

Asset Pricing with Experience Effects*

Ulrike Malmendier[†]
UC Berkeley, NBER, and CEPR

Demian Pouzo[‡]
UC Berkeley

Victoria Vanasco[§]
Stanford University

PRELIMINARY AND INCOMPLETE - DO NOT DISTRIBUTE

March 29, 2016

Abstract

How does the experience of a financial crisis and stock-market fluctuations alter the dynamics of financial markets? Recent evidence suggests that individuals overweight personal experiences of macroeconomic shocks when forming beliefs about risky outcomes and making investment decisions. We propose a simple OLG model of experience-based learning. Risk averse investors invest in a ‘Lucas tree’ and a risk-free asset. They form beliefs (1) based on data observed during their lifetimes so far and (2) exhibiting recency bias, the two components of experience effects. We show that, in equilibrium, prices depend on past dividends, but only those observed by the generations that are alive, and they are more sensitive to more recent dividends. Younger generations react more strongly to recent experiences than older generations and, hence, have higher demand for the risky asset than the old in good times, and lower demand in bad times. The model generates predictions for stock-market dynamics and trading volume: First, a recent crisis will increase the average age of stock-market participants, while periods of stock-market boom have the opposite effect. Second, the stronger the disagreement across generations (e.g. after a recent shock), the higher is the trade volume. We provide stylized facts from the *Survey of Consumer Finances* consistent with these predictions.

*We thank workshop participants at Pompeu Fabra, Stanford, UC Berkeley, as well as the NBER EFG Behavioral Macro, SITE (Psychology and Economics segment), SFB TR 15 (Special Research Program TR15, Tutzung, Germany) conferences for helpful comments. We also thank 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

1 Introduction

Economists and policy-makers alike have long wrestled with understanding the long-lasting effects of financial crises and other macroeconomic shocks. In the case of the Great Depression, [Friedman and Schwartz \(1963\)](#) argue that the experience of that time created a “mood of pessimism that for a long time affected markets.” In the case of the recent financial crisis, [Blanchard \(2012\)](#) maintains that “The crisis has left deep scars, which will affect both supply and demand for many years to come.”

The notion that longer-lasting crisis effects alter the dynamics of markets is consistent with growing empirical evidence on experience effects. This evidence suggests that personal experiences of macroeconomic shocks leave an imprint on individuals’ attitudes and willingness to take risk. For example, [Alesina and Fuchs-Schundeln \(2007\)](#) relate the personal experience of living in (communist) Eastern Germany to political attitudes post-reunification. [Weber, Bockenholt, Hilton, and Wallace \(1993\)](#) and [Hertwig, Barron, Weber, and Erev \(2004\)](#) show how doctors’ experience affect their future diagnoses. On the finance side, [Malmendier and Nagel \(2011\)](#) show that stock-market experiences predict future willingness to invest in the stock market, and [Kaustia and Knüpfer \(2008\)](#) argue the same for IPO experiences.

In this paper, we provide a theoretical framework that captures the long-term effects of personal experiences on stock-market participation and portfolio decisions of different cohorts in an economy. We derive implications for the predictability of stock prices, the composition of stock investors, as well as trade volume. The model illustrates that a deeper understanding of the influence of past experiences is important not only to improve the micro-modeling of financial risk-taking, but also for our understanding of the aggregate dynamics of financial decision-making and the long-run effects of macro-shocks.

We develop a stylized overlapping generations (OLG) general equilibrium model in which agents form their beliefs by overweighting their own experiences. Investors 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 to maximize their final wealth. Consumption takes place

at the end of their lives for tractability, as it is standard in much of the literature (see [Vives \(2008\)](#)). The risky asset (Lucas Tree) 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 risky dividends, but they can learn about it by observing the history of realized dividends. As in previous literature, we focus on affine equilibria, i.e., equilibria wherein prices are affine functions of current and past dividends.

The novel feature of the model is that investors are *experience-based learners*. Building on the psychology evidence on availability bias ([Tversky and Kahneman \(1974\)](#)), we model experience-based learners as overweighting outcomes they have experienced in their lives. Specifically, they (i) only use data observed during their lifetimes so far, and (ii) over-weight more recent observations when forming beliefs. Note that these agents observe the entire history of dividends; there is no asymmetric information. But they choose not to use it to form their beliefs. This assumption captures, in a stylized manner, the availability bias underlying experience-based learning, with the weights reflecting the recency bias. In generalizations of the model, and in empirical implementations, investors use more historical data, also from periods prior to their lifetimes. The essential feature is that they *overweight* their lifetime experiences.

