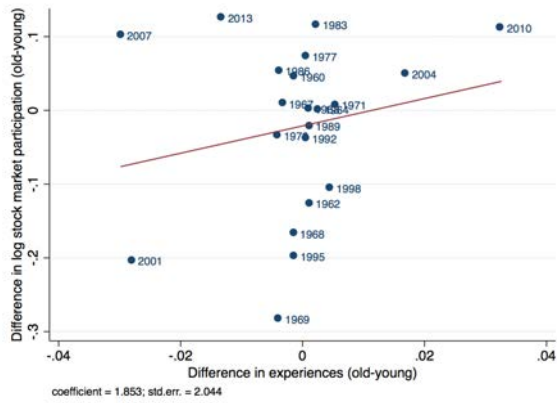
Difference in experienced dividends is calculated as the lifetime average experienced real dividends as given on Robert Shiller’s website, using declining weights with either $\lambda = 1$ or $\lambda = 3$ as in equation (5). *Stock-market participation* is measured as the fraction of households in the 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.



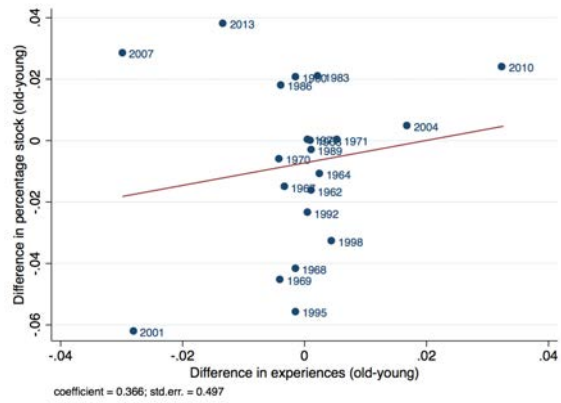
(a) Stock-market participation ($\lambda = 1$)



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



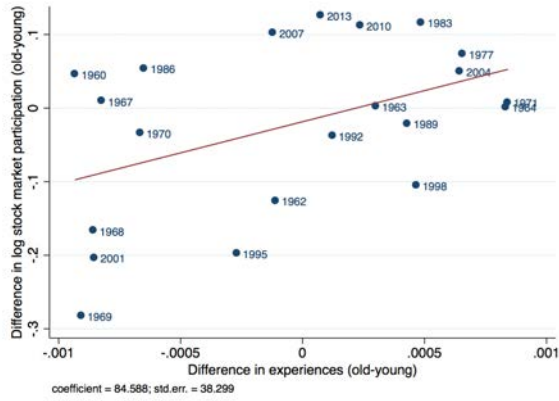
(c) Stock-market participation ($\lambda = 3$)



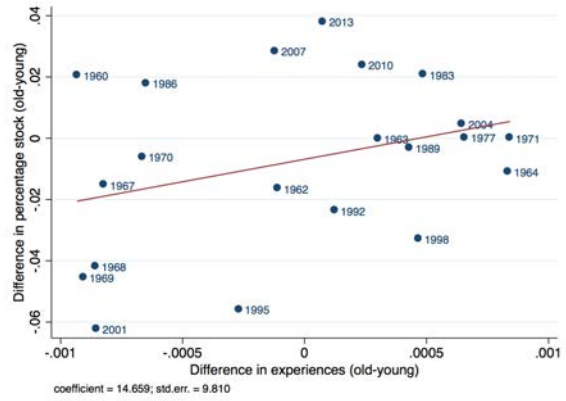
(d) Fraction invested in stock ($\lambda = 3$)

Figure 7: Experienced Earnings and Stock Holdings

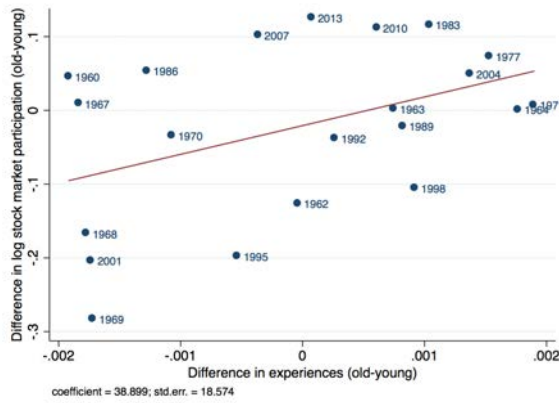
Difference in experienced earnings is calculated as the lifetime average experienced log real earnings as given on Robert Shiller’s website, using declining weights with either $\lambda = 1$ or $\lambda = 3$ as in equation (5). *Stock-market participation* is measured as the fraction of households in 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.



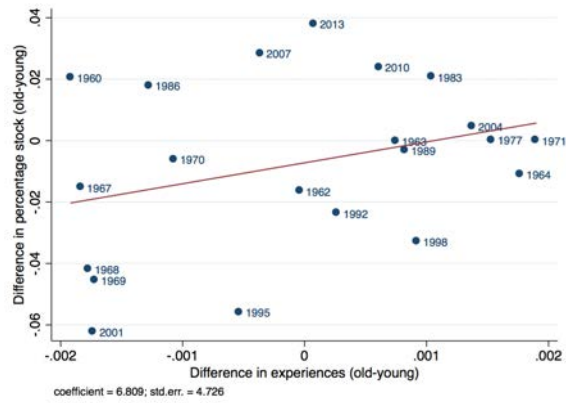
(a) Stock-market participation ($\lambda = 1$)



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



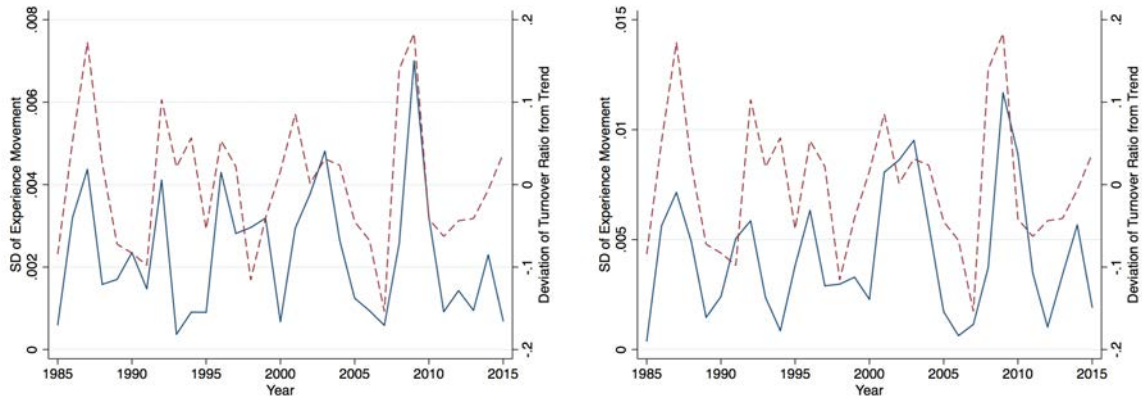
(c) Stock-market participation ($\lambda = 3$)



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

Figure 8: Experienced Log GDP and Stock Holdings

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 equation (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.



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

(b) Stock-market participation ($\lambda = 3$)

Figure 9: Trading Volume and Standard Deviation of Changes in Experienced Returns

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.¹⁹

As a measure of abnormal trade volume, we calculate the deviation of the turnover ratio from its trend. Following prior literature (Statman, Thorley, and Vorkink (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, Thorley, and Vorkink (2006)). Then, we aggregate these numbers into a market-wide turnover ratio, weighting firms by their market capitalization.²⁰ Since the turnover ratio is non-stationary, 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

¹⁹ For this, we obtain data on U.S. population by age between 1985 and 2015 from US Census Bureau.

²⁰ 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.

we keep frequencies between 2 and 8 years. We examine the co-movement between the aforementioned measure of disagreement, i. e., the standard deviation of the change in experienced stock returns, and the above measures of (abnormal) trade volume.

Figure 9 displays trading volume and disagreement in experienced returns between cohorts over time. Graph (a) shows the results when we apply linear weights for the calculation of experienced returns, graph (b) displays the case with super-linear 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, Figure 9 compares the variation 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 trade repeatedly based on experienced performance.

Table 3: Correlation between Trading Volume and S.D. of Changes in Experiences

Experiences constructed using:		Returns	Dividends	Log Earnings	Log GDP
$\lambda = 1$	correlation	0.5982	0.2008	0.3287	0.1699
	<i>p</i> -value	0.0004	0.2786	0.0710	0.3610
$\lambda = 3$	correlation	0.5330	0.1635	0.3168	0.1844
	<i>p</i> -value	0.0013	0.3794	0.0825	0.3206

Consistent with the predictions of our model, we observe a clear co-movement between disagreement among cohorts and trading volume. Table 3 reveals that the co-movement is statistically significant at 1%. The table presents the correlation between trading volume and our return disagreement variable, 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 positive, albeit (marginally) significant only for changes in disagreement in experienced earnings.

All empirical findings are consistent with experience-based learning. Alternative models of belief formation in macro-finance cannot easily explain these findings jointly.

7 Extension: Non-Myopic Agents

As a final step, we analyze the model predictions when the assumption of myopic agents is removed. We consider a version of our model where agents re-balance their portfolios every period to maximize the expected utility of consuming their final wealth. Thus, agents need to account for the dynamic nature of their portfolio problem. This approach to modeling dynamic portfolio choices follows a widely used approach in the literature (cf. [Vives \(2010\)](#)). The analysis reveals two additional determinants of risky-asset demand that are at work when we remove the myopia assumption: the “discount effect,” due to investment horizon considerations, and the “dynamic effect,” due to agents wanting to hedge their exposure to future changes in beliefs.

We characterize equilibrium prices in the new economy, and show that the result of [Proposition 4.2](#) continues to hold: Prices are affine functions of past dividends observed by the generations that are trading. To sharpen our predictions, we then return to the two-generation model, $q = 2$, and compare the findings with those with myopic agents.

7.1 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 . At time n , an n -generation agent solves the following problem:

$$\max_{\mathbf{x} \in \mathbb{R}^q} E_n^n [-\exp(-\gamma W_{n+q}^n(\mathbf{x}))] \quad (26)$$

$$\text{s.t. } W_{n+q}^n(\mathbf{x}) = \sum_{\tau=n}^{n+q} R^{n+q-\tau} x_\tau s_{\tau+1} \quad (27)$$

where $\mathbf{x} \in \mathbb{R}^q$ are the q trading decisions from n up to $n_q = n + q - 1$. We continue to assume that the initial wealth of all generations is zero, i. e., $W_n^n = 0 \forall n$.

