Investor Experiences and Financial Market Dynamics*

Ulrike Malmendier†

*UC Berkeley, NBER, and CEPR

Demian Pouzo‡

UC Berkeley

Victoria Vanasco§

Stanford University

July 13, 2017

Abstract

How do financial shocks affect investor behavior and market dynamics? Recent evidence suggests that individuals over-weigh personal experiences when forming beliefs and making investment decisions. To study the aggregate implications of such behavior, we propose an OLG model where agents learn from experience. We characterize a novel link between investor demographics and the dependence of prices on past dividends. Different cross-cohort experiences generate persistent heterogeneity in beliefs, portfolio choices, and trade. The model captures many features of asset prices, and produces new implications for the cross-section of asset holdings and market dynamics that are in line with the data.

*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.

†Department of Economics and Haas School of Business, University of California, 501 Evans Hall, Berkeley, CA 94720-3880, ulrike@berkeley.edu

‡Department of Economics, University of California, 501 Evans Hall, Berkeley, CA 94720-3880, dpouzo@berkeley.edu

§Stanford University, Graduate School of Business, 655 Knight Way, Stanford, CA 94305, vvanasco@stanford.edu
I HAVE NOTHING TO DISCLOSE.

Ulrike Malmendier
I HAVE NOTHING TO DISCLOSE.

Demian Pouzo
I HAVE NOTHING TO DISCLOSE.

Victoria Vanasco
1 Introduction

Recent crises in the stock and housing markets have stimulated a new wave of macrofinance 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 we have accumulated 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 above mentioned stylized facts. A core feature of these models is that agents over-weigh recent realizations of the relevant economic variables when forming beliefs.

The over-weighing of recent realizations in the belief formation process is tightly linked to another approach proposed in the literature, the notion of experience effects. This literature argues that personal experiences of economic shocks leave an imprint on individuals’ willingness to take risk. For example, lifetime experiences of stock-market outcomes predict individuals’ future willingness to invest in the stock market.\footnote{See Malmendier and Nagel (2011). Earlier work by Kaustia and Knüpfer (2008) argues the same for IPO experiences. There is also evidence of experience effects in non-financial settings. For example, Oreopoulos, von Wachter, and Heisz (2012) show the long-term effects of graduating in a recession on labor market outcomes, and Alesina and Fuchs-Schündeln (2007) relate the personal experience of living in (communist) Eastern Germany to political attitudes post-reunification. See also Giuliano and Spilimbergo (2013), who relate the effects of growing up in a recession to redistribution preferences.} Much of the empirical evidence on experience effects pertains directly to stated beliefs, e.g., beliefs about future stock returns (in the UBS/Gallup data, cf. Malmendier and Nagel (2011)), beliefs about future inflation (in the Michigan Survey of Consumers, cf. Mal-
mendier and Nagel (2016)), or beliefs about future unemployment rates and the outlook for durable consumption (also in the Michigan Survey of Consumers; cf. Malmendier and Shen (2015)). A key difference between experience-based learning and other recent models of belief formation is that the former generates cohort-specific differences in agents’ beliefs and in their response to a common shock due to different lifetime experiences.

In this paper, we develop an equilibrium model of asset markets to study the aggregate implications of learning from experience. We find that experience-based learning offers a unifying explanation both for financial-market features of prices and trade volume, and for the cross-section of market participation and portfolio decisions. We illustrate how the “living memory” of all cohorts who are active in the market affects equilibrium outcomes and establish a novel link between the empirical patterns of stock-market behavior and demographics. The model generates stylized facts such as excess volatility and return predictability, and it has novel implications for the cross-sectional differences in portfolio choice and time variation in excess trade volume, which we show are consistent with the data. Furthermore, since in our model agent’s extrapolate from past dividends, the model can rationalize survey evidence in Greenwood and Shleifer (2014).

The role of personal experiences in our model comes from the assumed belief formation process, which captures the main two empirical features of experience effects: First, agents over-weigh their lifetime experiences; and second, their beliefs exhibit recency bias. This approach builds closely on the psychology evidence on availability bias, initiated by Tversky and Kahneman (1974), and on the extensive evidence on the different effects of description versus experience. 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-weighing of personal experiences is a pervasive and robust psychological phenomenon. Given this, our model aims to provide a guide

---

for testing whether experience-based learning and demographic factors can enhance our understanding of market dynamics.

The key features of our model are as follows. We consider a stylized overlapping generations (OLG) equilibrium model where agents have CARA preferences and live for a finite number of periods. During their lifetimes, they choose portfolios of a risky and a risk-free security. We initially assume that agents maximize their per-period payoffs (i.e., are myopic). The risky asset is in unit net supply and pays random dividends every period. The risk-less asset is in infinitely elastic supply and pays a fixed return. Investors do not know the true mean of the distribution of dividends, but they 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 to the model. Now, agents form their beliefs by over-weighing their own experiences. We identify two channels through which past dividends affect market outcomes. The first channel is the belief-formation process: shocks to dividends shape agents’ beliefs about future dividends. As a result, individual demands depend on personal experiences, and thus 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 in the population: different lifetime experiences generate persistent belief heterogeneity. Thus, agents in this model “agree to disagree.” Furthermore, younger cohorts react more strongly than older cohorts to a

---

3 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)).

4 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)). We remove the myopia assumption in Section 7. As we show there, the first-order effects of experience-based learning are identical to those shown for myopic agents.
dividend shock, as it makes up a larger part of their lifetime experience. 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 dividend realization and, thus, their reactions to a given recent shock are similar. As a result, price volatility increases and 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.

We further explore the connection between demographics and the dependence of prices on past dividends by analyzing the effect of a one-time demographic shock to our economy. We find that the demographic composition of markets significantly influences the dependence of prices on past dividends. For example, when a demographic change increases the stock-market participation of the young relative to the old, 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 the previous 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. In terms of return predictability (Campbell and Shiller (1988)), we show that the CARA-model analogue, the dividend-price difference, exhibits predictive power for future
price changes. This return predictability stems solely from the experience-based learning mechanism rather than, say, a built-in dependence on dividends or past returns, and it depends on the demographic structure of the market. Similarly, the model generates excess volatility in prices and price changes as established by LeRoy and Porter (1981), Shiller (1981), and LeRoy (2005), which goes above and beyond the stochastic structure of the assumed dividend process.

Experience-based learning generates new predictions for the cross-section of asset holdings and trade volume, which we verify in the data. Using the representative sample of the Survey of Consumer Finance, merged with CRSP and historical data on stock-market performance, we first replicate and extend the evidence in Malmendier and Nagel (2011) and 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, the 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 also show that, in terms of abnormal trade volume, the de-trended turnover ratio is strongly correlated with differences in lifetime market experiences across cohorts.

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.\(^5\) This 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. By exploiting the CARA-Gaussian

---

\(^5\) This form of modeling dynamic portfolio choices is again following a widely used approach in the literature, see Vives (2010).
setup, however, we are able to show that the demand of experience-based learners coincides with the one in a 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 that studies the role of learning in explaining asset pricing puzzles. Closely related to our approach, 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,” as the authors put it. There are two important differences in our setup. First, agents in our model are not Bayesian. Second, in our OLG model, different cohorts have different (finite) experiences. Consequently, observations during an agents’ lifetime have a non-negligible effect on their beliefs and generate heterogeneity across cohorts.

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, where the latter believe that positive changes in prices will be followed by positive changes. In contrast, in our model, the heterogeneity in extrapolation across agents is linked to the demographic structure of the market. In addition, while cross-sectional heterogeneity arises in their framework from the presence of both “rational” and “extrapolative” infinitely-lived agents, in our model, it results from the different experiences of different cohorts of finitely-lived market participants. 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 literature in asset pricing 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 to ours exploring the role of learning in OLG models (Collin-Dufresne, Johannes, and Lochstoer (2016), Schraeder (2015)). The paper most closely related to ours is Ehling, Graniero, and Heyerdahl-Larsen (2015), who explore the role of experience in portfolio decisions and asset prices in a complete markets setting. They focus on trend chasing and the negative relationship between beliefs about expected returns and realized future returns, as shown by Greenwood and Shleifer (2014). Instead, in our incomplete markets setting, we study the cross-section of asset holdings and the relation between demographics and pricing and trading dynamics. Furthermore, our model allows for recency bias in agent’s beliefs formation process.

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

The remainder of the paper is organized as follows. In Section 2, we present the model setup and describe the notion of experience-based learning. In Section 3, we illustrate the mechanics of the model and main results in a simplified version of our model. We present our main results in Section 4, and we extend the model to study demographic
shocks in Section 5. In Section 6, we discuss the model’s empirical implications, and we remove the myopic-agents assumption in Section 7. Section 8 concludes. All proofs are relegated to the Appendix.

2 Model Set-Up

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

Agents have CARA preferences with risk aversion $\gamma$. 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-less asset, which is in perfectly elastic supply and has a gross return of $R > 1$ at all times. 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 $\theta$ and use past observations to estimate it. To keep the model tractable, we assume that the variance of dividends $\sigma^2$ is known at all times.

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

$$W_t^n = x_t^n p_t + a_t^n,$$

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

\[
W_{t+1}^n = x_t^n \left( p_{t+1} + d_{t+1} \right) + a_t^n R = x_t^n \left( p_{t+1} + d_{t+1} - p_t R \right) + W_t^n R.
\]  

(2)

We denote the excess payoff received in \( t + 1 \) from investing in one unit of the risky asset at time \( t \) as \( s_{t+1} \equiv p_{t+1} + d_{t+1} - p_t R \). This is the analogous to the equity premia 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 determinant of portfolio choice generated by EBL. In Section 7, we show that the same mechanism that shapes the results here is at work when this assumption is removed.

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

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

(3)

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

### 2.1 Experience-Based Learning

In this framework, experienced-based learning (EBL) means that agents over-weigh observations received 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. That is, even though they observe the entire history of dividends, EBL agents choose to disregard observations outside their lifetimes.