We can cast this problem iteratively — by solving from n_q backwards — as

$$V_{n_q}^n(d_{n_q-K:n_q}) = \max_{x \in \mathbb{R}} E_{n_q}^n [-\exp(-\gamma s_{n+q} x)] \quad \text{and} \quad (28)$$

$$V_\tau^n(d_{\tau-K:\tau}) = \max_{x \in \mathbb{R}} E_\tau^n [V_{\tau+1}^n(d_{\tau+1-K:\tau+1}) \exp(-\gamma s_{\tau+1} x)] \quad \forall \tau \in \{n, \dots, n_q - 1\} \quad (29)$$

Note that our definition of $V_{n_q}^n$ omits the term $\exp\{-\gamma R W_{n_q}^n\}$ since it does not affect the maximization. The same applies for all $\tau < n + q - 1$, and we thus omit the wealth at time τ as well. The formulation of the maximization problem shows that, while the n -generation's problem is a static portfolio problem at n_q , it is not static at any other time $\tau \in \{n, \dots, n_q - 1\}$ because $V_{\tau+1}^n$ is correlated with $s_{\tau+1}$ through dividends. Dividend realization $d_{\tau+1}$ impacts (i) the net payoff obtained from investing x_τ in the risky asset at time τ , and (ii) the continuation value $V_{\tau+1}^n(d_{\tau+1-K:\tau+1})$ by affecting the beliefs of the n -generation at $\tau + 1$ and the resulting portfolio decision. As a result, this dynamic portfolio problem cannot be expressed as a succession of static problems, as is standard in the literature for the rational model (see [Vives \(2010\)](#)).

Let $E_{N(\mu, \sigma^2)}[\cdot]$ and $V_{N(\mu, \sigma^2)}[\cdot]$ be the expectation and variance with respect to a Gaussian pdf with mean μ and variance σ^2 . We can then characterize agents' demands:

Proposition 7.1. *Suppose $p_t = \alpha + \sum_{k=0}^K \beta_k d_{t-k}$ with $\beta_0 \neq -1$. Then, the demand of any generation n at age $j \in \{0, \dots, q - 1\}$, i. e., in period $n + j$, is given by*

$$x_{n+j}^n = \frac{E_{N(m_j, \sigma_j^2)}[s_{n+j+1}]}{\gamma R^{q-1-j} V_{N(m_j, \sigma_j^2)}[s_{n+j+1}]}, \quad (30)$$

where $m_j \equiv \frac{\theta_{n+j}^n - \sigma^2 (b_j + \sum_{k=1}^K b_j(k) d_{n+j-k})}{2c_j \sigma^2 + 1}$, $\sigma_j^2 \equiv \frac{\sigma^2}{2c_j \sigma^2 + 1}$, and $\{\{b_j(k)\}_{k=1}^{q-1}, b_j, c_j\}$ are constants that change with the agent's age j . (See the proof for exact expressions.)

The intuition of the proof is as follows. Solving the problem backwards, we note that, at time n_q (see equation (28)), the problem is in fact static. In particular, $V_{n_q}^n$ is exponential-quadratic in d_{n_q} . (See Lemma B.1 in the Appendix.) We then show (in Lemma C.1) that the exponential-quadratic term times the Gaussian distribution of dividends imply a new Gaussian distribution with a slanted mean and variance. Thus the problem at time $n_q - 1$ can be viewed as a static problem with a modified Gaussian distribution. As a result, demands are of the form shown in (30), and $V_{n_q-1}^n$ is also exponential-quadratic. The same process continues backwards until time n .

After straightforward algebra, we can cast equation (30), as²¹

$$\begin{aligned} x_{n+j}^n &= \frac{1}{R^{q-1-j}} \cdot \frac{E_{N(\theta_{n+j}^n, \sigma^2)}[s_{n+j+1}]}{\gamma V_{N(\theta_{n+j}^n, \sigma^2)}[s_{n+j+1}]} - \frac{b_j + \sum_{k=1}^K b_j(k) d_{n+j-k}}{\gamma R^{q-1-j} (1 + \beta_0)} \\ &\equiv \frac{1}{R^{q-1-j}} \cdot \tilde{x}_{n+j}^n + \Delta_{n+j}^n \end{aligned} \quad (31)$$

The term \tilde{x}_{n+j}^n coincides with the demand of a *static* portfolio problem for an agent with beliefs θ_{n+j}^n (see Proposition 4.1). We coin this term the *myopic component* of the demand for risky assets. The scaling by $1/R^{q-1-j}$ arises because agents discount the future by R . The second term, $\Delta_{n+j}^n \equiv -\frac{b_j + \sum_{k=1}^K b_j(k) d_{n+j-k}}{\gamma R^{q-1-j} (1 + \beta_0)}$, is an adjustment that accounts for the dynamic nature of the problem. We denote it the *dynamic component*. It arises because agents understand that they are learning about the risky asset, and that the value function is therefore correlated with the one-period-ahead returns.

The following proposition shows that, in a linear equilibrium, prices at any time t depend on past dividends, but only on those observed by the generations trading at t :

Proposition 7.2. *For $R > 1$, the price in any linear equilibrium with $\beta_0 \neq -1$ is affine in the history of dividends observed by the oldest generation participating in the market.*

For any $t \in \mathbb{Z}$, $q \geq 1$, $p_t = \alpha + \sum_{k=0}^{q-1} \beta_k d_{t-k}$.

²¹ Note that $E_{N(b+a,s)}[s_{n+1}] = E_{N(a,s)}[s_{n+1}] + (1 + \beta_0)b$.

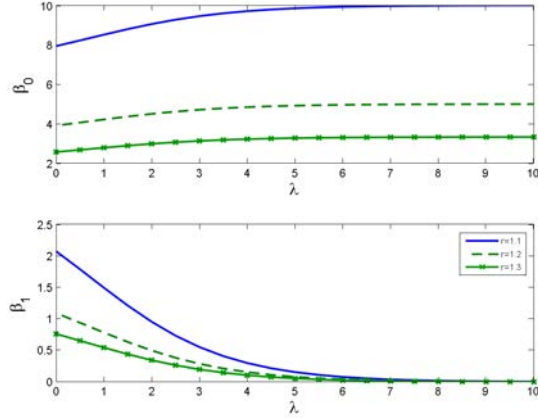


Figure 10: Comparative Statics: Sensitivity of Prices to Dividends for $q = 2$.

The proof follows along the same lines as the one for the myopic case.²² This result shows that the insights in Proposition 4.2 continue to hold in the general setup.

7.2 Illustration: The $q = 2$ Case

We now return to the special case of agents who live for two periods, $q = 2$, to sharpen our results on equilibrium prices and demands. We relegate the derivation of the coefficients of the price function, $\{\alpha, \beta_0, \beta_1\}$, to Online Appendix OA.2. (Lemma OA.2.1 solves the corresponding system of linear equations.) With this result, we are able to establish:

Proposition 7.3. *For $\lambda > 0$, prices react positively to dividends d_t and d_{t-1} , with $\alpha \leq 0$ and $0 < \beta_1 < R\beta_0$.*

This proposition is analogous to Proposition 4.3. It reflects the fact that the dividends at time t are observed by both generations, whereas d_{t-1} is only observed by the old generation. The relation between the recency bias λ and price loadings is as in the model

²² Regarding the restriction $\beta_0 = -1$, note that, heuristically, an equilibrium with $\beta_0 = -1$ is not well-defined since, in this case, the excess payoff in a given period is deterministic given the information in the prior period, and thus the agent will take infinite positions.

with myopic agents, and is depicted in Figure 10 for different risk-free rates.²³

The next proposition establishes that the demand of the young generation increases and the demand of the old generation decreases, when current dividends increase, and vice versa if current dividends decrease (as in Proposition 4.4). However, as we explain below, there are now several partly off-setting motives at work.

Proposition 7.4. *For $\lambda > 0$, (1) $\frac{\partial x_t^t}{\partial d_t} > 0 > \frac{\partial x_t^{t-1}}{\partial d_t}$, and (2) $\frac{\partial x_t^t}{\partial d_{t-1}} < 0 < \frac{\partial x_t^{t-1}}{\partial d_{t-1}}$.*

The basic intuition for this result remains the same as in the myopic setting. However, this is not the only effect to consider when agents solve a dynamic portfolio problem. There are additional effects due to the fact that the young are confronted with a different investment horizon; and that all agents have hedging motives due to the correlation between future returns and beliefs. We show in the proof of Proposition 7.1 that the force introduced by EBL dominates the other forces, generating the same prediction as in the model with myopic agents.

Overall, the results of this section illustrate that our main findings are still present when we remove the myopia assumption. At the same time, this setting reveals additional factors influencing agents' decision-making due to their understanding of how future returns relate to future learning and resulting investment decisions. An interesting question for future research is whether such influences are indeed at work or whether a lack thereof implies investor naivete about how future experiences affect future risk attitudes.

8 Conclusion

In this paper, we have proposed an OLG equilibrium framework to study the effect of personal experiences on market dynamics. We incorporate the two main empirical

²³ The values of $\{\beta_0, \beta_1\}$ are independent of the process for dividends, σ^2 , and of the coefficient of risk aversion, γ . Thus, the results shown in the figure do not depend on parameter values other than the ones used for comparative statics, (λ, R) .

features of experience effects, the over-weighting of lifetime experiences and recency bias, into the belief formation process of agents. We show that experience-based learning 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. We highlight two channels through which shocks have long-lasting effects on economic 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 illustrate how the demographic composition of an economy can have important implications for the extent to which prices depend on past dividends. We consider this paper to be a first step into the exploration of the role of demographics in understanding market dynamics.

References

- Adam, K., A. Marcet, and J. P. Nicolini (2012). Stock Market Volatility and Learning. *Journal of Finance*, forthcoming.
- Agarwal, S., J. C. Driscoll, X. Gabaix, and D. Laibson (2013). Learning in the Credit Card Market. *working paper*.
- Alesina, A. and N. Fuchs-Schundeln (2007). Good-bye Lenin (or Not?): The Effect of Communism on People's Preferences. *American Economic Review* 97, 1507–1528.
- Atkins, A. B. and E. A. Dyl (1997). Market structure and reported trading volume: NASDAQ versus the NYSE. *Journal of Financial Research* 20(3), 291–304.
- Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2015). X-CAPM: An extrapolative capital asset pricing model. *Journal of Financial Economics* 115(1), 1–24.
- Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2016). Extrapolation and bubbles. Working Paper, National Bureau of Economic Research.
- Barsky, R. and J. B. DeLong (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 on Demand.
- Campbell, J. Y. (2008). Household Finance. *Journal of Finance* 61(4), 1553–1604.
- Campbell, J. Y. and A. S. Kyle (1993). Smart Money, Noise Trading, and Stock Price Behavior. *Review of Economic Studies* 60, 1–34.
- Campbell, J. Y. and R. J. Shiller (1988). Stock Prices, Earnings, and Expected Dividends. *Journal of Finance* 43(3), 661–676.
- Cassella, S. and H. Gulen (2015). Extrapolation bias and the predictability of stock returns by price-scaled variables. *working paper*.
- Cecchetti, S. G., P. Lam, and N. C. Mark (2000). Asset Pricing with Distorted Beliefs: Are Equity Returns Too Good to Be True? *American Economic Review* 90(4), 787–805.
- Chiang, Y.-M., D. Hirshleifer, Y. Qian, and A. E. Sherman (2011). Do Investors Learn from Experience? Evidence from Frequent IPO Investors. *Review of Financial Studies* 24, 1560–1589.
- Christiano, L. J. and T. J. Fitzgerald (2003). The band pass filter. *international economic review* 44(2), 435–465.
- Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of Finance* 66(4), 1047–1108.
- Cogley, T. and T. J. Sargent (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., M. Johannes, and L. A. Lochstoer (2016). Asset Pricing when ‘This Time is Different’. *Swiss Finance Institute Research Paper* (13-73), 14–8.
- Ehling, P., A. Graniero, and C. Heyerdahl-Larsen (2015). Asset Prices and Portfolio Choice with Learning from Experience. *working paper*, NYU.
- Fama, E. F. and K. R. French (1988). Dividend yields and expected stock returns. *Journal of financial economics* 22(1), 3–25.
- Fuster, A., B. Hebert, and D. Laibson (2011). Natural Expectations, Macroeconomic Dynamics and Asset Pricing. In D. Acemoglu and M. Woodford (Eds.), *NBER Macroeconomics Annual 2011*. Chicago, IL: University of Chicago Press.
- Fuster, A., D. Laibson, and B. Mendel (2010). Natural Expectations and Macroeconomic Fluctuations. *Journal of Economic Perspectives* 24, 67–84.
- Giuliano, P. and A. Spilimbergo (2013). Growing Up in a Recession. *Review of Economic Studies* 81(2), 787–817.
- Greenwood, R. and A. Shleifer (2014). Expectations of Returns and Expected Returns. *Review of Financial Studies*, 714–746.

- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 357–384.
- Hamilton, J. D. (1994). Chapter 50 State-space models. Volume 4 of *Handbook of Econometrics*, pp. 3039 – 3080. Elsevier.
- Hertwig, R., G. Barron, E. U. Weber, and I. Erev (2004). Decisions from Experience and the Effect of Rare Events in Risky Choice. *Psychological Science* 15, 534 C539.
- Jin, L. J. (2015). A speculative asset pricing model of financial instability. *Working Paper*.
- Kaustia, M. and S. Knüpfer (2008). Do Investors Overweight Personal Experience? Evidence from IPO Subscriptions. *Journal of Finance* 63, 2679–2702.
- LeRoy, S. F. (2005). Excess Volatility. *working paper*.
- LeRoy, S. F. and R. D. Porter (1981). The present value relation: Tests based on implied variance bounds. *Econometrica* 49, 555–574.
- Lo, A. W. and J. Wang (2000). Trading volume: definitions, data analysis, and implications of portfolio theory. *Review of Financial Studies* 13(2), 257–300.
- Malmendier, U. and S. Nagel (2011). Depression Babies: Do Macroeconomic Experiences Affect Risk-Taking? *Quarterly Journal of Economics* 126, 373–416.
- Malmendier, U. and S. Nagel (2016). Learning from Inflation Experiences. *Quarterly Journal of Economics* 131, 53–87.
- Malmendier, U. and L. S. Shen (2017). Scarred Consumption. Working paper, UC-Berkeley.
- Oreopoulos, P., T. von Wachter, and A. Heisz (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*, 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., N. Karlsson, G. Loewenstein, and D. Ariely (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., S. Thorley, and K. Vorkink (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. *The Review of Economic Studies* 63(4), 523–557.
- Tversky, A. and D. Kahneman (1974). Judgment under uncertainty: Heuristics and biases. *Science* 4157, 1124–1131.
- Vives, X. (2010). *Information and Learning in Markets: The Impact of Market Microstructure*. Princeton University Press.
- Weber, E. U., U. Böckenholt, D. J. Hilton, and B. Wallace (1993). Determinants of diagnostic hypothesis generation: effects of information, base rates, and experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19(5), 1151.
- Woodford, M. (2013). Macroeconomic Analysis without the Rational Expectations Hypothesis. *Annual Review of Economics* 5(1), forthcoming.

Appendix A 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 $k \leq age'$, we note that $\Delta(k) > 0 \iff Q(k) := \frac{w(k, \lambda, age)}{w(k, \lambda, age')} > 1$. Hence, it remains to be shown that $\exists k_0 \in \{0, \dots, age'\}$ 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 $w(k, \lambda, age)$ are independent of k (see the definition in (5)), we absorb them in a constant $c \in \mathbb{R}^+$ and rewrite

$$Q(k) = c \cdot \frac{(age + 1 - k)^\lambda}{(age' + 1 - k)^\lambda} = c \cdot \left[\frac{age + 1 - k}{age' + 1 - k} \right]^\lambda = c \cdot \alpha(k)^\lambda \quad \forall k \in \{0, \dots, age'\}. \quad (32)$$

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

Appendix B Proofs for Results in Section 4

Proposition 4.1 directly follows from the following Lemma.

Lemma B.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. See Online Appendix OA.1.1. \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 omitted for simplicity.

From Lemma B.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:

$$x_t^n = \frac{(1 + \beta_0) \theta_t^n + \alpha + \beta_1 d_t + \dots + \beta_q d_{t-q+1} - p_t R}{\gamma (1 + \beta_0)^2 \sigma^2} \quad (33)$$

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

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

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:

$$-\frac{\gamma (1 + \beta_0)^2 \sigma^2}{R - 1} = \alpha \quad (34)$$

$$(1 + \beta_0) \frac{1}{q} \sum_{n=t-q+1}^{t-k} w(k, \lambda, t-n) + \beta_{k+1} = \beta_k R \quad \forall k \in \{0, 1, \dots, q-1\} \quad (35)$$

$$0 = \beta_q R. \quad (36)$$

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 equation (36) we obtain:

$$(1 + \beta_0) w_k + \beta_{k+1} = \beta_k R \quad \forall k \in \{0, 1, \dots, q-2\} \quad (37)$$

$$(1 + \beta_0) w_{q-1} = \beta_{q-1} R \quad (38)$$

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 (37) and (38) as

$$(1 + \beta_0) \sum_{j=0}^{k-1} w_{q-(k-j)}/R^j = \beta_{q-k} R \quad \text{for } k = 1, \dots, q. \quad (39)$$

The last expression (39) 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 (34) allows us to obtain the expression for α from (15) in Proposition 4.2. And expression (39) 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 equation (16) of the Proposition. The latter also subsumes equation (38), 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 equations (37) and (38). In addition, note that by construction, $w_k < w_{k-1}$ for $\lambda > 0$ since for all generations, $w(k, \lambda, age)$ 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

$$\beta_{q-1} = \frac{1}{R} (1 + \beta_0) w_{q-1} < \frac{1}{R} [(1 + \beta_0) w_{q-2} + \beta_{q-1}] = \beta_{q-2} \quad (40)$$

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

$$\beta_{k-1} = \frac{1}{R} [(1 + \beta_0) w_{k-1} + \beta_k] < \frac{1}{R} [(1 + \beta_0) w_{k-2} + \beta_{k-1}] = \beta_{k-2} \quad (41)$$

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 $G'_q(\lambda) > 0 \quad \forall q > 0$ and $\forall \lambda > 0$. After some algebra, the terms in $G_q(\cdot)$ can be re-organized as follows:

$$G_q(\lambda) = \sum_{age=0}^{q-1} \frac{1}{q} \sum_{k=0}^{age} w(k, \lambda, age) / R^{k+1} \quad (42)$$

Note that for any $age \in \{0, \dots, q-1\}$: (i) $\sum_{k=0}^{age} w(k, \lambda, age) = 1$ and (ii) for any λ_1, λ_2 such that $\lambda_1 > \lambda_2 > 0$, $\sum_{k=j}^{age} w(k, \lambda_1, age) < \sum_{k=j}^{age} w(k, \lambda_2, age)$. 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 $age \in \{0, \dots, q-1\}$, $\sum_{k=0}^{age} c^{k+1} w(k, \lambda_1, age) > \sum_{k=0}^{age} c^{k+1} w(k, \lambda_2, age)$, and thus $G_q(\lambda_1) > G_q(\lambda_2)$.

To show the limit results, note that $\lim_{\lambda \rightarrow \infty} w(0, \lambda, age) = 1$, while $\lim_{\lambda \rightarrow \infty} w(k, \lambda, age) = 0$ 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 $k \in \{0, \dots, q-1\}$,

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

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 , $w(k, \lambda, age) = 0$ for all $k \geq age$.) 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 B.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. See Online Appendix OA.1.1. \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 to one.

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

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

These quantities, as functions of m , are non-decreasing 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,

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

If the weights are non-decreasing, 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 B.2 with $a = t-n > t-n-k = a'$.

If the weights are non-increasing, 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 $n \leq t$,

$$x_t^n = \frac{1}{\gamma\sigma^2(1+\beta_0)^2} \left(\alpha_0(1-R) + (1+\beta_0)\theta_t^n - R\beta_0 d_t + \sum_{k=1}^{q-1} \beta_k (d_{t+1-k} - R d_{t-k}) \right). \quad (43)$$

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

$$x_t^n - x_{t-1}^n = \frac{(1+\beta_0)(\theta_t^n - \theta_{t-1}^n) + \mathcal{T}(d_{t:t-q})}{\gamma\sigma^2(1+\beta_0)^2} \quad (44)$$

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

$$q^{-1} \left(\sum_{n=t-q}^t x_t^n - x_{t-1}^n \right) = 0. \quad (45)$$

This expression and the expression in (44) imply that

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

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 (44) as follows:

$$x_t^n - x_{t-1}^n = \chi \left[(\theta_t^n - \theta_{t-1}^n) - \frac{1}{q} \sum_{n=t-q}^t (\theta_t^n - \theta_{t-1}^n) \right], \quad \forall n \in \{t, \dots, t - q\} \quad (46)$$

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 C Proofs for Results in Section 7

In this section, we present the proofs to Propositions 7.1-7.4. In order to prove the first two propositions, we first establish the following four lemmas. (The proofs of the lemmas are relegated to the Online Appendix OA.1.2.)

Lemma C.1. *Suppose $z \sim N(\mu, \sigma^2)$, then for any $A, B \in \mathbb{R}$ and $C \geq 0$, $z \mapsto K^{-1} \exp\{-A - Bz - Cz^2\} \phi(z; \mu, \sigma^2)$ is Gaussian with mean $m \equiv -\Sigma^2 B + \Sigma^2 \sigma^{-2} \mu$ and $\Sigma^2 \equiv \frac{\sigma^2}{2C\sigma^2 + 1}$, where*

$$K = E_{N(\mu, \sigma^2)}[\exp\{-A - Bz - Cz^2\}] = \frac{1}{\sqrt{2\sigma^2 C + 1}} \exp\{-(A + 0.5\sigma^{-2}\mu^2) + \frac{m^2}{2\Sigma^2}\}.$$

Lemma C.2. *Demands for the risky asset in the last two period of a cohort- n agent's life are given by $x_{n+q}^n = 0$ and $x_{n+q-1}^n = \frac{E_{n+q-1}^{[s_{n+q}]}}{\gamma\sigma^2(1+\beta_0)^2} \quad \forall n \in \mathbb{Z}, q \geq 1$.*

Lemma C.3. *Let $z \sim N(\mu, \sigma^2)$. Let $A, B \in \mathbb{R}$ and $C \geq 0$, and $z \mapsto h(z) \equiv f + ez$ for*

any $e, f \in \mathbb{R}$. Then

$$\begin{aligned} & \max_x E[-\exp\{-A - Bz - Cz^2\} \exp\{-axh(z)\}] \\ &= -\frac{1}{\sqrt{2\sigma^2C + 1}} \exp\left[-A - \frac{1}{2}\left(\frac{\mu^2}{\sigma^2} - \frac{m^2}{s^2}\right)\right] \exp\left[-\frac{1}{2}\frac{\tilde{\mu}(m, s^2)^2}{\tilde{\sigma}^2(m, s^2)}\right] \end{aligned}$$

$$\text{and } \arg \max_x E[-\exp\{-A - Bz - Cz^2\} \exp\{-axh(z)\}] = \frac{\tilde{\mu}(m, s^2)}{a\tilde{\sigma}^2(m, s^2)},$$

with $m = s^2[\sigma^{-2}\mu - B]$, $s^2 = \frac{\sigma^2}{2C\sigma^2 + 1}$, $\tilde{\mu}(m, s^2) = E_{N(m, s^2)}[h(z)]$, $\tilde{\sigma}^2(m, s^2) = V_{N(m, s^2)}[h(z)]$.

For Lemma C.4, let $\beta(k) = \beta_{k+1} - r\beta_k$ for $k \in \{0, \dots, K-1\}$ and $\beta(K) = -r\beta_K$.

Lemma C.4. *Suppose $p_t = \alpha + \sum_{k=0}^K \beta_k d_{t-k}$ with $\beta_0 \neq -1$. Then the demand for risky assets of any cohort alive at time t is an affine function of past dividends, where the coefficients associated with a given dividend depend on the agent's age. That is,*

$$x_t^{t-\text{age}} = \delta(\text{age}) + \sum_{k=0}^K \delta_k(\text{age}) d_{t-k} \quad \text{for } \text{age} \in \{0, \dots, q\} \quad (47)$$

with

$$\delta(q) = \delta_k(q) = 0 \quad \forall k \in \{0, \dots, K\} \quad (48)$$

$$\delta(q-1) = \frac{\alpha(1-R)}{\gamma((1+\beta_0)\sigma)^2} \quad (49)$$

$$\delta_k(q-1) = \frac{(1+\beta_0)w(k, \lambda, q-1) + \beta(k)}{\gamma((1+\beta_0)\sigma)^2} \quad \forall k \in \{0, \dots, q-1\} \quad (50)$$

$$\delta_k(q-1) = \frac{\beta(k)}{\gamma((1+\beta_0)\sigma)^2} \quad \forall k \in \{q, \dots, K\}, \quad (51)$$

and for $\text{age} \in \{0, \dots, q-2\}$,

$$\delta(\text{age}) = \frac{\alpha(1-R) - s_{\text{age}}^2(1+\beta_0)\delta_0(\text{age}+1)\delta(\text{age}+1)(R^{q-1-(\text{age}+1)}\gamma)^2((1+\beta_0)s_{\text{age}+1})^2}{R^{q-1-(\text{age})}\gamma((1+\beta_0)s_{\text{age}})^2}, \quad (52)$$

$$\delta_k(\text{age}) = \frac{(1+\beta_0)s_{\text{age}}^2(\sigma^{-2}w(k, \lambda, \text{age}) - [(R^{q-1-(\text{age}+1)}\gamma)^2((1+\beta_0)s_{\text{age}+1})^2\delta_{k+1}(\text{age}+1)\delta_0(\text{age}+1)]) + \beta(k)}{R^{q-1-(\text{age})}\gamma((1+\beta_0)s_{\text{age}})^2} \quad (53)$$

$$\forall k \in \{0, \dots, q-1\},$$

$$\delta_k(\text{age}) = \frac{-(1+\beta_0)s_{\text{age}}^2[(R^{q-1-(\text{age}+1)}\gamma)^2((1+\beta_0)s_{\text{age}+1})^2\delta_{k+1}(\text{age}+1)\delta_0(\text{age}+1)] + \beta(k)}{R^{q-1-(\text{age})}\gamma((1+\beta_0)s_{\text{age}})^2}, \quad (54)$$

$$\forall k \in \{q, \dots, K-1\}$$

$$\delta_K(\text{age}) = \frac{\beta(K)}{R^{q-1-(\text{age})}\gamma((1+\beta_0)s_{\text{age}})^2}, \quad (55)$$

$$\text{and } s_{q-1} = \sigma \text{ and } s_{age}^2 \equiv \frac{\sigma^2}{(R^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2(\delta_0(age+1))^2\sigma^2+1}$$

We now prove Proposition 7.1. We first note that the expressions for $b_j, b_j(k)$ and c_j for $j \in \{0, \dots, q-1\}$ in the proposition are:

$$\begin{aligned} b_j &\equiv (R^{q-1-j}\gamma)^2((1+\beta_0)\sigma_j)^2\delta(j)\delta_0(j) \\ b_j(k) &\equiv \delta_k(j)\delta_0(j)(R^{q-1-j}\gamma)^2((1+\beta_0)\sigma_j)^2 \end{aligned}$$

and, $c_{q-1} = 1$ and

$$c_{j-1} = 0.5(R^{q-1-(j+1)}\gamma)(1+\beta_0)\sigma_{j+1}\delta_0(j+1)$$

for $j \in \{0, \dots, q-2\}$.

Proof of Proposition 7.1. By Lemma B.1, it follows that $x_{t+q-1}^t = \frac{E_{N(m_{q-1}, \sigma_{q-1}^2)}^{[s_{t+q}]}}{\gamma V_{N(m_{q-1}, \sigma_{q-1}^2)}^{[s_{t+q}]}}$ with $m_{q-1} = \theta_{t+q-1}^t$ and $\sigma_{q-1} = \sigma$, and that

$$V_{t+q-1}^t = -\exp\{-0.5((1+\beta_0)\sigma\gamma x_{t+q-1}^t)^2\}.$$

By Lemma C.4, x_{t+q-1}^t is affine in $d_{t+q-1-K:t_q}$ and thus $V_{t+q-1}^t = -\exp\{-A - Bd_{t+q-1} - C(d_{t+q-1})^2\}$, where A, B , and C depend on primitives and on $d_{t+q-1-K:t+q-2}$. In particular, B is affine in $d_{t+q-1-K:t+q-1-1}$ and C is constant with respect to $d_{t+q-1-K:t+q-1}$:

$$\begin{aligned} C &\equiv \frac{1}{2}\gamma^2((1+\beta_0)\sigma_{q-1})^2(\delta_0(q-1))^2 \\ B &\equiv \gamma^2((1+\beta_0)\sigma_{q-1})^2 \left(\delta(q-1) + \sum_{j=1}^K \delta_k(q-1)d_{t+q-1-j} \right) \delta_0(q-1) \\ A &\equiv \frac{1}{2}\gamma^2((1+\beta_0)\sigma_{q-1})^2 \left(\delta(q-1) + \sum_{j=1}^K \delta_k(q-1)d_{t+q-1-j} \right)^2. \end{aligned}$$

(See Lemma C.4 for the expressions for $\delta(q-1)$ and $(\delta_k(q-1))_{k=1}^K$.)

At time $t+q-2$, by equation (29),

$$x_{t+q-2}^t = \arg \max_{x \in \mathbb{R}} E_{t+q-2}^t [V_{t+q-1}^t(d_{t+q-1-K:t+q-1}) \exp(-\gamma s_{t+q-1}x)]$$

where the expectation is taken with respect $N(\theta_{t+q-2}^t, \sigma^2)$. Hence, by Lemma C.1, this problem can be cast as

$$x_{t+q-2}^t = \arg \max_{x \in \mathbb{R}} E_{N(m_{q-2}, \sigma_{q-2})} [-\exp(-R\gamma s_{t+q-1}x)],$$

where $m_{q-2} = \sigma_{q-2}(\frac{\theta_{t+q-2}^t}{\sigma^2} - B)$ and $\sigma_{q-2}^2 = \frac{\sigma^2}{2C\sigma^2+1}$. Hence, by lemma B.1

$$x_{t+q-2}^t = \frac{E_{N(m_{q-2}, \sigma_{q-2}^2)}[s_{t+q-1}]}{\gamma R V_{N(m_{q-2}, \sigma_{q-2}^2)}[s_{t+q-1}]}.$$

Also, by Lemma B.1, $V_{t+q-2}^t = -\exp\{-0.5 (V_{N(m_{q-2}, \sigma_{q-2}^2)}[s_{t+q-1}] R \gamma x_{t+q-2}^t)^2\}$. By Lemma C.4, x_{t+q-2}^t is affine and thus $V_{t+q-2}^t = -\exp\{-A - B d_{t+q-2} - C(d_{t+q-2})^2\}$, where A , B , and C depend on primitives and on $d_{t+q-2-K:t+q-3}$. In particular B is affine in $d_{t+q-2-K:t+q-3}$ and C is constant with respect to $d_{t+q-1-K:t+q-1}$:

$$\begin{aligned} C &\equiv \frac{1}{2}(R\gamma)^2((1 + \beta_0)\sigma_{q-2})^2 (\delta_0(q-2))^2 \\ B &\equiv (R\gamma)^2((1 + \beta_0)\sigma_{q-2})^2 \left(\delta(q-2) + \sum_{j=1}^K \delta_k(q-2)d_{t+q-2-j} \right) \delta_0(q-2) \\ A &\equiv \frac{1}{2}(R\gamma)^2((1 + \beta_0)\sigma_{q-2})^2 \left(\delta(q-2) + \sum_{j=1}^K \delta_k(q-2)d_{t+q-2-j} \right)^2. \end{aligned}$$

(Observe that the A , B , and C here are not the same as the previous ones; the expressions for $\delta(q-2)$ and $(\delta_k(q-2))_{k=1}^K$ can be found in the statement of Lemma C.4).

The result for $j \in \{0, \dots, q-3\}$ follows by iteration. \square

Proof of Proposition 7.2. Market Clearing and Lemma C.4 imply that, for all $k \in \{0, \dots, K\}$, $\sum_{age=0}^{q-1} \delta_k(age) = 0$ and $\sum_{age=0}^{q-1} \delta(age) = q$.

For $k = K$, it follows from equations (51) and (55)

$$\sum_{age=0}^{q-1} \delta_K(age) = \beta(K) \left(\sum_{age=0}^{q-1} \frac{1}{R^{q-1-age} \gamma ((1 + \beta_0) s_{age})^2} + \frac{1}{\gamma ((1 + \beta_0) \sigma)^2} \right)$$

therefore $\beta(K) = 0$ which implies that $\beta_K = 0$ and $\beta(K-1) = -R\beta_{K-1}$ and $\delta_K(age) = 0$ for any age .

For $k = K-1$, by equations (51) and (54), we have

$$\sum_{age=0}^{q-1} \delta_{K-1}(age) = \beta(K-1) \left(\sum_{age=0}^{q-2} \frac{1}{R^{q-1} \gamma ((1 + \beta_0) s_{age})^2} + \frac{1}{\gamma ((1 + \beta_0) \sigma)^2} \right),$$

and thus $\beta(K-1) = 0$, which implies that $\beta_{K-1} = 0$ and $\beta(K-2) = -R\beta_{K-2}$ and $\delta_{K-1}(age) = 0$ for any age .

By induction, for any $k \in \{q, \dots, K-2\}$, taking $\beta_{k+1} = 0$, it follows by equations (51)

and (54), that

$$\sum_{age=0}^{q-1} \delta_k(age) = \beta(k) \left(\sum_{age=0}^{q-2} \frac{1}{R^{q-1-age} \gamma((1+\beta_0)s_{age})^2} + \frac{1}{\gamma((1+\beta_0)\sigma)^2} \right),$$

and thus $\beta(k) = 0$ which implies $\beta_k = 0$ and $\beta(k-1) = -R\beta_{k-1}$ and $\delta_k(age) = 0$ for any $age \in \{q, \dots, K\}$. □

Proof of Proposition 7.3. Throughout the proof, let $w_0 \equiv w(0, \lambda, 0)$.