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 features a passive learning problem in the sense that actions of the players do not affect the information they receive. This would be different if we had, say, a participation decision that would link 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 Malmendier and Nagel (2010):

$$\theta^n_t \equiv \sum_{k=0}^{age} w(k, \lambda, age) d_{t-k},$$

---

6. All we need for our results to hold is that agents discount their pre-lifetime history relative to their experienced history when forming beliefs.

7. 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.
where \( \text{age} = t - n \), and where, for all \( k \leq \text{age} \),

\[
w(k, \lambda, \text{age}) = \frac{(\text{age} + 1 - k)\lambda}{\sum_{k'=0}^{\text{age}}(\text{age} + 1 - k')\lambda}
\]  

(5)

denotes the weight an agent aged \( \text{age} \) assigns to the dividend observed \( k \) periods earlier, and \( w(k, \lambda, \text{age}) \equiv 0 \) for all \( k > \text{age} \).

The denominator in (5) is a normalizing constant that depends only on \( \text{age} \) and on the parameter that regulates the recency bias, \( \lambda \). 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 \( 0 \leq k, k + j \leq \text{age} \),

\[
w(k + j, 1, \text{age}) - w(k, 1, \text{age}) = -\frac{j}{\sum_{k'=0}^{\text{age}}(\text{age} + 1 - k')}
\]

**Example 2.2** (Equal Weights, \( \lambda = 0 \)). For \( \lambda = 0 \), lifetime observations are equal-weighted, i.e., for any \( 0 \leq k \leq \text{age} \), \( w(k, 0, \text{age}) = \frac{1}{\text{age} + 1} \).

**Example 2.3.** For \( \lambda \to \infty \), the weight assigned to the most recent observation converges to 1, and all other weights converge to 0, i.e., for any \( 0 \leq k \leq \text{age} \), \( w(k, \lambda, \text{age}) \to 1_{(k=0)} \).

Observe that by construction, \( \theta_t^\alpha \sim N(\theta, \sigma^2 \sum_{k=0}^{\text{age}}(w(k, \lambda, \text{age}))^2) \). Hence, \( \theta_t^\alpha \) does not necessarily converge to the truth as \( t \to \infty \); it depends on whether \( \sum_{k=0}^{\text{age}}(w(k, \lambda, \text{age}))^2 \to 0 \). This in turn depends on how fast the weights for “old” observations decay to zero (i.e., how small \( \lambda \) 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, ..., 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 to FBL from 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.\(^8\)

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 \( d_k \) available since time 0. The key difference to EBL agents is that different past experiences do not play a role: there is no heterogeneity in beliefs. In addition, beliefs of FBL agents are non-stationary (i.e., depend on the time period), and as \( t \to \infty \), the posterior mean converges (almost surely) to the true mean. That is,

\(^8\)The assumption of Gaussianity is also not needed but simplifies the exposition greatly.
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 \( \tilde{\theta}^n_t \), is given by

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

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 of our learners from experience \( \theta^n_t \) only 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., \( \tau \to \infty \), then \( \tilde{\theta}^{n+a}_t \) coincides with the \( \theta^{n+a}_t \) of EBL agents with \( \lambda = 0 \). The same is true as \( \text{age} \to \infty \).

### 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 \( \{x^n_t\} \), a demand profile for the riskless asset \( \{a^n_t\} \), and a price schedule \( \{p_t\} \) such that:

1. Given the price schedule, \( \{(a^n_t, x^n_t) : t \in \{n, ..., n_q\}\} \) solve the generation-\( n \) problem.

2. The market clears in all periods: \( 1 = \frac{1}{q} \sum_{n=t-q+1}^{t} x^n_t \) 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 \( k \in \{0, ..., K\} \) such that:

\[
\begin{align*}
p_t &= \alpha + \sum_{k=0}^{K} \beta_k d_{t-k}. \\
(6)
\end{align*}
\]

**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, \( \theta \), is known by all agents, i.e., \( E_n^t[d_t] = \theta \forall n, t \). In this scenario, there are no disagreements across cohorts, and the demand of any cohort trading at time \( t \) is

\[
x^n_t \in \arg \max_{x \in \mathbb{R}} E[-\exp(-\gamma x s_{t+1})],
\]

(7)

The solution to this problem is standard and given by,

\[
x^n_t = \frac{E[s_{t+1}]}{\gamma V[s_{t+1}]}
\]

(8)

for all \( n \in \{t-q+1, ..., t\} \), and zero otherwise. Since there is no heterogeneity in cohorts’ demands and there is a unit supply of the risky asset, in any equilibrium, \( x^n_t = 1 \) for all \( n \in \{t-q+1, ..., t\} \), 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 and to highlight, in a simple environment, the main results of the paper, we first solve the model for \( q = 2 \). We generalize the model to any \( q > 1 \) in the next section. (And we solve 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 show that (i) prices depend on the history of dividends, and (ii) 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

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

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

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

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

The difference among cohorts’ demand arises from their different beliefs about future
dividends, $E^t_t[d_{t+1}]$ and $E^{t-1}_t[d_{t+1}]$, given by

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

These formulas illustrate the mechanics of EBL and the cause of heterogeneity among agents. Here, 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 its weighing scheme. An implication of these formulas is that the younger generations react more optimistically (pessimistically) than older generations to positive (negative) changes in recent dividends. In Section 4.3 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, $\lambda$. In Section 4.4, we exploit this observation to link movements in the volume of trade to belief disagreements.

Finally, we impose the market clearing condition, $\frac{1}{2}(x^t_t + x^{t-1}_t) = 1$, 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,

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

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

$$\beta_1 = \frac{R\left(1-\frac{2\lambda}{1+2\lambda}\right)}{(R-1)\left(1+2R-\frac{2\lambda}{1+2\lambda}\right)}.$$
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 expressions for the unconditional price volatility, \( \sigma(p_t) = (\beta_0^2 + \beta_1^2)^{1/2} \sigma \), and price auto-correlation, \( \rho(p_t, p_{t+j}) = \beta_0 \beta_1 \) for \( j = 1 \) and zero for \( j > 1 \). The variance of prices is increasing in the recency bias 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. We show that the general model replicates the results obtained in the toy model, and 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

For any \( s, t \in \mathbb{Z} \), let \( d_{st} = (d_s, ..., d_t) \) 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., \( W^n = 0 \) for all \( n \in \mathbb{Z} \). At time \( t \in \{n, ..., n_q\} \), an agent of generation \( n \) determines her demand for the risky asset maximizing \( E_t^n \left[ -\exp(-\gamma x_{s_{t+1}}) \right] \), 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 $t \in \{n, ..., n_q\}$, demands for the risky asset are given by

$$x^n_t = \frac{E^n_t[s_{t+1}]}{\gamma V[s_{t+1}]} = \frac{E^n_t[s_{t+1}]}{\gamma(1 + \beta_0)^2 \sigma^2}. \quad (13)$$

**Proof.** The result follows by Lemma B.1 in Appendix B. \qed

### 4.2 Characterization of Equilibrium Prices

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 that, 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 $t \in \mathbb{Z}$

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

$$\alpha = -\frac{1}{\left(1 - \sum_{j=0}^{q-1} w_j R^{j+1}\right)^2 R - 1}, \quad \beta_k = \frac{\sum_{j=0}^{q-1-k} w_{k+j} R^{j+1}}{1 - \sum_{j=0}^{q-1} w_j R^{j+1}} \quad k \in \{0, ..., q - 1\} \quad (15)$$

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

Proposition 4.2 establishes a novel link between the factors influencing asset prices and
demographic composition. For each $k = \{0, 1, ..., q - 1\}$, 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, $\beta_k$, 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, $\lambda$.\footnote{In our baseline model we assume cohorts are equally weighted, but 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.}

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 $\lambda$. 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 $\gamma$ decreases the equilibrium price by lowering $\alpha$.

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

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

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, \( \beta_0 \), is increasing in \( \lambda \), with \( \lim_{\lambda \to \infty} \beta_0(\lambda) = 1/(R-1) \) and \( \lim_{\lambda \to \infty} \beta_k(\lambda) = 0 \) for \( k > 0 \).

As \( \lambda \to \infty \), the average weights \( w_k \) (defined in Proposition 4.2) converge to \( 1_{\{k=0\}} \) for all \( k = \{0,1,...,K\} \). Therefore, \( \beta_k \to 0 \) for all \( k > 0 \) and \( \beta_0 \to \frac{1}{R-1} \). In other words, under extreme recency bias (i.e., \( \lambda \to \infty \)), 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, \( \beta_0 \) 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 that we explore in Section 6.

4.3 Cross-Section of Asset Holdings

In this section, we study the implications of EBL for the cross-section of asset holdings. We show that positive shocks (booms) induce a large representation of younger investors while downward shocks (crashes) have the opposite effect. To illustrate this, we first establish that younger generations react more optimistically than older generations 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. \( \frac{\partial x^m_t}{\partial d_{t-k}} \geq \frac{\partial x^n_t}{\partial d_{t-k}} \) for \( 0 \leq k \leq k_0 \) and

2. \( \frac{\partial x^{m}_{t}}{\partial d_{t-k}} \leq \frac{\partial x^{n}_{t}}{\partial d_{t-k}} \) for \( k > k_0 \)
2. \( \frac{\partial x_{m+k}^n}{\partial d_{t-k}} \leq \frac{\partial x_{n+k}^n}{\partial d_{t-k}} \) 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.

We denote the discrepancy between positions of generations \( n \) and \( n+k \), in terms of their investment in the risky asset, as \( \xi(n,k,t) \equiv x_{n+k} - x_{n-k} \). By Proposition 4.1, and some simple algebra, it follows that:

\[
\xi(n,k,t) = \frac{E^n_t[\theta] - E^{n+k}_t[\theta]}{\gamma(1+b_0)\sigma^2} \tag{17}
\]

for any \( k = \{0, ..., t-n\} \) and \( n = \{t-q+1, ..., t\} \). This formulation illustrates that the discrepancy between the positions of different generations is entirely explained by the discrepancy in beliefs. For instance, if for some \( a > 0 \), \( d_{n+a} \approx d_{n+a+t-a} \), then \( \xi(n+a,k,t+a) \approx \xi(n,k,t) \).\(^{10}\)

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 those times, the younger generations exhibit a particularly low willingness to invest in the risky asset, relative to older generations born during those times.

\(^{10}\) 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 \( \left( \sum_{j=0}^{t-n-k} w(j,\lambda,t-n) - w(j,\lambda,t-n-k) \right)/(d_{t+a-j} - d_{t-j}) + \sum_{j=t-n-k+1}^{t-n-1} w(j,\lambda,t-n)(d_{t+a-j} - d_{t-j})/(\gamma(1+b_0)\sigma^2) \).
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 ($x^n_t$) is lower than the demand of the younger generation ($x^{n+k}_t$) 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.

4.4 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:

$$TV_t \equiv \left( \frac{1}{q} \sum_{n=t-q}^{t} (x^n_t - x^n_{t-1})^2 \right)^{\frac{1}{2}}$$

with $x^n_{t-1} = 0$. 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

$$TV_t = \left( \chi^2 \frac{1}{q} \sum_{n=t-q}^{t} \left( \theta^n_t - \theta^n_{t-1} \right) - \frac{1}{q} \sum_{\tilde{n}=t-q}^{t} \left( \theta^{\tilde{n}}_t - \theta^{\tilde{n}}_{t-1} \right) \right)^2 + \frac{1}{q} (x^t_t)^2 + \frac{1}{q} (x^{t-q}_{t-1})^2 \right)^{\frac{1}{2}}, \quad (19)$$

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

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 on agents’ demands, 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 realization. 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 \( d_{t_0} = d_{t_0} + 1 = \ldots = d_{t-1} = \bar{d} \) for \( t - t_0 > q \) 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 \( n \in \{ t - 1 - q, \ldots, t - 1 \} \) and thus trade volume in \( t - 1 \) is simply given by the demand of the youngest -entering,- and the oldest -exiting,- agents.

What happens if, at time \( t \), a dividend \( d_t \neq \bar{d} \) is observed? In that case, for each generation \( n \) trading at time \( t \) and \( t - 1 \), i.e., for \( n = \{ t - q + 1, \ldots, t - 1 \} \), beliefs are given by \( E_t^n [d_{t+1}] = w(0, \lambda, t - n)(d_t - \bar{d}) + \bar{d} \) 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:

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

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, independently of whether the change is positive or negative, and 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. For example, as \( \lambda \to \infty \), the dispersion in weights decreases as \( w(0, \lambda, \text{age}) \to 1 \) for all \( \text{age} \in \{0, \ldots, q - 1\} \). Thus, our results suggest that higher recency bias, \( \lambda \), should generate lower trade volume responses for a given shock to dividends, and vice-versa.

5 Demographics and Equilibrium Prices

In this section, we explore the link between demographics and asset markets by considering an unexpected demographic shock to our economy with \( q = 2 \).\(^{11}\) We denote the mass of young agents at any time \( t \) by \( y_t \), and the total mass of agents at time \( t \) by \( m_t = y_t + y_{t-1} \). We assume that \( y_t = y \) and thus \( m_t = 2y = m \) for all \( t < \tau \) and \( t > \tau + 1 \).

We then consider the effect of a one-time unexpected demographic shock to the mass of agents entering the economy at time \( t = \tau \). We analyze two types of such shocks: a positive shock, \( y_\tau > y \) (e.g. baby-boom), and a negative shock, with \( y_\tau < y \) (e.g. war).

We know from our previous results that when the demographic structure has equal-sized cohorts, prices are given by \( p_t = \alpha + \beta_0 d_t + \beta_1 d_{t-1} \), where \( \{\alpha, \beta_0, \beta_1\} \) are given by (10)-(12). Thus, prices follow this path for \( t > \tau + 1 \) and, since the shock at time \( \tau \) is assumed to be unexpected, for \( t < \tau \) we well.\(^{12}\) For these time periods the economy is as the one described in Section 3. We are left to characterize demands and prices for \( \tau \) and \( \tau + 1 \), when the generation of the demographic shock is young and old respectively. We

\(^{11}\) We have also analyzed the implications of having a growing population, as opposed to a one-time demographic shock. In Online Appendix OA.3, we show that population growth generates a positive trend in prices, but that it is independent of experience effects: The growing mass of agents increases the demand for the asset, and hence prices adjust to clear markets, since asset supply is assumed to be constant. EBL affects the path of the prices. In particular, we find that the relative reliance of prices on the most recent dividend is increasing in the population growth rate.

\(^{12}\) We use the assumption that the shock is unexpected to construct the series in Figure ??; but none of the result for \( t \geq \tau \) depend on this assumption.
make the following guesses:

\[
p_\tau = a_\tau + b_{0,\tau}d_\tau + b_{1,\tau}d_{\tau-1}
\]

\[
p_{\tau+1} = a_{\tau+1} + b_{0,\tau+1}d_{\tau+1} + b_{1,\tau+1}d_\tau
\]

We solve the problem by backwards induction. Note that the form of agent’s demands remains unchanged. By imposing market clearing in \(\tau + 1\), with mass \(y\) of young agents and \(y_\tau\) of old agents, and using the method of undetermined coefficients we obtain

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

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

\[
b_{1,\tau+1} = \beta_1 \frac{y_\tau}{m_\tau} \frac{m_\tau}{y}
\]

where \(\omega \equiv \frac{2^\lambda}{1+2^\lambda}\) and \(m_\tau = y + y_\tau\). Note that for \(y_\tau = y\), the coefficients are as those in the baseline model in equations (10)-(12). The above expressions show that the total mass of agents \(m_\tau\) only affects the price constant, while the price loadings depend solely on the fraction of young agents in the economy. As before, we impose market clearing in \(\tau\), with mass \(y_\tau\) of young agents and \(y\) of old agents. With the method of undetermined coefficients we obtain

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

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

\[
b_{1,\tau} = \frac{1}{R} \left( 1 + b_{0,\tau+1} \right) \frac{m_\tau - y_\tau}{m_\tau} (1 - \omega).
\]

Figure 2 shows how the reliance of prices on past dividends changes with the size and direction of the demographic shock. From the first two panels, we can see that a positive
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 \( y_\tau \). The results are for \( y = 0.5, \gamma = 1, \lambda = 3, \sigma = 1, \) and \( R = 1.1 \).

demographic shock generates a stronger response of prices to the contemporaneous dividend and a weaker response to past dividends. Furthermore, these responses increase in the size of the demographic shock. The more young people are in the market, who pay no attention to past dividends, the more do current dividends matter relative to past dividends. Consistent with this, we also see that when the \( \tau \)-generation of ‘baby boomers’ becomes old, prices depend less on contemporaneous dividends and more on past dividends than in the baseline case. In addition, the third panel shows that a positive demographic shock generates a level increase in prices that is captured in an increase in the 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 for a negative demographic shock, as can be seen at the left side of each graph.

The results in this section generalize beyond the demographic shocks such as baby booms and wars, which affect cohort size, to shocks affecting the market participation of cohorts, even when their total size is stable. For instance, if other factors increase the market participation of young generations relative to old generations, ceteris paribus, our model predicts that the reliance of equilibrium prices on the most recent dividends
relative to past dividends goes up.

6 Empirical Implications

In this section, we analyze the empirical implications of our model. We divide the analysis into two parts. First, we show that the model is able to account for several asset pricing features that have been identified in the data. These include the predictability of stocks returns (Fama and French (1988)), and the excess volatility puzzle (LeRoy (2005) and Shiller (1981)). Second, we verify that the model’s new predictions about cross-sectional differences in portfolio choices, trading decisions, and resulting trade volume are in line with empirical observations.

6.1 Asset Pricing

We first take the qualitative predictions regarding the above-mentioned asset-pricing features and asset-pricing puzzles to the data. Even though the CARA-Normal framework is not well suited for a thorough calibration exercise, we can follow the approach in 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 work with quantities defined in terms of differences as opposed to ratios. For example, instead of stock returns we measure price changes, \( \Delta P \), and instead of the dividend-price ratio we study the difference \( D - P \).

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. All of our parameters are either observable (such as the risk-free rate), or have been estimated from micro-data (as the recency bias). For our numerical
solutions, and following Barberis, Greenwood, Jin, and Shleifer (2015), we choose the following parameter values: 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 $\lambda \in \{1, 3, 10\}$ and for $q \in \{5, 10\}$.

**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. EBL rationalizes such predictability and, at the same time, limits it to those dividend realizations experienced by the oldest cohorts participating in the market. To analyze the predictability generated in our model, we compute the following measures of co-movement between (i) $D_t - P_t$ and price changes $P_{t+z} - P_t$, and (ii) $D_t - P_t$ and future equity premia, computed as $D_{t+z} + P_{t+z} - RP_t$,

$$B_{DP}(z) \equiv \frac{\text{Cov}(D_t - P_t, P_{t+z} - P_t)}{\text{Var}(D_t - P_t)}, \quad \text{and} \quad B_{DP}^{ep}(z) \equiv \frac{\text{Cov}(D_t - P_t, D_{t+z} + P_{t+z} - RP_t)}{\text{Var}(D_t - P_t)},$$

where $z$ denotes the horizon.

Figure 3 illustrates the predictive power of the dividend-price difference for future price changes and future equity premia for different horizons, different number of cohorts, and different levels of recency bias. The figure shows that, as evidenced in the data (Cochrane 2011), the model generates a positive (and strong) relation between $D - P$ and future price changes, which increases with the horizon.

The predictability of excess returns is an equilibrium phenomenon that stems solely from our learning mechanism and not from, say, a built-in dependence on dividends or past returns. This is why, as the figure illustrates, the predictability of returns increases in $\lambda$, the level of recency bias. 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, however, has the additional
implication that different demographic structures generate different $\beta$’s, which directly determine the level of predictability (or extrapolation) of stock returns.