We know from Lemma OA.2.1 in the Online Appendix OA.2 that $\{\alpha, \beta_0, \beta_1\}$ solve the system of equations given by (74) and (75) and (73) in Online Appendix OA.2.

STEP 1. By equation (73),

$$2R\gamma(1+\beta_0)^2\sigma^2 = \alpha(1-R) \left[R + \frac{\sigma^2}{s^2} - \frac{[(1+\beta_0)w(0, \lambda, 0) + \beta_1 - R\beta_0]}{1+\beta_0} \right].$$

We note that $R > 1 \geq w(0, \lambda, 0)$. Thus, if $0 < \beta_1 < R\beta_0$ and $1 + \beta_0 > 0$, then

$$\left[R + \frac{\sigma^2}{s^2} - \frac{[(1+\beta_0)w(0, \lambda, 0) + \beta_1 - R\beta_0]}{1+\beta_0} \right] > 0$$

and $\alpha \leq 0$.

STEP 2. We show that if $1 + \beta_0 > 0$, then $0 < \beta_1 < R\beta_0$.

For $1 + \beta_0 > 0$, equation (75) implies $\beta_1 > 0$ and $l(1, 1) > 0$. Now assume that $\beta_1 - R\beta_0 > 0$. This implies that $l(0, 1) > 0$. For equation (74) to hold it must be that $1 - \frac{1}{R} \frac{l(1,1)}{1+\beta_0} < 0$; but

$$1 - \frac{1}{R} \frac{l(1,1)}{(1+\beta_0)^2} = 1 - \frac{1}{R} (1 - w_0) + \frac{\beta_1}{1+\beta_0} > 0 \quad (56)$$

since $R > 1$, $w_0 < 1$, and $\beta_1 > 1$. Contradiction. Then, $1 + \beta_0 > 0 \Rightarrow \beta_1 - R\beta_0 < 0$.

STEP 3. We now show that $1 + \beta_0 > 0$. Let $\phi \equiv \frac{\sigma^2}{s^2} > 1$. From equation (75), we know

$$\frac{(1+\beta_0)(1-w_0)}{\phi+R} = \beta_1.$$

We plug this into equation (74) and obtain

$$\begin{aligned} & \phi \left(-\beta_0 R + \frac{(1+\beta_0)(1-w_0)}{\phi+R} \right) + R \left[\frac{(1+\beta_0)(1-w_0)}{\phi+r} + (1+\beta_0)w_0 - \beta_0 R \right] + \dots \\ & + \left[1 + \beta_0 - \frac{\phi(1-w_0)(1+\beta_0 - \beta_0\phi R - \beta_0 R^2 + (1+\beta_0)(\phi+R-1)w_0)}{(\phi+R)^2} \right] = 0. \end{aligned}$$

Note that this is a linear equation in β_0 . Therefore,

$$\beta_0 = -\frac{2 - w_0(1 - R) - \frac{\phi(1-w_0)(1+(\phi+R-1)w_0)}{(\phi+R)^2}}{2 - w_0(1 - R) - \frac{\phi(1-w_0)(1+(\phi+R-1)w_0)}{(\phi+R)^2} - (R\phi + R^2) \left[1 - \frac{\phi(1-w_0)}{(\phi+R)^2}\right]} \equiv -\frac{A}{A - x},$$

where $A \equiv 2 - w_0(1 - R) - \frac{\phi(1-w_0)(1+(\phi+R-1)w_0)}{(\phi+R)^2}$ and $x \equiv (R\phi + R^2) \left[1 - \frac{\phi(1-w_0)}{(\phi+R)^2}\right] > 0$.

Note that for $x = 0 \Rightarrow \beta_0 = -1$. Then, it suffices to show that $\frac{\partial \beta_0}{\partial x} = \frac{A}{(A-x)^2} \geq 0$, that is, $A \geq 0$. For $w_0 = 0.5$, which corresponds to $\lambda = 0$, A is positive, i.e., $A(0.5) > 0$. In addition, $\frac{\partial A}{\partial w_0} = \frac{(\phi+R-1)(R^2+\phi(R-2(1-w_0)))}{(\phi+R)^2} > 0$ for $w_0 \geq 0.5$. Therefore, $A > 0$ for $w_0 \geq 0.5$.

If we are interested in $\lambda < 0$ cases, since $A(0) > 0$, all we need to ensure that A is positive, and thus the result holds for $w_0 \in [0, 0.5)$, is that $R \geq 2(1 - w_0)$. \square

In order to show Proposition 7.4, we need the two following lemmata. (Their proofs are relegated to the Online Appendix OA.1.2.)

Lemma C.5. For $\lambda \geq 0, 1 + \beta_0 + \beta_1 - r\beta_0 > 0$.

Lemma C.6. Given our linear guess for prices (6), when $q = 2$, at time t :

$$x_t^{t-1} = \frac{E_t^{t-1}[s_{t+1}]}{\gamma R(1 + \beta_0)\sigma^2} = \frac{\alpha(1 - R)}{\gamma(1 + \beta_0)^2\sigma^2} + \frac{l(0, 1)}{\gamma(1 + \beta_0)^2\sigma^2}d_t + \frac{l(1, 1)}{\gamma(1 + \beta_0)^2\sigma^2}d_{t-1} \quad (57)$$

$$x_t^t = \frac{E_{\Phi(m, s^2)}[s_{t+1}]}{R(1 + \beta_0)s^2} = \delta(0) + \delta_0(0)d_t + \delta_1(0)d_{t-1} \quad (58)$$

with $l(0, 1) \equiv [(1 + \beta_0)w(0, \lambda, 0) + \beta_1 - R\beta_0]$ and $l(1, 1) \equiv [(1 + \beta_0)w(1, \lambda, 0) - R\beta_1]$, and

$$\delta(0) = \frac{\alpha(1-R)\left[1 - \frac{s^2}{\sigma^2} \frac{l(0,1)}{(1+\beta_0)}\right]}{\gamma R(1+\beta_0)^2 s^2}, \quad \delta_0(0) = \frac{\beta_1 - R\beta_0 + (1+\beta_0)\frac{s^2}{\sigma^2} \left(1 - \frac{l(1,1)l(0,1)}{(1+\beta_0)^2}\right)}{\gamma R(1+\beta_0)^2 s^2}, \quad \text{and} \quad \delta_1(0) = -\frac{R\beta_1}{R\gamma(1+\beta_0)s^2}.$$

Proof of Proposition 7.4. By Lemma C.6 and market clearing, it follows that

$$\delta_0(0) + \frac{l(0, 1)}{\gamma(1 + \beta_0)^2\sigma^2} = 0,$$

and

$$\delta_1(0) + \frac{l(1, 1)}{\gamma(1 + \beta_0)^2\sigma^2} = 0.$$

These expressions and Lemma C.6 imply that $\frac{\partial x_t^{t-1}}{\partial d_t} = \frac{l(0,1)}{\gamma(1+\beta_0)^2\sigma^2} = -\frac{\partial x_t^t}{\partial d_t}$, and $\frac{\partial x_t^t}{\partial d_{t-1}} = \delta_1(0) = -\frac{\partial x_t^t}{\partial d_{t-1}}$. Therefore, it suffices to show that $l(0, 1) < 0$ and $\delta_1(0) < 0$.

By Proposition 7.3, $\beta_1 > 0$ and $\beta_0 > 0$ and thus $\delta_1(0) = -\frac{R\beta_1}{R\gamma(1+\beta_0)s^2} < 0$. So it only remains

to show that $l(0, 1) < 0$. To show this, note that from the equilibrium condition (74) we have

$$0 = \left[R - \frac{l(1, 1)}{(1 + \beta_0)} \right] l(0, 1) + \frac{l(0, 1)^2}{(1 + \beta_0)^2} (\beta_1 - R\beta_0) + [1 + \beta_0 + \beta_1 - R\beta_0].$$

From Lemma C.5, $1 + \beta_0 + \beta_1 - R\beta_0 > 0$. Let $x = \frac{l(0, 1)}{1 + \beta_0}$, then

$$0 = [R(1 + \beta_0) - l(1, 1)]x + x^2(\beta_1 - R\beta_0) + [1 + \beta_0 + \beta_1 - R\beta_0],$$

or equivalently $F(x) \equiv ax^2 + bx + c$, with $a = \beta_1 - R\beta_0 < 0$ (by Proposition 7.3), $b = R(1 + \beta_0) - l(1, 1) = R(1 + \beta_0) - (1 + \beta_0)w(1, \lambda, 0) + R\beta_1 > 0$ (by Proposition 7.3) and $c = 1 + \beta_0 + \beta_1 - R\beta_0 > 0$ (by Lemma C.5). Thus, F is concave and $F(0) = c > 0$. By definition of F , $l(0, 1)/(1 + \beta_0)$ must be a root of F . Let $x^* = \arg \max_{x \in \mathbb{R}} F(x)$, which is given by $x^* = -\frac{b}{2a} > 0$. Therefore, $F(\cdot)$ has two roots x_1, x_2 with $x_1 < 0 < x^* < x_2$.

We now show that $x_2 = \frac{l(0, 1)}{1 + \beta_0}$ cannot hold. Suppose not, that is, assume that our solution is the positive root $\frac{l(0, 1)}{1 + \beta_0} = x_2$. Then, since $x^* < x_2$ and $a < 0$, we have $\frac{b}{2} < -a\frac{l(0, 1)}{1 + \beta_0}$, or equivalently, $\frac{R(1 + \beta_0) - l(1, 1)}{2} < l(0, 1)\frac{R\beta_0 - \beta_1}{1 + \beta_0}$.

Let $Z \equiv -\frac{\beta_1 - R\beta_0}{1 + \beta_0}$. Then the last inequality implies that $R(1 + \beta_0) - (1 + \beta_0)(1 - w_0) + R\beta_1 < 2l(0, 1)Z$. (Recall that $w_0 \equiv w(0, \lambda, 0)$ and $w(1, \lambda, 0) = 1 - w_0$.) By replacing $l(0, 1)$ and some algebra, it follows that $R(1 + \beta_0) - (1 + \beta_0)(1 - w_0) + R\beta_1 < 2Z[(1 + \beta_0)w_0 + \beta_1 - R\beta_0]$, or equivalently $\frac{1}{2}w_0 + \frac{1}{2}\left[R - 1 + R\frac{\beta_1}{1 + \beta_0}\right] < Z(w_0 - Z)$. We note that the RHS is bounded by $w_0/4$ since no matter the value of Z , the function $z \mapsto z(w_0 - z)$ cannot be larger than $(w_0)^2/4 < w_0/4$ (since $w_0 \in [0, 1]$). Therefore,

$$\frac{w_0}{4} > \frac{w_0}{2} + \frac{1}{2}\left[R - 1 + R\frac{\beta_1}{1 + \beta_0}\right].$$

However, $\frac{1}{2}\left[R - 1 + R\frac{\beta_1}{1 + \beta_0}\right] > 0$; thus a contradiction follows. The solution must be the negative root. \square

Online Appendix

Appendix OA.1 Proofs of Supplementary Lemmas

OA.1.1 Proofs of Supplementary Lemmas in the Appendix B

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

$$\begin{aligned} x^* &= \arg \max_x - \exp \left(-axE[z] + \frac{1}{2}a^2x^2V[z] \right) \\ &= \arg \max_x ax\mu - \frac{1}{2}a^2x^2\sigma^2 \end{aligned}$$

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 Lemma B.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