This prediction is in line with the findings in Cassella and Gulen (2015), who estimate the extent to which investors’ recent return experiences, relative to older return experiences, help predict future returns. The authors first provide evidence that, when the level of extrapolation bias is high in the market, the predictive power of the price-dividend ratio for future returns goes up. Second, and in line with our model predictions, they find a positive relation between their market-wide measure of extrapolation and the relative participation of young versus old investors in the stock market. That is, as the number of young agents in the market increases, the level of extrapolation observed in the data increases. Our model with EBL goes beyond rationalizing evidence on agents extrapolating from past dividends (cf. also Greenwood and Shleifer (2014)) in that it puts structure on the extent of such extrapolation that is aligned with empirical observations.

Price Dynamics. The second set of asset pricing implications are related to the dynamics of prices. We focus our analysis on price volatility, price change volatility, and on the autocorrelation of the price-dividend difference. Table 1 presents the first set of...
results, regarding excess volatility. As shown in the table, our model generate both excess volatility in prices and in $P - D$. In particular, that price volatility increases with the level of recency bias, and decreases with the number of cohorts in the market.

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$\sigma(P-D)$</th>
<th>$\sigma(P)$</th>
<th>$\sigma(D) / \sigma(P)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q=5$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda=1$</td>
<td>3.07</td>
<td>3.24</td>
<td>0.08</td>
</tr>
<tr>
<td>$\lambda=3$</td>
<td>3.71</td>
<td>4.10</td>
<td>0.06</td>
</tr>
<tr>
<td>$\lambda=10$</td>
<td>4.38</td>
<td>5.08</td>
<td>0.05</td>
</tr>
<tr>
<td>$q=10$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda=1$</td>
<td>2.31</td>
<td>2.24</td>
<td>0.11</td>
</tr>
<tr>
<td>$\lambda=3$</td>
<td>2.90</td>
<td>3.08</td>
<td>0.08</td>
</tr>
<tr>
<td>$\lambda=10$</td>
<td>4.66</td>
<td>4.52</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Table 1: Excess Volatility.

Figure 4 displays the autocorrelation of $P - D$ and the volatility of price changes for different horizons. As in the data, $P - D$ is highly autocorrelated at short lags, with this correlation being zero at longer horizons. Furthermore, the $P - D$ autocorrelation decreases in the recency bias, and increases with the number of cohorts in the market. A direct implication is that the $P - D$ is only positively correlated with lags that are bellow the number of cohorts in the market, and that the correlation vanishes as $\lambda \to \infty$ for any number of cohorts. Most importantly, the Figure shows that the model can generates ample volatility in price differences, and that this excess volatility is increasing in the recency bias and decreasing with the number of cohorts in the market. In our calibration, the volatility of dividends is between 0.03 and 0.05 that of the change in prices, consistent with evidence in Campbell and Kyle (1993) who estimate it at 0.032.

6.2 Cross-Section of Asset Holdings and Trade Volume

We now turn to the novel empirical predictions 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-asset 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 on the extensive margin, i.e., stock market participation, and on the intensive margin, i.e., amount of liquid assets invested in the stock market. The second prediction is that large changes in the cross-section of experience-based beliefs generate trade (Proposition 4.6).

To test these model implications, we combine historical data on stock-market performance, obtained from Robert Shiller’s website, with data on stock holdings from the Survey of Consumer Finances (SCF) and on stock turnover data from the Center for Research in Security Prices (CRSP).

The key variable to construct is a measure of lifetime experiences of dividends. Dividends in our model do not translate one-to-one to the dividend payments recorded in CRSP, though. Theoretically, dividends in the Lucas-tree economy capture the performance of the risky asset, or the stock market. Empirically, however, dividend payments do not necessarily reflect how well firms are doing. For example, firms have incentives to smooth dividend payments, or they may retain earnings rather than distribute them.
to shareholders. We therefore use an array of empirical measures to capture the performance of the risky asset in our model. Specifically, we use annual stock market returns, real dividends, real earnings, and 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 capturing the performance of the risky asset at medium frequencies. Therefore, for all non-stationary series (i.e., for the dividend, earnings, and GDP time series), we use the Christiano and Fitzgerald (2003) band-pass filter and remove stochastic cycles at frequencies higher than 2 years and lower than 8 years, which are the default frequencies for the CF-filter.\footnote{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.}

Next, we construct experienced returns, dividends, earnings, and GDP of different generations over the course of their life. To do so, we apply the formula from equation (5) in the calculation of generation-specific weighted averages, and employ both linearly declining weights ($\lambda = 1$), and a steeper weighting function (with $\lambda = 3$), corresponding to the range of empirical estimates in Malmendier and Nagel (2011).

**Stock market participation.** To test the first prediction, we relate the differences in lifetime experiences between older cohorts above 60 years of age (i.e., ages 61 to 74) and younger cohorts below 40 years of age (i.e., ages 24 to 39) to the differences in their stock market investment. Our source of household-level micro data is the Survey of Consumer Finances (SCF), which provides repeated cross-sectional observations on asset holdings and various household background characteristics. Our variable construction follows Malmendier and Nagel (2011). We use all waves of the modern, triannual SCF, available from...
the Board of Governors of the Federal Reserve System since 1983, i.e., extend the analysis of Malmendier and Nagel (2011) 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. The precursor survey starts in 1947, but includes age (rather than 5- or 10-year brackets) only since 1960. We use all survey waves that also offer stock-market participation information, i.e., the 1960, 1962, 1963, 1964, 1967, 1968, 1969, 1970, 1971, and 1977 surveys.

For the extensive margin of stock holdings, we construct a binary variable for stock-market participation. It indicates whether 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).\textsuperscript{15}

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 calculated with all surveys from 1960-2013 other than 1971. Liquid assets are defined as stock holdings plus bonds plus cash and short-term instruments (checking and savings accounts, money market mutual funds, certificates of deposit). In those analyses, we drop all households that have no money in stocks.

For both the young and old age group, we calculate the experienced performance and stock market investment as a weighted average across cohorts, where weights are determined by the weight variable provided in the SCF. The weighted estimates are representative of the U.S. population.\textsuperscript{16}

\textsuperscript{15} 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.

\textsuperscript{16} The 1983-2013 waves of the SCF oversample high-income households with significant stock holdings. The oversampling is helpful for our analysis of asset allocation, but could also induce selection bias. By weighting the data using SCF sample weights, we undo the over-weighing of high-income households and
We present the results graphically. The figures depict the relationship between stock holdings (both extensive and intensive margin) and experienced returns (Figure 5.1), dividends (Figure 5.2), earnings (Figure 5.3), and GDP (Figure 5.4). (Note that graphs (a) and (c) of Figure 5.1 update the evidence on the extensive margin and experienced 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 Figure 5.1, graph (a) shows 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, and that the opposite holds when the returns experienced by 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 appears to make 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 relatively higher share of the their liquid assets into stocks 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 5.2 to 5.4 present the corresponding results for experienced dividends, earnings, and GDP. Across 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 adds credibility to the link between our theoretical model and the empirical facts, and ameliorates con-
also adjust for non-response bias.
Figure 5.1. 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.
Figure 5.2. 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 5.3. 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.
Figure 5.4. 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.
concerns about dividends in our model not translating one-to-one into a specific empirical performance measure.

**Trade volume.** We now turn to the second prediction that relates trade volume with the dispersion of changes in the level of disagreement among investors. We calculate this 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.\(^{17}\)

As a measure of abnormal trade volume, we calculate the deviation of the turnover ratio from its trend as follows. Following the prior literature (Statman, Thorley, and Vorkink (2006), Lo and Wang (2000)), we first compute firm-level turnover ratio, i.e. the ratio of number of shares traded and shares outstanding, on a monthly basis. We require that firms be listed on the NYSE or AMEX. We exclude NASDAQ-listed firms due to the concern that 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

---

\(^{17}\) For this, we obtain data on U.S. population by age between 1985 and 2015 from US Census Bureau.
firms by their market capitalization.\textsuperscript{18} 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, in order to 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 6 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, for a given year, the average of the turnover ratio in December of that year and in January of the following year as our measure for trading volume. That is, Figure 6 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, i.e., up to the mid-1980s, making it less likely that (individual) investors trade repeatedly based on experienced performance.

Table 2: Correlation between Trading Volume and S.D. of Changes in Experiences

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

Consistent with the predictions of our model, we observe a clear and statistically significant (at 1%, see Table 2) co-movement between disagreement among cohorts and trading

\textsuperscript{18} 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.
volume. Finally, Table 2 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. dividends, earnings, or GDP. In each case, the correlation coefficient is positive. It is significant at 10% for changes in disagreement in experienced earnings; it is however insignificant for the other two measures.

All sets of empirical findings are consistent with our model of experience-based learning, while it appears that 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. Now, agents make their portfolio decisions to maximize the expected utility of consuming their final wealth, as is standard in the literature (see (Vives 2010)). Thus, agents need to account for the dynamic nature of their portfolio problem. The analysis illustrates 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,” since agents may want 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 focus on the two-generations model, $q = 2$, to compare the remaining findings with those of the model with myopic agents.
7.1 Characterization of risky demands

For any $s, t \in \mathbb{Z}$, let $d_{s:t} = (d_s, ..., d_t)$ denote the history of dividends from time $s$ up to time $t$. At time $n$, an $n$-generation agent solves the following problem:

$$\max_{x \in \mathbb{R}^q} E^n_n \left[ - \exp \left( -\gamma W^n_{n+q}(x) \right) \right]$$

$$\text{s.t. } W^n_{n+q}(x) = \sum_{\tau = n}^{n+q-1} R^{n+q-\tau} x_{\tau+1} s_{\tau+1}$$

where $x \in \mathbb{R}^q$ are the $q$ trading decisions from $n$ up to $n + q - 1$. 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

$$V^n_{n_q}(d_{n_q-K:n_q}) = \max_{x \in \mathbb{R}} E^{n_q}_{n_q} \left[ - \exp \left( -\gamma s_{n+q} x \right) \right] \ \text{and}$$

$$V^n_{\tau}(d_{\tau-K:\tau}) = \max_{x \in \mathbb{R}} E^{\tau}_{\tau} \left[ V^n_{\tau+1}(d_{\tau+1-K:\tau+1}) \exp \left( -\gamma s_{\tau+1} x \right) \right] \ \forall \tau \in \{n, ..., n_q - 1\}$$

Note that our definition of $V^n_{n_q}$ omits the term $\exp\{ -\gamma RW^n_{n_q} \}$ since it does not affect the maximization. The same applies for all $\tau < n + q - 1$, and we thus omit the wealth at time $\tau$ as well. The formulation of the maximization problem in equations (23) and (24) shows that, while the $n$-generation’s problem is a static portfolio problem at $n_q$, it is not static at any other time $\tau \in \{n, ..., n_q - 1\}$ because $V^n_{\tau+1}$ is correlated with $s_{\tau+1}$ through dividends. Dividend realization $d_{\tau+1}$ impacts (i) the net payoff obtained from investing $x_{\tau}$ in the risky asset at time $\tau$, and (ii) the continuation value $V^n_{\tau+1}(d_{\tau+1-K:\tau+1})$ 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 $E_{N(\mu, \sigma^2)}[.]$ and $V_{N(\mu, \sigma^2)}[.]$ be the expectation and variance with respect to a Gaussian pdf with mean $\mu$ and $\sigma^2$. The following proposition characterizes 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 \( j \in \{0, \ldots, q-1\} \), i.e., in period \( n+j \), is given by

\[
x^n_{n+j} = \frac{E_N(m_j, \sigma^2_j)[s_{n+j+1}]}{\gamma R^{q-1-j} V_N(m_j, \sigma^2_j)[s_{n+j+1}]}
\]

where \( m_j \equiv \theta_{n+j}^2 - \sigma^2 (b_j + \sum_{k=1}^{K} b_j(k)d_{n+j-k}) \), \( \sigma^2_j \equiv \frac{\sigma^2}{2c_j\sigma^2+1} \), 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 (23)), the problem is in fact a static one. In particular, we show that \( V^n_{n_q} \) is exponential-quadratic in \( d_{n_q} \). (See Lemma B.1 in the Appendix.) We then show that the exponential-quadratic term times the Gaussian distribution of dividends imply a new Gaussian distribution with an slanted mean and variance. (See Lemma C.1 in the Appendix.) Thus the problem at time \( n_q - 1 \) can be viewed as a static problem with a modified Gaussian distribution, and consequently demands are of the form shown in (25), and \( V^n_{n_q-1} \) is also exponential-quadratic. The process thus continues until time \( n \).

After straightforward algebra, we can cast equation (25), as

\[
x^n_{n+j} = \frac{1}{R^{q-1-j}} \cdot \frac{E_N(m_j, \sigma^2_j)[s_{n+j+1}]}{\gamma V_N(m_j, \sigma^2_j)[s_{n+j+1}]} - \frac{b_j + \sum_{k=1}^{K} b_j(k)d_{n+j-k}}{\gamma R^{q-1-j} (1 + \beta_0)}
\]

\[
\equiv \frac{1}{R^{q-1-j}} \cdot \tilde{x}^n_{n+j} + \Delta^n_{n+j}
\]

The term \( \tilde{x}^n_{n+j} \) coincides with the demand of a static portfolio problem for an agent with beliefs \( \theta^n_{n+j} \) (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, \( \Delta^n_{n+j} \equiv -\frac{b_j + \sum_{k=1}^{K} b_j(k)d_{n+j-k}}{\gamma R^{q-1-j} (1 + \beta_0)} \), is an adjustment which accounts for the dynamic nature of the problem; we thus denote it the dynamic component. It arises

\[19\] Note that \( E_N(b+a,s)[s_{n+1}] = E_N(a,s)[s_{n+1}] + (1 + \beta_0)b \).
because agents understand that they are learning about the risky asset, and that therefore
the value function is correlated with the one-period-ahead returns.

7.2 Characterization of equilibrium prices

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 time \( t \):

**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 \( t \in \mathbb{Z}, q \geq 1 \),

\[
p_t = \alpha + \sum_{k=0}^{q-1} \beta_k d_{t-k}.
\]

The proof follows along the same lines as the one for the myopic case.\(^{20}\) This result
shows that the insights in Proposition 4.2 continue to hold in the general setup.

7.3 Illustration: The \( q = 2 \) Case

We now return to the special case of two-period lived agents, i.e, \( q = 2 \), to sharpen our
results regarding the behavior of prices and risky demands in equilibrium. We relegate the
derivation of the coefficients on the price function, \( \{\alpha, \beta_0, \beta_1\} \), to the Online Appendix; see
Section OA.2. (In Lemma OA.2.1, we solve the corresponding system of linear equations.)
With this result, we are able to establish that prices react positively to dividends \( d_t \) and
\( d_{t-1} \). Specifically,

**Proposition 7.3.** For \( \lambda > 0, \alpha \leq 0 \) 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

\(^{20}\) 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, say, \( s_{t+q-1} \) is deterministic given the information at
time \( t + q - 2 \), and thus the agent will take infinite positions.
The relation between the recency bias, \( \lambda \), and price loadings is as the one shown for the model with myopic agents, and is depicted in Figure 7 for different risk-free rates.\(^{21}\) The next proposition establishes that the demand of the young generation increases (decreases), while the one of the old generation decreases (increases), when current dividends increase (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) \( \frac{\partial x^*_t}{\partial d_t} > 0 > \frac{\partial x^*_t}{\partial d_{t-1}} \), and (2) \( \frac{\partial x^*_t}{\partial d_{t-1}} < 0 < \frac{\partial x^*_t}{\partial d_{t-1}} \).

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.

\(^{21}\) The values of \( \{\beta_0, \beta_1\} \) are independent of the process for dividends, \( \sigma^2 \), and of the coefficient of risk aversion, \( \gamma \). Thus, the results shown in the figure do not depend on parameter values other than the ones used for comparative statics, \( (\lambda, R) \).
Overall, the results of this section suggest that our main results are still present in a setting without the myopia assumption. At the same time, this setting reveals additional factors influencing agents decision-making due to agents’ understanding of how their 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 a stylized OLG equilibrium framework to study the effect of personal experiences on market dynamics. We incorporate the two main empirical features of experience effects, the over-weighing of lifetime experiences and recency bias, into the belief formation process of agents. We find that EBL not only generates several asset pricing features observed in the data, but it also produces new empirical implications about the relation between demographics, price and 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, since all agents update their beliefs about the future after observing a given shock. The second is the cross-sectional heterogeneity in the population, since 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


Appendix A  Proofs for Results in Section 2

Proof of Lemma 2.1. Let \( \Delta(k) \equiv w(k, \lambda, \text{age}) - w(k, \lambda, \text{age}') \) for all \( k \in \{0, \ldots, \text{age}\} \). We need to show that \( \exists k_0 \in \{0, \ldots, \text{age}\} \) 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 \( k > \text{age}' \), \( \Delta(k) > 0 \) since \( w(k, \lambda, \text{age'}) \equiv 0 \), and hence \( \Delta(k) = w(k, \lambda, \text{age}) > 0 \), for all \( k \in \{\text{age}'+1, \ldots, \text{age}\} \).

For \( k \leq \text{age}' \), we note that \( \Delta(k) > 0 \iff Q(k) := \frac{w(k, \lambda, \text{age})}{w(k, \lambda, \text{age'})} > 1 \). Hence, it remains to be shown that \( \exists k_0 \in \{0, \ldots, \text{age}'\} \) such that \( Q(k) < 1 \) for all \( k \leq k_0 \), and \( Q(k) \geq 1 \) for all \( k > k_0 \).

Since the normalizing constants used in the weights \( w(k, \lambda, \text{age}) \) are independent of \( k \) (see the definition in (5)), we absorb them in a constant \( c \in \mathbb{R}^+ \) and rewrite

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

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

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 \( p_t = \alpha + \beta_0d_t + \ldots + \beta_Kd_{t-K} \) 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 \( x_t^n = \frac{E^n[p_{t+1}]}{\gamma V^n[t+1]} \). Plugging in our guess for prices, and for \( \beta_0 \neq -1 \), we obtain:

\[
x_t^n = \frac{(1 + \beta_0) \theta_t^n + \alpha + \beta_1d_t + \ldots + \beta_qd_{t-q+1} - p_tR}{\gamma (1 + \beta_0)^2 \sigma^2}
\]

(28)
By market clearing, \( \frac{1}{q} \sum_{n=t+1}^{q} x_{t}^{n} = 1 \), which implies that

\[
\frac{(1 + \beta_{0}) \frac{1}{q} \sum_{n=t+q+1}^{t} \theta_{t}^{n}}{\gamma (1 + \beta_{0})^{2} \sigma^{2}} + \frac{\alpha + \beta_{1} d_{t} + \ldots + \beta_{q} d_{t-q+1} - p_{t} R}{\gamma (1 + \beta_{0})^{2} \sigma^{2}} = 1.
\]

By straightforward algebra and the definition of \( \theta_{t}^{n} \), it follows that

\[
(1 + \beta_{0}) \frac{1}{q} \sum_{n=t-q+1}^{t} \left[ \sum_{k=0}^{t-n} w(k, \lambda, t-n) d_{t-k} \right] + \left[ \alpha - \gamma (1 + \beta_{0})^{2} \sigma^{2} \right] + \beta_{1} d_{t} + \ldots + \beta_{q} d_{t-q+1} = p_{t} R.
\]