OA.1.2 Proof of Supplementary Lemmas in the Appendix C

Proof of Lemma C.1. Let $\varphi(z) \equiv K \exp\{-(A + Bz + Cz^2)\}\phi(z; \mu, \sigma^2)$. By definition of K , $\int \varphi(z)dz = 1$ and $\varphi \geq 0$, so it is a pdf. Moreover,

$$\begin{aligned} \varphi(z) &= \frac{K^{-1}}{\sqrt{2\pi\sigma}} \exp\{-A - Bz - Cz^2 - 0.5\sigma^{-2}(z - \mu)^2\} \\ &= \frac{1}{K\sqrt{2\pi\sigma}} \exp\{-z^2(C + 0.5\sigma^{-2}) - 2z(0.5B - 0.5\sigma^{-2}\mu) - (A + 0.5\sigma^{-2}\mu^2)\} \\ &= \frac{1}{K\sqrt{2\pi\sigma}} \exp\{-(A + 0.5\sigma^{-2}\mu^2)\} \exp\{-0.5(2C + \sigma^{-2}) \left(z^2 - 2z \frac{(-B + \sigma^{-2}\mu)}{(2C + \sigma^{-2})} \right)\}. \end{aligned}$$

Let $\Sigma^2 \equiv (2C + \sigma^{-2})^{-1}$, $m \equiv \Sigma^2(\sigma^{-2}\mu - b)$, and $K = \frac{1}{\sqrt{2\sigma^2C+1}} \exp\{-(A + 0.5\sigma^{-2}\mu^2) + \frac{m^2}{2\Sigma^2}\}$:

$$\begin{aligned} \varphi(z) &= \frac{1}{K\sqrt{2\pi\sigma}} \exp\left\{-\left(a + 0.5\frac{\mu^2}{\sigma^2}\right) + \frac{m^2}{2\Sigma^2}\right\} \exp\left\{-\frac{z^2 - 2zm + m^2}{2\Sigma^2}\right\} \\ &= \frac{1}{K\sqrt{2\pi\sigma}} \exp\left\{-\left(a + 0.5\sigma^{-2}\mu^2\right) + \frac{m^2}{2\Sigma^2}\right\} \exp\left\{-\frac{(z - m)^2}{2\Sigma^2}\right\} = \frac{1}{\sqrt{2\pi\Sigma}} \exp\left\{-\frac{(z - m)^2}{2\Sigma^2}\right\} \\ &= \frac{1}{\sqrt{2\pi\Sigma^2}} \exp\left\{-\frac{(z - m)^2}{2\Sigma^2}\right\} \end{aligned}$$

□

Proof of Lemma C.2. At time $n + q$, an agent born in n is in the last period of his life, consuming all of its wealth. Therefore, he will sell all of his claims to the assets he holds and consume. The gain from saving is zero, and therefore the holding of financial assets is also zero by the end of this period: $x_{n+q}^n = 0, a_{n+q}^n = 0$. Given this, we can compute the portfolio choice of an agent of age $q - 1$, who will want to save for next period when all wealth will be consumed. The agent's problem is a standard static portfolio problem, with initial wealth $W_{n_q}^n$:

$$\max_x E_{n+q-1}^t [-\exp(-\gamma(W_{n+q-1}^n + xs_{n+q}))] = \max_x E_{n+q-1}^n [-\exp(-\gamma xs_{n+q})] \quad (59)$$

At time $n + q - 1$, the only random variable is d_{n+q} , which is normally distributed, and thus $s_{n+q} \sim N(E_{n+q-1}^n[s_{n+q}]; (1 + \beta_0)\sigma^2)$. Given this, the agent's problem becomes:

$$\max_x \left[-\exp\left(-\gamma x E_n^{n+q-1}[s_{n+q}] + \frac{1}{2}\gamma^2 x^2 (1 + \beta_0)\sigma^2\right) \right] \quad (60)$$

$$\iff \max_x x E_n^{n+q-1}[s_{n+q}] - \frac{1}{2}\gamma x^2 (1 + \beta_0)^2 \sigma^2. \quad (61)$$

And therefore, by FOC:

$$x_{n+q-1}^n = \frac{E_{n+q-1}^n[s_{n+q}]}{\gamma\sigma^2(1 + \beta_0)^2}. \quad (62)$$

□

Proof of Lemma C.3. Note that $E[-\exp\{-A - Bz - Cz^2\} \exp\{-axh(z)\}]$ can be written as:

$$\int \exp\{-axh(z)\} - \exp\{-A - Bz - Cz^2\} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\frac{z - \mu}{\sigma^2}\right\} dz$$

By Lemma C.1, we know that this can be re-written as:

$$\frac{1}{\sqrt{2\sigma^2 C + 1}} \exp\left\{-A - 0.5\left(\frac{\mu^2}{\sigma^2} - \frac{m^2}{s^2}\right)\right\} \int -\exp\{-axh(z)\} \Phi(m, s^2) dz$$

with $m = -s^2 B + s\sigma^{-2}\mu$ and $s^2 = \frac{\sigma^2}{2C\sigma^2 + 1}$. Therefore, the maximization problem becomes:

$$\max_x E_{N(m, s^2)}[-\exp\{-axh(z)\}]$$

with $E_{N(m, s^2)}[\cdot]$ being the expectations operator over $z \sim N(m, s^2)$. Since $h(z)$ is linear, we know that $h(z) \sim N(\tilde{\mu}(m, s^2), \tilde{\sigma}(m, s^2)^2)$, with $\tilde{\mu}(m, s^2) = E_{N(m, s^2)}[h(z)]$,

$\tilde{\sigma}(m, s^2)^2 = V_{N(m, s^2)}[h(z)]$, by Lemma B.1, we know that

$$\begin{aligned} \arg \max_x E[-\exp\{-A - Bz - Cz^2\} \exp\{-axh(z)\}] &= \frac{\tilde{\mu}(m, s^2)}{a\tilde{\sigma}(m, s^2)^2} \\ \max_x E[-\exp\{-A - Bz - Cz^2\} \exp\{-axh(z)\}] &= -\frac{1}{\sqrt{2\sigma^2C + 1}} \exp\left[-A - 0.5\left(\frac{\mu^2}{\sigma^2} - \frac{m^2}{s^2}\right)\right] \\ &\quad \times \exp\left[-0.5\frac{\tilde{\mu}(m, s^2)^2}{\tilde{\sigma}(m, s^2)^2}\right] \end{aligned}$$

□

Let $t \mapsto \rho(t) \equiv \gamma t^2$ and let

$$\begin{aligned} \Lambda(d_{t-K}, \dots, d_t) &\equiv \alpha(1 - R) + \sum_{k=1}^K \beta_k d_{t+1-k} - R \sum_{k=0}^K \beta_k d_{t-k} \\ &= \alpha(1 - R) + \sum_{j=0}^{K-1} \beta_{j+1} d_{t-j} - R \sum_{k=0}^K \beta_k d_{t-k} = \alpha(1 - R) + \sum_{k=0}^K \beta(k) d_{t-k} \end{aligned} \quad (63)$$

with $\beta(k) = \beta_{k+1} - R\beta_k$ for $k \in \{0, \dots, K-1\}$ and $\beta(K) = -R\beta_K$. We use Λ_τ to denote $\Lambda(d_{\tau-K}, \dots, d_\tau)$.

Proof of Lemma C.4. We divide the proof into several steps.

STEP 1 It is straightforward that demand for risky assets can only be positive for a generation that is alive. From Lemma C.2, we know that $x_t^{t-q} = 0$ and that $x_t^{t-q+1} = \frac{E_t^{t-q+1}[s_{t+1}]}{\gamma((1+\beta_0)\sigma)^2}$. Therefore,

$$\begin{aligned} \delta(q) &= \delta_k(q) = 0, \quad \forall k \in \{0, \dots, K\} \\ \delta(q-1) &= \frac{\alpha(1-R)}{\gamma((1+\beta_0)\sigma)^2}, \quad \delta_k(q-1) = \frac{(1+\beta_0)w(k, \lambda, q-1) + \beta(k)}{\gamma((1+\beta_0)\sigma)^2}, \quad \forall k \in \{0, \dots, q-1\} \\ \delta_k(q-1) &= \frac{\beta(k)}{\gamma((1+\beta_0)\sigma)^2}, \quad \forall k \in \{q, \dots, K\}. \end{aligned}$$

We also know from Lemma B.1 that

$$V^{q-1}(d_{t-K}, \dots, d_t) = -\exp\left(-\frac{1}{2}\left(d_t \delta_0(q-1) + \delta(q-1) + \sum_{j=1}^K \delta_k(q-1) d_{t-j}\right)^2 \gamma^2((1+\beta_0)s_{q-1})^2\right)$$

where $s_{q-1} = \sigma^2$. Henceforth, we denote $V^{q-1}(d_{t-K}, \dots, d_t)$ by V_t^{t-q+1} . In particular, $V_{t+1}^{t+1-q+1} = V_{t+1}^{t-q+2} = V^{q-1}(d_{t+1-K}, \dots, d_{t+1})$.

STEP 2. We now derive the risky demand and continuation value for generation aged

$q - 2$. The problem of generation aged $q - 2$ at time t is given by,

$$\max_x E_t^{t-q+2} [V_{t+1}^{t-q+2} \exp(-\gamma R x s_{t+1})]. \quad (64)$$

By the calculations in step 1, and using Λ_t as defined in (63), this problem becomes:

$$\begin{aligned} & V^{q-2}(d_{t-K}, \dots, d_t) \quad (65) \\ & = \max_x E_t^{t-q+2} \left[-\exp \left(-\frac{1}{2} (x_t^{q-1})^2 \gamma^2 ((1 + \beta_0) s_{q-1})^2 - \gamma R x ((1 + \beta_0) d_{t+1} + \Lambda_t) \right) \right]. \quad (66) \end{aligned}$$

with $x_t^{q-1} = d_{t+1} \delta_0 (q - 1) + \delta (q - 1) + \sum_{j=1}^K \delta_k (q - 1) d_{t+1-j}$.