Plugging in (again) our guess for \( p_{t} \) and using the method of undetermined coefficients, we find the expressions for \( \alpha \) and the \( \beta \)'s:

\[
-\frac{\gamma (1 + \beta_{0})^{2} \sigma^{2}}{R - 1} = \alpha \tag{29}
\]

\[
(1 + \beta_{0}) \frac{1}{q} \sum_{n=t-q+1}^{t} \left[ \sum_{k=0}^{t-n} w(k, \lambda, t-n) \right] + \beta_{k+1} = \beta_{k} R \quad \forall k \in \{0, 1, \ldots, q-1\} \tag{30}
\]

\[
0 = \beta_{q} R. \tag{31}
\]

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 (31) we obtain:

\[
(1 + \beta_{0}) w_{k} + \beta_{k+1} = \beta_{k} R \quad \forall k \in \{0, 1, \ldots, q-2\} \tag{32}
\]

\[
(1 + \beta_{0}) w_{q-1} = \beta_{q-1} R \tag{33}
\]

By solving this system of equations, we obtain the expressions in the proposition. In particular, 

\[
(1 + \beta_{0}) \left( w_{q-2} + w_{q-1}/R \right) = \beta_{q-2} R \quad \text{for} \quad k = q - 2, \quad (1 + \beta_{0}) \left( w_{q-3} + w_{q-2}/R + w_{q-1}/R^{2} \right) = \beta_{q-3} R \quad \text{for} \quad k = q - 3, \quad \text{and so on.}
\]

This allows us to express (32) and (33) as

\[
(1 + \beta_{0}) \sum_{j=0}^{k-1} w_{q-(k-j)}/R^{j} = \beta_{q-k} R \quad \text{for} \quad k = 1, \ldots, q. \tag{34}
\]

The last expression (34) implies \( \beta_{0} = \frac{\sum_{j=0}^{q-1} w_{j}/R^{j}}{R - \sum_{j=0}^{q-1} w_{j}/R^{j+1}} = \frac{\sum_{j=0}^{q-1} w_{j}/R^{j+1}}{1 - \sum_{j=0}^{q-1} w_{j}/R^{j+1}} \) (from plugging in \( k = q \)), which in turn, plugged into (29) allows us to obtain the expression for \( \alpha \) from (15) in Proposition 4.2. And expression (34) implies \( \beta_{k} = \frac{\sum_{j=0}^{q-1} w_{k+j}/R^{j+1}}{1 - \sum_{j=0}^{q-1} w_{j}/R^{j+1}} \) (from substituting \( k \) with \( q - k \)), and using the expression for \( \beta_{0} \) as expressed in equation (16) of the Proposition. The latter also subsumes equation (33), solved for \( \beta_{q-1} \), and the above formula for \( \beta_{0} \), and hence holds for \( k = 0, \ldots q - 1 \). \( \square \)

Proof of Proposition 4.3. For this proof, we use equations (32) and (33). In addition, note that by construction, \( w_{k} < w_{k-1} \) for \( \lambda > 0 \) since for all generations, \( w(k, \lambda, qge) \) 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
\[ \beta_0 > 0 \text{ then } \beta_{q-1} > 0 \text{ and} \]

\[ \beta_{q-1} = \frac{1}{R} (1 + \beta_0) w_{q-1} < \frac{1}{R} [(1 + \beta_0) w_{q-2} + \beta_{q-1}] = \beta_{q-2} \quad (35) \]

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

\[ \beta_{k-1} = \frac{1}{R} [(1 + \beta_0) w_{k-1} + \beta_k] < \frac{1}{R} [(1 + \beta_0) w_{k-2} + \beta_{k-1}] = \beta_{k-2} \quad (36) \]

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

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

\[ G_q(\lambda) = \sum_{age}^{q-1} \sum_{k=0}^q w(k, \lambda, age)/R^{k+1} \quad (37) \]

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

To show the limit results, note that \( \lim_{\lambda \to \infty} w(0, \lambda, age) = 1 \), while \( \lim_{\lambda \to \infty} w(k, \lambda, age) = 0 \) 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 \( \{t - q + 1, \ldots, t\} \) and any \( k \in \{0, \ldots, q-1\} \),

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

We note that, for any \( n \in \{t - q + 1, \ldots, t\} \), \( \frac{\partial m}{\partial d_{t-k}} = w(k, \lambda, n - t) \) if \( k \in \{0, \ldots, t - n\} \), and \( \frac{\partial m}{\partial d_{t-k}} = 0 \) if \( k \in \{t - n + 1, \ldots, q - 1\} \). (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 \( k \in \{0, \ldots, q-1\} \). (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 any \( k \in \{0, \ldots, k_0\} \) and \( w(k, \lambda, t - n) \geq w(k, \lambda, t - m) \) for the rest of the \( k \)'s, \( k \in \{k_0 + 1, \ldots, q-1\} \). \( \square \)

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

**Lemma B.2.** For any \( a \in \{0, 1, \ldots\} \), \( a' < a \) and any \( m \in \{0, \ldots, a\} \), 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 \( m \in \{0, \ldots, a\} \).

**Proof.** See Online Appendix OA.1.1. \( \square \)
Proof of Proposition 4.5. We first introduce some notation. For any \( j \in \{t-n-k+1, ..., t-n \} \), 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, \( \sum_{j=0}^{t-n-k} w(j, \lambda, t-n-k)dt_{t-j} = \sum_{j=0}^{t-n} w(j, \lambda, t-n-k)dt_{t-j} \).

In addition, we note that \((w(j, \lambda, t-n-k))_{t-n}^{t-n} \) and \((w(j, \lambda, t-n))_{t-n}^{t-n} \) are sequences of positive weights that add to one.

Let for any \( m \in \{0, ..., t-n \} \),

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

These quantities, as functions of \( m \), are non-decreasing and \( F(t-n, t-n-k) = F(t-n, t-n) = 1 \). Moreover, \( F(m+1, t-n) - F(m, t-n) = w(m+1, \lambda, t-n) \) 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,

\[
\xi(n, k, t) = \sum_{m=0}^{t-n}(F(m, t-n) - F(m-1, t-n))d_{t-m} - \sum_{m=0}^{t-n}(F(m, t-n-k) - F(m-1, t-n-k))d_{t-m}
\]

\[= \frac{\sum_{j=0}^{t-n-1}(d_{t-j} - d_{t-j-1})(F(j, t-n) - F(j, t-n-k))}{\gamma(1 + \beta_0)\sigma^2}.
\]

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

\[
x^n_t = \frac{1}{\gamma\sigma^2(1 + \beta_0)^2} \left( \alpha_0(1 - R) + (1 + \beta_0)\theta^n_t - R\beta_0 d_t + \sum_{k=1}^{q-1} \beta_k(d_{t+1-k} - Rd_{t-k}) \right).
\]

Thus, for \( n \in \{t-q+1, ..., t-1\} \),

\[
x^n_t - x^n_{t-1} = \frac{(1 + \beta_0)\left(\theta^n_t - \theta^n_{t-1}\right) + \mathcal{T}(d_{t-q})}{\gamma\sigma^2(1 + \beta_0)^2}
\]

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

The fact that \( x^n_t - x^n_{t-1} = x^n_t \) and \( x^n_t - x^n_{t-q} = -x^{t-q}_{t-1} \), and market clearing imply \( q^{-1}\left(\sum_{n=t-q}^{t} x^n_t - x^n_{t-1}\right) = \)
0. This expression and the expression in (39) imply that
\[
\frac{1}{q} \left( \sum_{n=t-q+1}^{t-1} \frac{(1 + \beta_{0})(\theta^n_t - \theta^n_{t-1})}{\gamma \sigma^2 (1 + \beta_{0})^2} + x^n_t - x^{t-q}_t \right) = -\frac{1}{q} \sum_{n=t-q}^{t-1} \frac{T(d_{t:t-q})}{\gamma \sigma^2 (1 + \beta_{0})^2} = -\frac{T(d_{t:t-q})}{\gamma \sigma^2 (1 + \beta_{0})^2}.
\]

Letting \(\theta^n_{t-1} = \theta^n_t - q = 0\), it follows that
\[
\frac{1}{q} \left( \sum_{n=t-q}^{t-1} (1 + \beta_{0})(\theta^n_t - \theta^n_{t-1}) \right) = -T(d_{t:t-q}).
\]

Thus, we can express the change in individual demands for those agents with \(n = \{t-q+1, \ldots, t-1\}\) in expression (39) as follows:
\[
x^n_t - x^{t-q}_t = \chi \left[ (\theta^n_t - \theta^n_{t-1}) - \frac{1}{q} \sum_{n=t-q}^{t-1} (\theta^n_t - \theta^n_{t-1}) \right], \quad \forall n \in \{t, \ldots, t-q\}
\]
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 \(x^n_{n+q} = 0\) and \(x^n_{n+q-1} = \frac{E_{n+q-1}[s_{n+q}]}{\gamma \sigma^2 (1 + \beta_{0})^2}\) \(\forall n \in \mathbb{Z}, q \geq 1\).

**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

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

and

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

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

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

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

with

\[
\delta(q) = \delta_k(q) = 0 \quad \forall k \in \{0, ..., K\}
\]

\[
\delta(q - 1) = \frac{\alpha(1 - R)}{\gamma((1 + \beta_0)\sigma)^2}
\]

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

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

and for \( \text{age} \in \{0, ..., q - 2\} ,

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

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

\[
\forall k \in \{0, ..., q - 1\},
\]

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

\[
\forall k \in \{q, ..., K - 1\}
\]

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

55
and \( s_{q-1} = \sigma \) and \( s_{\text{age}}^2 \equiv \frac{\sigma^2}{(R^{t-1-(\text{age}+1)}\gamma)^2((1+\beta_0)\kappa_{\text{age}+1})^2(\delta_0(\text{age}+1))\sigma^2+1} \)

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

\[
\begin{align*}
b_j & \equiv (R^{t-1-j}\gamma)^2((1+\beta_0)\sigma_j)^2\delta(j)\delta_0(j) \\
b_j(k) & \equiv \delta_j(k)\delta_0(j)(R^{t-1-j}\gamma)^2((1+\beta_0)\sigma_j)^2
\end{align*}
\]

and, \( c_{q-1} = 1 \) and

\[ c_{j-1} = 0.5(R^{t-1-(j+1)}\gamma)(1+\beta_0)\sigma_{j+1}\delta_0(j+1) \]

for \( j \in \{0, \ldots, q-2\} \).

**Proof of Proposition 7.1.** By Lemma B.1, it follows that \( x_{t+q-1}^t = \frac{E_{N(mq-1,\sigma^2_q)}[\theta t]}{\gamma V_{N(mq-1,\sigma^2_q)}[\theta t]} \) with \( m_q-1 = \theta t_{t+q-1} \) 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 \( d_{t+q-1-K:t+q-1} \) 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 \( d_{t+q-1-K:t+q-1} \):

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

(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 (24),

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

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

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

56
where \( m_{q-2} = \sigma_{q-2}(\frac{\theta_{t+q-2}}{\sigma^2} - B) \) and \( \sigma^2_{q-2} = \frac{\sigma^2}{2c} \). Hence, by lemma B.1

\[
x_{t+q-2}^t = \frac{E_N(m_{q-2}, \sigma^2_{q-2})[s_{t+q-1}]}{\gamma VN(m_{q-2}, \sigma^2_{q-2})[s_{t+q-1}]}.
\]

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

\[
C \equiv \frac{1}{2}(R\gamma)^2((1 + \beta_0)\sigma_{q-2})^2(\delta_0(q - 2))^2
\]

\[
B \equiv (R\gamma)^2((1 + \beta_0)\sigma_{q-2})^2\left(\delta(q - 2) + \sum_{j=1}^{K} \delta_k(q - 2)d_{t+q-2-j}\right)\delta_0(q - 2)
\]

\[
A \equiv \frac{1}{2}(R\gamma)^2((1 + \beta_0)\sigma_{q-2})^2\left(\delta(q - 2) + \sum_{j=1}^{K} \delta_k(q - 2)d_{t+q-2-j}\right)^2.
\]

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

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

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

For \( k = K \), it follows from equations (45) and (49)

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

therefore \( \beta(K) = 0 \) which implies that \( \beta_k = 0 \) and \( \beta(K - 1) = -R\beta_{K-1} \) and \( \delta_k(\text{age}) = 0 \) for any \( \text{age} \).

For \( k = K - 1 \), by equations (45) and (48), we have

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

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

By induction, for any \( k \in \{q, ..., K - 2\} \), taking \( \beta_{k+1} = 0 \), it follows by equations (45)
and (48), 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 \( age \in \{q, \ldots, K\} \).

\[ \square \]

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 (70) and (71) and (69) in Online Appendix OA.2.

**Step 1.** By equation (69),
\[
2R \gamma (1 + \beta_0)^2 \sigma^2 = \alpha \left( 1 - R \right) \left[ R + \frac{\sigma^2}{s^2} - \frac{\{(1 + \beta_0)w(0, \lambda, 0) + \beta_1 - R \beta_0\}}{1 + \beta_0} \right].
\]

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 (71) 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 (70) to hold it must be that \( 1 - \frac{l(1, 1)}{R (1 + \beta_0)^2} < 0 \); but
\[
1 - \frac{l(1, 1)}{R (1 + \beta_0)^2} = 1 - \frac{1}{R} \left( 1 - w_0 \right) + \frac{\beta_1}{1 + \beta_0} > 0 \tag{50}
\]
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 (71), we know
\[
\frac{(1 + \beta_0)(1 - w_0)}{\phi + R} = \beta_1.
\]
We plug this into equation (70) and obtain
\[
\phi \left( -\beta_0 R + \frac{(1 + \beta_0)(1 - w_0)}{\phi + R} \right) + R \left[ \frac{(1 + \beta_0)(1 - w_0)}{\phi + R} + (1 + \beta_0) w_0 - \beta_0 R \right] + \ldots
\]
\[
+ \left[ 1 + \beta_0 - \frac{\phi(1 - w_0)(1 + \beta_0 - \beta_0 \phi R - \beta_0 R^2 + (1 + \beta_0)(\phi + R - 1) w_0)}{(\phi + R)^2} \right] = 0.
\]

58
Note that this is a linear equation in $\beta_0$. Therefore,

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

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 $R \geq 2(1 - w_0)$.

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 $t$:

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

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

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

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

**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 $\frac{\partial x_{t-1}^t}{\partial d_t} = \frac{l(0,1)}{\gamma (1 + \beta_0)^2 \sigma^2} = - \frac{\partial x_t^t}{\partial d_t}$, and $\frac{\partial x_{t-1}^t}{\partial d_{t-1}} = \delta_1(0) = - \frac{R\beta_1}{R^2 (1 + \beta_0) \sigma^2}$. 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 $\delta_1(0) = - \frac{R\beta_1}{R^2 (1 + \beta_0) \sigma^2} < 0$. So it only remains
to show that \( l(0,1) < 0 \). To show this, note that from the equilibrium condition (70) we have

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

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 \( l(0,1)/(1 + \beta_0) \) 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, \( F(.) \) 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 re-write the problem as follows:

\[
x^* = \arg\max_x -axE[z] + \frac{1}{2}a^2x^2V[z]
\]

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

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 \( j \in \{0, \ldots m\} \). For \( m > j_0 \), the result follows from the fact that \( w(j, \lambda, a') < w(j, \lambda, a) \) for all \( j \in \{m, \ldots a\} \) and \( F(a, a) = F(a', a') = 1 \).

OA.1.2 Proof of Supplementary Lemmas in the Appendix C

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

\[
\phi(z) = \frac{1}{\sqrt{2\pi\sigma}} \exp\{-A - Bz - Cz^2 - 0.5\sigma^{-2}(z - \mu)^2\}
\]

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

\[
\phi(z) = \frac{1}{K\sqrt{2\pi\sigma}} \exp\{-a + 0.5\mu^2 + \frac{m^2}{2\Sigma^2}\} \exp\{-\frac{z^2 - 2zm + m^2}{2\Sigma^2}\}
\]

\[
= \frac{1}{K\sqrt{2\pi\sigma}} \exp\{-a + 0.5\sigma^{-2}\mu^2 + \frac{m^2}{2\Sigma^2}\} \exp\{-\frac{(z - m)^2}{2\Sigma^2}\} = \frac{1}{\sqrt{2\pi\Sigma}} \exp\{-\frac{(z - m)^2}{2\Sigma^2}\}
\]
**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 $W_n^n$:

$$\max_x E_{n+q-1}^n [-\exp (-\gamma (W_{n+q-1}^n + xs_{n+q}))] = \max_x E_{n+q-1}^n [-\exp (-\gamma xs_{n+q})]$$  (53)

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

$$\max_x x E_{n+q-1}^n [s_{n+q}] - \frac{1}{2} \gamma^2 x^2 (1 + \beta_0)^2 \sigma^2.$$  (54)

And therefore, by FOC:

$$x_{n+q-1}^n = \frac{E_{n+q-1}^n [s_{n+q}]}{\gamma \sigma^2 (1 + \beta_0)^2}.$$  (56)

**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 his 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 $s^2 = \frac{\sigma^2}{2C + 1}$. Therefore, the maximization problem becomes:

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

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

2
\[ \tilde{\sigma}(m, s^2)^2 = V_{N(m, s^2)}[h(z)], \] by Lemma B.1, we know that
\[
\begin{align*}
\sum \max_{x} & E[-\exp\{-A - Bz - Cz^2\} \exp\{-axh(z)\}] = \frac{\tilde{\mu}(m, s^2)}{a\tilde{\sigma}(m, s^2)^2} \\
\max_{x} & E[-\exp\{-A - Bz - Cz^2\} \exp\{-axh(z)\}] = -\frac{1}{\sqrt{2\pi}C + 1} \exp\left[-A - 0.5 \left(\frac{\mu^2}{\sigma^2} - \frac{m^2}{s^2}\right)\right] \\
& \times \exp\left[-0.5 \frac{\tilde{\mu}(m, s^2)^2}{\tilde{\sigma}(m, s^2)^2}\right].
\end{align*}
\]

Let \( t \mapsto \rho(t) \equiv \gamma t^2 \) and let
\[
\Lambda(d_{t-K}, \ldots, d_t) \equiv \alpha(1 - R) + \sum_{k=1}^{K} \beta_k d_{t+1-k} - R \sum_{k=0}^{K} \beta_k d_{t-k}
= \alpha(1 - R) + \sum_{j=0}^{K-1} \beta_{j+1} d_{t-j} - R \sum_{k=0}^{K} \beta_k d_{t-k} = \alpha(1 - R) + \sum_{k=0}^{K} \beta(k) d_{t-k}
\]
with \( \beta(k) = \beta_{k+1} - R \beta_k \) for \( k \in \{0, \ldots, K - 1\} \) and \( \beta(K) = -R \beta_K \). We use \( \Lambda_r \) to denote \( \Lambda(d_{t-K}, \ldots, d_t) \).

**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 \( x_t^{t-q+1} = E_t^{t-q+1}[s_{t+1}] \gamma((1+\beta_0)\sigma)^2 \). Therefore,
\[
\begin{align*}
\delta(q) &= \delta_k(q) = 0, \quad \forall k \in \{0, \ldots, K\} & (57) \\
\delta(q-1) &= \frac{\alpha(1-R)}{\gamma((1+\beta_0)\sigma)^2}, \quad \delta_k(q-1) = \frac{(1+\beta_0)w(k, \lambda, q-1) + \beta(k)}{\gamma((1+\beta_0)\sigma)^2}, \quad \forall k \in \{0, \ldots, q-1\} & (58) \\
\delta_k(q-1) &= \frac{\beta(k)}{\gamma((1+\beta_0)\sigma)^2}, \quad \forall k \in \{q, \ldots, K\}. & (59)
\end{align*}
\]

We also know from Lemma B.1 that
\[
V_t^{q-1}(d_{t-K}, \ldots, d_t) = -\exp\left(-\frac{1}{2} \left(d_t \delta_0(q-1) + \delta(q-1) + \sum_{j=1}^{K} \delta_k(q-1)d_{t-j}\right)^2 \gamma^2((1+\beta_0)s_{q-1})^2\right)
\]
where \( s_{q-1} = \sigma^2 \). Henceforth, we denote \( V_t^{q-1}(d_{t-K}, \ldots, d_t) \) by \( V_t^{t-q+1} \). In particular, \( V_t^{t-q+1} = V_{t+1}^{t-q+2} = V_{t+1}^{t-q+1}(d_{t+1-K}, \ldots, d_{t+1}) \).