Observe that

$$\begin{aligned} & -\frac{1}{2} \left(d_{t+1} \delta_0 (q - 1) + \delta (q - 1) + \sum_{j=1}^K \delta_k (q - 1) d_{t+1-j} \right)^2 \gamma^2 ((1 + \beta_0) s_{q-1})^2 \\ & = -\frac{1}{2} \gamma^2 ((1 + \beta_0) s_{q-1})^2 \left(\delta (q - 1) + \sum_{j=1}^K \delta_k (q - 1) d_{t+1-j} \right)^2 \\ & \quad - \gamma^2 ((1 + \beta_0) s_{q-1})^2 \left(\delta (q - 1) + \sum_{j=1}^K \delta_k (q - 1) d_{t+1-j} \right) \delta_0 (q - 1) d_{t+1} \\ & \quad - \frac{1}{2} \gamma^2 ((1 + \beta_0) s_{q-1})^2 (\delta_0 (q - 1))^2 d_{t+1}^2, \end{aligned}$$

and that future dividends are the only random variable, with $d_{t+1} \sim N(\theta_t^{t-q+2}, \sigma^2)$. Therefore, by Lemma C.3, and with:

$$\begin{aligned} A &= \frac{1}{2} \gamma^2 ((1 + \beta_0) s_{q-1})^2 \left(\delta (q - 1) + \sum_{j=1}^K \delta_k (q - 1) d_{t+1-j} \right)^2 \\ B &= \gamma^2 ((1 + \beta_0) s_{q-1})^2 \left(\delta (q - 1) + \sum_{j=1}^K \delta_k (q - 1) d_{t+1-j} \right) \delta_0 (q - 1) \\ C &= \frac{1}{2} \gamma^2 ((1 + \beta_0) s_{q-1})^2 (\delta_0 (q - 1))^2 \end{aligned}$$

we obtain:

$$x_t^{t-(q-2)} = \frac{(1 + \beta_0) s_{q-2}^2 (\sigma^{-2} \theta_t^{t-(q-2)} - B) + \Lambda_t}{R \gamma ((1 + \beta_0) s_{q-2})^2}$$

with $s_{q-2}^2 \equiv \frac{\sigma^2}{\gamma^2((1+\beta_0)s_{q-1})^2(\delta_0(q-1))^2\sigma^2+1}$. Therefore,

$$\delta(q-2) = \frac{\alpha(1-R) - s_{q-2}^2(1+\beta_0)\delta_0(q-1)\delta(q-1)\gamma^2((1+\beta_0)s_{q-1})^2}{R\gamma((1+\beta_0)s_{q-2})^2}$$

For $k \in \{0, \dots, q-1\}$:

$$\delta_k(q-2) = \frac{(1+\beta_0)s_{q-2}^2(\sigma^{-2}w(k, \lambda, q-2) - [\gamma^2((1+\beta_0)s_{q-1})^2\delta_{k+1}(q-1)\delta_0(q-1)]) + \beta(k)}{R\gamma((1+\beta_0)s_{q-2})^2},$$

$$\delta_k(q-2) = \frac{-(1+\beta_0)s_{q-2}^2[\gamma^2((1+\beta_0)s_{q-1})^2\delta_{k+1}(q-1)\delta_0(q-1)] + \beta(k)}{R\gamma((1+\beta_0)s_{q-2})^2},$$

$$\text{and } \delta_K(q-2) = \frac{\beta(K)}{R\gamma((1+\beta_0)s_{q-2})^2}.$$

By lemma C.1, $d_{t+1} \sim N(m_t, s_{q-2}^2)$ with $m_t \equiv -s_{q-2}^2B + s_{q-2}^2\sigma^{-2}\theta_t^{t-q+2}$. Thus, invoking lemma B.1 for this distribution for dividends and $a = R\gamma(1+\beta_0)$ implies that

$$\begin{aligned} V^{q-2}(d_{t-K}, \dots, d_t) &\asymp -\exp\left(-\frac{1}{2}\left(x_t^{t-(q-2)}\right)^2 (R\gamma)^2((1+\beta_0)s_{q-2})^2\right) \\ &= -\exp\left(-\frac{1}{2}\left(d_t\delta_0(q-2) + \delta(q-2) + \sum_{j=1}^K \delta_k(q-2)d_{t-j}\right)^2 (R\gamma)^2((1+\beta_0)s_{q-2})^2\right) \end{aligned}$$

(the symbol \asymp means that equality holds up to a positive constant).

STEP 3. We now consider the problem for agents of age $age \leq q-3$. Suppose the problem at age $age+1$ is solved, that is, suppose

$$\begin{aligned} V_{t+1}^{t-age-1} &= V^{age+1}(d_{t+1-K}, \dots, d_{t+1}) \\ &\asymp -\exp\left\{-\frac{1}{2}\left(d_{t+1}\delta_0(age+1) + \delta(age+1) + \sum_{j=1}^K \delta_j(age+1)d_{t+1-j}\right)^2 (R^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2\right\}. \end{aligned}$$

The maximization problem is given by:

$$V^{age}(d_{t-K}, \dots, d_t) \equiv \max_x E_t^{t-age} [V_{t+1}^{t-age-1} \exp(-\gamma R^{q-1-age} x((1+\beta_0)d_{t+1} + \Lambda_t))]. \quad (67)$$

By similar calculations to step 2 and Lemma C.3,

$$x_t^{t-age} = \frac{(1+\beta_0)s_{age}^2(\sigma^{-2}\theta_t^{t-age} - B) + \Lambda_t}{R^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2}$$

with $s_{age}^2 \equiv \frac{\sigma^2}{(R^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2(\delta_0(age+1))^2\sigma^2+1}$, and

$$B \equiv (R^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2 \left(\delta(age+1) + \sum_{j=1}^K \delta_j(age+1)d_{t+1-j} \right) \delta_0(age+1).$$

Therefore

$$\begin{aligned} \delta(age) &= \frac{\alpha(1-R) - s_{age}^2(1+\beta_0)\delta_0(age+1)\delta(age+1)(R^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2}{R^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2}, \\ \delta_k(age) &= \frac{(1+\beta_0)s_{age}^2(\sigma^{-2}w(k, \lambda, age) - [(R^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2\delta_{k+1}(age+1)\delta_0(age+1)]) + \beta(k)}{R^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2} \\ &\quad k \in \{0, \dots, q-1\}, \\ \delta_k(age) &= \frac{-(1+\beta_0)s_{age}^2[(R^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2\delta_{k+1}(age+1)\delta_0(age+1)] + \beta(k)}{R^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2}, \quad k \in \{q, \dots, K-1\} \\ \delta_K(age) &= \frac{\beta(K)}{R^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2}. \end{aligned}$$

By lemma C.1, $d_{t+1} \sim N(m_t, s_{age}^2)$ with $m_t \equiv -s_{age}^2 B + s_{age}^2 \sigma^{-2} \theta_t^{t-q+2}$. Thus, invoking lemma B.1 for this distribution for dividends and $a = R^{q-1-age}\gamma(1+\beta_0)$ implies that

$$\begin{aligned} V^{age}(d_{t-K}, \dots, d_t) &\asymp -\exp\left(-\frac{1}{2}\left(x_t^{t-(age)}\right)^2 (R^{q-1-(age)}\gamma)^2((1+\beta_0)s_{age})^2\right) \\ &= -\exp\left(-\frac{1}{2}\left(d_t\delta_0(age) + \delta(age) + \sum_{j=1}^K \delta_k(age)d_{t-j}\right)^2 (R^{q-1-(age)}\gamma)^2((1+\beta_0)s_{age})^2\right). \end{aligned}$$

□

Proof of Lemma C.5. Assume it is not the case, i.e. $1 + \beta_0 + \beta_1 - R\beta_0 \leq 0$. This implies that $l(0, 1) = (1 + \beta_0)w_0 + \beta_1 - R\beta_0 \leq 0$. From condition (74) we have:

$$0 = \left[R - \frac{l(1, 1)}{(1 + \beta_0)} \right] l(0, 1) + \frac{l(0, 1)^2}{(1 + \beta_0)^2} (\beta_1 - R\beta_0) + [1 + \beta_0 + \beta_1 - R\beta_0]$$

Then, since $\beta_1 - R\beta_0 \leq 0$ by proposition 7.3, for the previous equation to hold it must be that $\left[R - \frac{l(1, 1)}{(1 + \beta_0)} \right] \leq 0$. However, replacing $l(1, 1)$, this inequality implies that

$$\left[R - \frac{(1 + \beta_0)(1 - w_0) - R\beta_1}{(1 + \beta_0)} \right] = \left[R + \frac{R\beta_1}{1 + \beta_0} - (1 - w_0) \right] > 0.$$

By Proposition 7.3, $\frac{R\beta_1}{1+\beta_0} > 0$ and $R > 1$ by assumption, so we obtained a contradiction. Hence, it must be that $[1 + \beta_0 + \beta_1 - R\beta_0] > 0$. □

Proof of Lemma C.6. From Lemma B.1, we know that $x_t^{t-1} = \frac{E_t^{t-1}[s_{t+1}]}{\gamma(1+\beta_0)\sigma^2}$. Therefore, given

our guess for prices and Lemma 7.2, we have:

$$x_t^{t-1} = \frac{E_t^{t-1}[d_{t+1} + p_{t+1} - p_t R]}{\gamma(1 + \beta_0)\sigma^2} \quad (68)$$

$$= \frac{(1 + \beta_0)\theta_t^{t-1} + \alpha(1 - R) + (\beta_1 - R\beta_0)d_t - R\beta_1 d_{t-1}}{\gamma(1 + \beta_0)\sigma^2} \quad (69)$$

since $\theta_t^{t-1} = w_0 d_t + (1 - w_0)d_{t-1}$, we obtain equation (57), where $l(0, 1) = (1 + \beta_0)w_0 + \beta_1 - R\beta_0$ and $l(1, 1) = (1 + \beta_0)(1 - w_0) - R\beta_1$. We also know from Lemma C.2 that

$$\begin{aligned} V_t^{t-1} &= -\exp\left(-\frac{1}{2} \frac{E_t^{t-q+1}[s_{t+1}]^2}{\gamma(1 + \beta_0)\sigma^2}\right) \\ &= -\exp\left(-\frac{1}{2} \frac{(\alpha(1 - R) + l(1, 1)d_{t-1} + l(0, 1)d_t)^2}{\gamma(1 + \beta_0)\sigma^2}\right) \\ &= -\exp\left(-\frac{1}{2} \frac{(L_t(1, 1) + l(0, 1)d_t)^2}{\gamma(1 + \beta_0)\sigma^2}\right) \end{aligned}$$

where $L_t(1, 1) \equiv \alpha(1 - R) + l(1, 1)d_{t-1}$. Thus, we can write the value function of the generation that is investing for the last time on the market as follows:

$$V_t^{t-1} = -\exp(-A_t - B_t d_t - C d_t^2) \quad (70)$$

where $A_t \equiv \frac{L_t(1, 1)^2}{2\gamma(1 + \beta_0)^2\sigma^2}$, $B_t \equiv \frac{L_t(1, 1)l(0, 1)}{\gamma(1 + \beta_0)^2\sigma^2}$, $C \equiv \frac{l(0, 1)^2}{2\gamma(1 + \beta_0)^2\sigma^2}$. Using these results to obtain V_{t+1}^t , the problem of the young generation at time t is given by

$$\max_x E_t^t [V_{t+1}^t \exp(-\gamma R x s_{t+1})]. \quad (71)$$

From Lemma C.3

$$x_t^t = \frac{\tilde{\mu}(m, s^2)}{\gamma R \tilde{\sigma}(m, s^2)^2},$$

where

$$\begin{aligned} \tilde{\mu}(m, s^2) &= E_{N(m, s^2)}[h(z)] = \alpha(1 - R) + (\beta_1 - R\beta_0)d_t - R\beta_1 d_{t-1} + (1 + \beta_0)m \\ \tilde{\sigma}(m, s^2)^2 &= V_{N(m, s^2)}[h(d_{t+1})] = (1 + \beta_0)^2 s^2 \end{aligned}$$

with $m = \frac{\theta_t^t - \sigma^2 B_{t+1}}{2C\sigma^2 + 1}$, $s^2 = \frac{\sigma^2}{2C\sigma^2 + 1}$. Incorporating the fact that $B_{t+1} = \frac{(\alpha(R-1) + l(1, 1)d_t)l(0, 1)}{(1 + \beta_0)^2\sigma^2}$ and $\theta_t^t = d_t$ we obtain equation (58) and the respective δ 's. □

Appendix OA.2 Proposition 7.2 for the $q = 2$ case

The next lemma specializes the results in Proposition 7.2 for the $q = 2$ case. It helps illustrate the expressions needed to compute α , β_0 and β_1 .

Lemma OA.2.1. *For $R > 1$ in any linear equilibrium prices are given by:*

$$p_t = \alpha + \beta_0 d_t + \beta_1 d_{t-1} \quad \forall t \in \mathbb{Z} \quad (72)$$

where the coefficients $\{\alpha, \beta_0, \beta_1\}$ are uniquely determined by the following set of non-linear equations:

$$0 = \alpha(1 - R) \left[R + \frac{\sigma^2}{s^2} - \frac{l(0, 1)}{1 + \beta_0} \right] - 2R\gamma(1 + \beta_0)^2 \sigma^2 \quad (73)$$

$$0 = l(0, 1) + \frac{1}{R} \frac{\sigma^2}{s^2} (\beta_1 - R\beta_0) + \frac{1}{R} (1 + \beta_0) \left(1 - \frac{l(1, 1)l(0, 1)}{(1 + \beta_0)^2} \right) \quad (74)$$

$$0 = l(1, 1) - \frac{\sigma^2}{s^2} \beta_1 \quad (75)$$

where $l(0, 1) \equiv [(1 + \beta_0)w(0, \lambda, 0) + \beta_1 - R\beta_0]$ and $l(1, 1) \equiv [(1 + \beta_0)w(1, \lambda, 0) - R\beta_1]$.

Proof of Lemma OA.2.1. By Proposition 7.1, we have the following demands:

$$x_t^{t-2} = 0 \quad (76)$$

$$x_t^{t-1} = \frac{E_t^{t-1}[s_{t+1}]}{\gamma R(1 + \beta_0)\sigma^2} = \frac{\alpha(1 - R) + l(0, 1)d_t + l(1, 1)d_{t-1}}{\gamma(1 + \beta_0)^2 \sigma^2} \quad (77)$$

$$x_t^t = \frac{E_{\Phi(m, s^2)}[s_{t+1}]}{\gamma R(1 + \beta_0)s^2} = \frac{\alpha(1 - R) + (\beta_1 - R\beta_0)d_t - R\beta_1 d_{t-1} + (1 + \beta_0)m}{\gamma R(1 + \beta_0)^2 s^2} \quad (78)$$

where $l(0, 1) \equiv (1 + \beta_0)w(0, \lambda, 0) + \beta_1 - R\beta_0$, $l(1, 1) \equiv (1 + \beta_0)w(1, \lambda, 0) - R\beta_1$,

$$m = \frac{s^2}{\sigma^2} [d_t - \sigma^2 B_{t+1}(1)]$$

$$s^2 = \frac{\sigma^2}{2C(1)\sigma^2 + 1},$$

and

$$B_{t+1}(1) = \frac{\alpha(1 - R)l(0, 1)}{(1 + \beta_0)^2 \sigma^2} + \frac{l(1, 1)l(0, 1)}{(1 + \beta_0)^2 \sigma^2} d_t$$

$$C(1) = \frac{l(0, 1)^2}{(1 + \beta_0)^2 \sigma^2}$$

Therefore:

$$m = \frac{s^2}{\sigma^2} \left[d_t - \frac{\alpha(1-R)l(0,1)}{(1+\beta_0)^2} - \frac{l(1,1)l(0,1)}{(1+\beta_0)^2} d_t \right] = \frac{s^2}{\sigma^2} \left[-\frac{\alpha(1-R)l(0,1)}{(1+\beta_0)^2} + \left(1 - \frac{l(1,1)l(0,1)}{(1+\beta_0)^2} \right) d_t \right]$$

$$s^2 = \frac{\sigma^2}{2\frac{l(0,1)^2}{(1+\beta_0)^2\sigma^2} + 1} = \frac{(1+\beta_0)^2}{l(0,1)^2 + (1+\beta_0)^2} \sigma^2.$$

Plugging this in the expression for x_t^t , it follows that

$$x_t^t = \frac{\alpha(1-R) + (\beta_1 - R\beta_0) d_t - R\beta_1 d_{t-1} + (1+\beta_0) \frac{s^2}{\sigma^2} \left[-\frac{\alpha(1-R)l(0,1)}{(1+\beta_0)^2} + \left(1 - \frac{l(1,1)l(0,1)}{(1+\beta_0)^2} \right) d_t \right]}{\gamma R (1+\beta_0)^2 s^2}$$

$$= \frac{\alpha(1-R) \left[1 - \frac{s^2}{\sigma^2} \frac{l(0,1)}{(1+\beta_0)} \right] + \left[\beta_1 - R\beta_0 + (1+\beta_0) \frac{s^2}{\sigma^2} \left(1 - \frac{l(1,1)l(0,1)}{(1+\beta_0)^2} \right) \right] d_t - R\beta_1 d_{t-1}}{\gamma R (1+\beta_0)^2 s^2}.$$

By Market clearing:

$$1 = \frac{1}{2} \left(\frac{\alpha(1-R) + l(0,1) d_t + l(1,1) d_{t-1}}{\gamma (1+\beta_0)^2 \sigma^2} \right)$$

$$+ \frac{1}{2} \left(\frac{\alpha(1-R) \left[1 - \frac{s^2}{\sigma^2} \frac{l(0,1)}{(1+\beta_0)} \right] + \left[\beta_1 - R\beta_0 + \frac{s^2}{\sigma^2} (1+\beta_0) \left(1 - \frac{l(1,1)l(0,1)}{(1+\beta_0)^2} \right) \right] d_t - R\beta_1 d_{t-1}}{\gamma R (1+\beta_0)^2 s^2} \right)$$

$$= \frac{1}{2} \left(\frac{\alpha(1-R) + l(0,1) d_t + l(1,1) d_{t-1}}{\gamma (1+\beta_0)^2 \sigma^2} \right)$$

$$+ \frac{1}{2} \left(\frac{\alpha(1-R) \frac{\sigma^2}{s^2} \left[1 - \frac{s^2}{\sigma^2} \frac{l(0,1)}{(1+\beta_0)} \right] + \left[\frac{\sigma^2}{s^2} (\beta_1 - R\beta_0) + (1+\beta_0) \left(1 - \frac{l(1,1)l(0,1)}{(1+\beta_0)^2} \right) \right] d_t - \frac{\sigma^2}{s^2} R\beta_1 d_{t-1}}{\gamma R (1+\beta_0)^2 \sigma^2} \right),$$

which implies

$$2\gamma(1+\beta_0)^2 \sigma^2 = (\alpha(1-R) + l(0,1) d_t + l(1,1) d_{t-1})$$

$$+ \frac{1}{R} \left[\alpha(1-R) \frac{\sigma^2}{s^2} \left[1 - \frac{s^2}{\sigma^2} \frac{l(0,1)}{(1+\beta_0)} \right] \right]$$

$$+ \frac{1}{R} \left[\left[\frac{\sigma^2}{s^2} (\beta_1 - R\beta_0) + (1+\beta_0) \left(1 - \frac{l(1,1)l(0,1)}{(1+\beta_0)^2} \right) \right] d_t - \frac{\sigma^2}{s^2} R\beta_1 d_{t-1} \right]$$

$$= \alpha(1-R) \frac{1}{R} \left[R + \frac{\sigma^2}{s^2} - \frac{l(0,1)}{(1+\beta_0)} \right]$$

$$+ \left[l(0,1) + \frac{1}{R} \frac{\sigma^2}{s^2} (\beta_1 - R\beta_0) + \frac{1}{R} (1+\beta_0) \left(1 - \frac{l(1,1)l(0,1)}{(1+\beta_0)^2} \right) \right] d_t + \left[l(1,1) - \frac{\sigma^2}{s^2} \beta_1 \right] d_{t-1}.$$

Therefore $\{\alpha, \beta_0, \beta_1\}$ solve the following system of equations:

$$0 = \alpha (1 - R) \left[R + \frac{\sigma^2}{s^2} - \frac{l(0, 1)}{1 + \beta_0} \right] - 2R\gamma (1 + \beta_0)^2 \sigma^2 \quad (79)$$

$$0 = l(0, 1) + \frac{1}{R} \frac{\sigma^2}{s^2} (\beta_1 - R\beta_0) + \frac{1}{R} (1 + \beta_0) \left(1 - \frac{l(1, 1)l(0, 1)}{(1 + \beta_0)^2} \right) \quad (80)$$

$$0 = l(1, 1) - \frac{\sigma^2}{s^2} \beta_1 \quad (81)$$

where $l(0, 1) \equiv [(1 + \beta_0)w(0, \lambda, 0) + \beta_1 - R\beta_0]$ and $l(1, 1) \equiv [(1 + \beta_0)w(1, \lambda, 0) - R\beta_1]$. \square

Appendix OA.3 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 Online 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:

$$\begin{aligned} 1 &= y_t \frac{E_t^t [p_{t+1} + d_{t+1}] - Rp_t}{\gamma V [p_{t+1} + d_{t+1}]} + y_{t-1} \frac{E_t^{t-1} [p_{t+1} + d_{t+1}] - Rp_t}{\gamma V [p_{t+1} + d_{t+1}]} \iff \\ 1 &= \frac{y_0 (1 + g)^{t-1}}{\gamma (1 + \beta_0)^2 \sigma^2} \left[(1 + \beta_0) [(1 + g) E_t^t [d_{t+1}] + E_t^{t-1} [d_{t+1}]] + (2 + g) [\alpha_0 (1 + g)^{-(t+1)} + \beta_1 d_t - Rp_t] \right] \end{aligned}$$

and after simple algebra,

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

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:

$$\begin{aligned} \alpha_0 &= -\frac{\gamma (1 + \beta_0)^2 \sigma^2 (1 + g)}{R - \frac{1}{1+g}} \frac{1}{y_0 (2 + g)} \\ R\beta_0 &= (1 + \beta_0) \left(\frac{1 + g}{2 + g} + \frac{1}{2 + g} \omega \right) + \beta_1 \\ R\beta_1 &= (1 + \beta_0) \frac{1 - \omega}{2 + g} \end{aligned}$$

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

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

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

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

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 generates a positive trend in prices. The relative reliance of prices on the most recent experiences (dividends) is increasing in the population growth rate.

Appendix OA.4 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 Survey of Consumer Finance 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 1 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, $\mathbb{I}(\text{Fraction of young investors' wealth}_t > \text{Sample average})$,²⁴ i. e.,

$$\text{Fraction of young investors' wealth}_t = \frac{\sum_j \mathbb{I}(\text{age}_{j,t} < 50) \cdot w_{j,t}^{\text{scf}} \cdot \text{Wealth}_{j,t}}{\sum_j w_{j,t}^{\text{scf}} \cdot \text{Wealth}_{j,t}}$$

For robustness, we also consider thresholds 0.55 and 0.60 for the age-based first proxy, and $0.9 \times \text{Sample average}$ and $1.1 \times \text{Sample average}$ for the age- and wealth-based second proxy. Results are presented in Online-Appendix Table OA.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.

²⁴ The SCF documents 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.

Table OA.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 1 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 1 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 is from the SCF, stock data from CRSP.

	Dependent variable: $P_t - D_t$			
	Y_t age-based (1)	Y_t age-based (2)	Y_t age/wealth-based (3)	Y_t age/wealth-based (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

Robust standard errors in parentheses. * significant at 10%; ** significant at 5%.