**Step 2.** We now derive the risky demand and continuation value for generation aged
The problem of generation aged \( q - 2 \) at time \( t \) is given by,

\[
\max_x E_t^{t-q+2} \left[ V_{t+1}^{t-q+2} \exp \left( -\gamma R x_{t+1} \right) \right].
\]

By the calculations in step 1, and using \( \Lambda_t \) as defined in (57), this problem becomes:

\[
V^{q-2}(d_{t-K}, \ldots, d_t)
= \max_x E_t^{t-q+2} \left[ -\exp \left( -\frac{1}{2} \left( x_t^{q-1} \right)^2 \gamma^2 ((1 + \beta_0)s_{q-1})^2 - \gamma R x_t ((1 + \beta_0)d_{t+1} + \Lambda_t) \right) \right].
\]

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

\[
-\frac{1}{2} \left( d_{t+1}\delta_0(q - 1) + \delta(q - 1) + \sum_{j=1}^K \delta_k(q - 1)d_{t+1-j} \right)^2 \gamma^2 ((1 + \beta_0)s_{q-1})^2
\]

\[
= -\frac{1}{2} \gamma^2 ((1 + \beta_0)s_{q-1})^2 \left( \delta(q - 1) + \sum_{j=1}^K \delta_k(q - 1)d_{t+1-j} \right)^2
\]

\[
- \gamma^2 ((1 + \beta_0)s_{q-1})^2 \left( \delta(q - 1) + \sum_{j=1}^K \delta_k(q - 1)d_{t+1-j} \right) \delta_0(q - 1)d_{t+1}
\]

\[
- \frac{1}{2} \gamma^2 ((1 + \beta_0)s_{q-1})^2 (\delta_0(q - 1))^2 d_{t+1}^2,
\]

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: \( d_{t+1} \sim N(\theta_t^{t-q+2}, \sigma^2) \).

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

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

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

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

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

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 $m_t \equiv -s_{q-2}^2 B + s_{q-2}^2 \sigma^2 \theta^{1-q+2}$. Thus, invoking lemma B.1 for this distribution for dividends and $a = R\gamma(1+\beta_0)$ implies that

$$V^{q-2}(d_{t-K}, \ldots, d_t) \sim \exp \left( -\frac{1}{2} \left( \frac{d_{t}^{q-2}}{(R\gamma)^2((1+\beta_0)s_{q-2})^2} \right) \right)$$

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

(the symbol $\sim$ 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

$$V_t^{age-1} = V^{age+1}(d_{t+1-K}, \ldots, d_{t+1})$$

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

The maximization problem is given by:

$$V^{age}(d_{t-K}, \ldots, d_t) \equiv \max_x E^{t-\text{age}}_t \left[ V_{t+1}^{age-1} \exp \left( -\gamma R^{q-1-age}x((1+\beta_0)d_{t+1} + \Lambda_t) \right) \right]. \quad (63)$$

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

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

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

\[
\delta_k(age) = \frac{(1 + \beta_0)s_{age}^2(\sigma^{-2}w(k,\lambda,age) - [(R^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2\delta_k(age+1)\delta_0(age+1)) + \beta(k)}{R^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2}, \quad k \in \{0, ..., q-1\},
\]

\[
\delta_K(age) = \frac{\beta(K)}{R^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2}.
\]

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

\[
V_{age}(d_{t-K},...,d_t) \sim -\exp \left( -\frac{1}{2} \left( x_t^{-(age)} \right)^2 (R^{q-1-(age)}\gamma)^2((1+\beta_0)s_{age})^2 \right)
= -\exp \left( -\frac{1}{2} \left( d_0\delta_0(age) + \delta(age) + \sum_{j=1}^{K} \delta_k(age)d_{t-j} \right)^2 (R^{q-1-(age)}\gamma)^2((1+\beta_0)s_{age})^2 \right).
\]

**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 (70) 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 \( x_t^{q-1} = \frac{E_t^{q-1}[\sigma_{t+1}]}{(1+\beta_0)s_{age}} \). Therefore, given
our guess for prices and Lemma 7.2, we have:

\[
x_t^{t-1} = \frac{E_t^{t-1}[d_{t+1} + p_{t+1} - p_t R]}{\gamma(1 + \beta_0)\sigma^2} = \frac{(1 + \beta_0)\theta_t^{t-1} + \alpha(1 - R) + (\beta_1 - R\beta_0)d_t - R\beta_1d_{t-1}}{\gamma(1 + \beta_0)\sigma^2}
\]  \hspace{1cm} (64)

since \(\theta_t^{t-1} = w_0d_t + (1 - w_0)d_{t-1}\), we obtain equation (51), 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

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

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 who is investing for the last time on the market as follows:

\[
V_t^{t-1} = -\exp(-A_t - B_t d_t - Cd_t^2)
\]  \hspace{1cm} (66)

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 this results to obtain \(V_{t+1}^t\), the problem of the young generation at time \(t\) is given by:

\[
\max_x E^t_x \left[V_{t+1}^t \exp\left(-\gamma R x s_{t+1}\right)\right]
\]  \hspace{1cm} (67)

From Lemma C.3:

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

Where,

\[
\bar{\mu}(m, s^2) = E_{\Phi(m, s^2)}[h(z)] = \alpha(1 - R) + (\beta_1 - R\beta_0)d_t - R\beta_1d_{t-1} + (1 + \beta_0)m
\]

\[
\tilde{\sigma}(m, s^2)^2 = V_{\Phi(m, s^2)}[h(d_{t+1})] = (1 + \beta_0)^2 s^2
\]

with \(m = \frac{\theta_t^{t-1} - \sigma^2 B_{t+1}}{2C\sigma^2} + 1\), \(s^2 = \frac{\sigma^2}{2C\sigma^2 - 1}\). 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 (52) and the respective \(\delta s\).
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 $\alpha$, $\beta_0$ and $\beta_1$.

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

$$ p_t = \alpha + \beta_0 d_t + \beta_1 d_{t-1} \quad \forall t \in \mathbb{Z} \quad (68) $$

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

$$ 0 = \alpha (1 - R) \left[ R + \frac{\sigma^2 s^2}{\gamma (1 + \beta_0)^2 \sigma^2} - \frac{l(0, 1)}{1 + \beta_0} \right] - 2R \gamma (1 + \beta_0)^2 \sigma^2 \quad (69) $$

$$ 0 = l(0, 1) + \frac{1}{R} \sigma^2 \left( \beta_1 - R \beta_0 \right) + \frac{1}{R} (1 + \beta_0) \left( 1 - \frac{l(1, 1)l(0, 1)}{(1 + \beta_0)^2} \right) \quad (70) $$

$$ 0 = l(1, 1) - \frac{\sigma^2}{s^2} \beta_1 \quad (71) $$

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:

$$ x_t^{t-2} = 0 \quad (72) $$

$$ x_t^{t-1} = \frac{E_t^{-1} [s_{t+1}]}{\gamma R (1 + \beta_0) \sigma^2} = \frac{\alpha (1 - R) + l(0, 1) d_t + l(1, 1) d_{t-1}}{\gamma (1 + \beta_0)^2 \sigma^2} \quad (73) $$

$$ x_t^t = \frac{E_t \Phi(m, s^2) [s_{t+1}]}{\gamma R (1 + \beta_0) s^2} = \frac{\alpha (1 - R) + (\beta_1 - R \beta_0) d_t - R \beta_1 d_{t-1} + (1 + \beta_0) m}{\gamma R (1 + \beta_0)^2 s^2} \quad (74) $$

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

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

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

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:

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

\[ s^2 = \frac{\sigma^2}{2 \frac{\rho(0,1)^2}{(1+\beta_0)^2} \sigma^2 + 1} = \frac{(1 + \beta_0)^2}{\rho(0,1)^2 + (1 + \beta_0)^2}. \]

Plugging this in the expression for \( x_t \), it follows that

\[ x_t' = \frac{\alpha (1 - R) + (\beta_1 - R \beta_0) d_t - R \beta_1 d_{t-1} + (1 + \beta_0) \frac{s^2}{\sigma^2} \left[ -\frac{\alpha (1 - R) l (0, 1)}{(1 + \beta_0)^2} + \left( 1 - \frac{l(1,1)l(0,1)}{(1 + \beta_0)^2} \right) d_t \right]}{\gamma R (1 + \beta_0)^2 s^2} \]

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

By Market clearing:

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

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

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

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

which implies

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

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

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

\[ + \left[ l (0, 1) + \frac{1}{R} \frac{s^2}{\sigma^2} \beta_1 - R \beta_0 + \frac{1}{R} (1 + \beta_0) \left( 1 - \frac{l(1,1)l(0,1)}{(1 + \beta_0)^2} \right) d_t + \left[ l (1, 1) - \frac{s^2}{\sigma^2} \beta_1 \right] d_{t-1}. \]
Therefore \( \{\alpha, \beta_0, \beta_1\} \) solve the following system of equations:

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

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

\[
0 = l(1, 1) - \frac{\sigma^2}{s^2} \beta_1 
\]

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

**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:

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

and after simple algebra,

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

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

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

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

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

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

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

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

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

The latter expressions reveal that the total mass of agents in the market is reflected only in the price constant, while the fraction of young people in the market determines the dividend loadings \( \beta_0 \) and \( \beta_1 \). 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.