Experience-based learning generates long-lasting effects of economic shocks on equilibrium prices and asset demand through direct and indirect channels: First, shocks to dividends shape agents beliefs about future dividends, thereby affecting equilibrium prices and demands. Second, investors who are confronted with the same macroeconomic shock in the current period but who have had different experiences in their lives so far (i.e., different life-time sequences of dividends) have different reactions to the shock. This differential reaction to “booms” and “recessions” is an important determinant of equilibrium quantities such as trade volume and the demographic composition of the stock-market investor base. Third, depending on the foresight of experience-based learners, the different investment horizons of different generations also affect anticipated future trading behavior. If agents overweight

personal experiences but fully understand the model and everyone’s belief formation process, experience-based learning introduces a correlation between future returns and continuation values: Agents understand that shocks will generate disagreements that everybody will exploit in the market. In response to this, and for a given set of beliefs, agents distort their portfolio decisions. We will refer to this last channel as the *hedging motive*, and distinguish it from the channel of belief heterogeneity.

The model is stylized, which allows us to fully isolate the forces introduced by the presence of experience-based learners. In the comparison case of “correct expectations,” i.e., the case where agents know the true mean of dividends, the model setup generates the standard result that equilibrium prices are constant, and individual demands for the risky asset only change as a response to the change in horizon of different cohorts. Hence, any departure from constant equilibrium prices and purely horizon-dependent trade levels can be cleanly attributed to experience-based learning.

For ease of exposition, we first solve the model where agents are myopic. That is, agents take actions to maximize only the current per-period payoff. By doing so the agents omit the correlation between next period’s risky payoff and the value function. This simplified model directly generates the first-order effect of experience-based learning: Young generations react more strongly than old ones to current dividend shocks. The key intuition is simple: the younger generation has experienced a shorter life so far and will thus put a higher weight on the current realization in updating their beliefs (*beliefs effect*).

We also solve the non-myopic case, in which agents solve a standard dynamic portfolio problem: In every period, they optimize their portfolio to maximize last period’s consumption. In a rational-expectations framework, we can partition agents’ multi-period investment problem into a sequence of one-period ones (Vives (2008)). Under experience-based learning, instead, future beliefs and portfolio decisions and, as a result, prices in the distant future, depend on current dividends. Thus, investors’ wealth in the distant future is correlated with next period’s returns and the simplification to a sequence of static problem no longer ap-

plies. However, exploiting the CARA-Gaussian setup, 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; but for each of these, the probability distribution differs from the original one.

In this non-myopic case, we can decompose agents' demands for the risky asset into a beliefs effect, a hedging effect, and a horizon effect. The *beliefs effect* matches the demand of myopic agents. Thus, all the forces that are present in the myopic case are captured by this term. The dynamics inherent to the multi-period problem generate the two other effects. The *hedging effect* captures that agents anticipate that they will learn about the risky asset from future dividends. In order to hedge their exposure to changes in beliefs, they distort their portfolio decisions. The *horizon effect* captures that young agents react less aggressively (in their beliefs) to a change in dividends due to their longer remaining investment horizon. In other words, their longer investment horizon makes them behave in a more risk-averse fashion. We show that the beliefs effect always dominates. Hence, the qualitative results presented in the myopic case hold in the non-myopic world.

The link between past dividends and beliefs has important implications for price volatility. The resulting price volatility goes above and beyond the volatility of the assumed dividend process. Moreover, if we let recency bias go to its extreme (i.e., $\lambda \rightarrow \infty$), the weight placed on the most recent observation will assume its maximum value and price volatility will also be at its maximum, reflecting that all generations fully adjust their beliefs to the most recent change. For the same reason, autocorrelation of prices vanishes. Both features are a direct result of experience-based learning. If agents know the true mean of dividends (or their beliefs converge to the truth), prices are constant.

We also derive the implications of experience-based expectations for stock-market investment and trade volume. We show that experience-based learning affects the demographic composition of stock-market investors. A positive shock induces younger cohorts to partic-

ipate relatively more, while a negative shock tilts the composition towards older cohorts. Moreover, disagreements between cohorts generates positive trade volume in equilibrium. The mechanism is intuitive: an increase (decrease) in dividends induces trade since young agents become more optimistic (pessimistic) than old agents, and disagreement generates gains from trade. The reaction of trade volume to changes in dividends is tightly linked to the extent of recency bias present in agent's beliefs formation, captured by λ . When λ is high, the recency bias is strong, and thus the reaction of trade volume is smaller, since all agents pay a lot of attention to the change in dividends. Since higher λ also generates higher price volatility we obtain the following prediction: economies where prices are highly volatile and autocorrelated should be associated with smoother trade volumes and vice-versa.

Finally, we present stylized facts on portfolio decisions and trade volume, consistent with the predictions generated by our model. Using the representative sample of the *Survey of Consumer Finance*, CRSP, and historical data on stock-market performance, we are able to show that both stock market participation (extensive margin) and the fraction of liquid assets invested in the stock market (intensive margin) differ across cohorts and over time in the same way lifetime stock-market experiences differ. In terms of abnormal trade volume, we show that the detrended turnover ratio is strongly correlated with differences in lifetime experiences of stock-market returns across cohorts.

In summary, our paper provides a simple formalization reveals of experience effects. It generates testable implications for individual financial decision-making and the resulting stock-market dynamics, including the long-term effects of crisis experiences.

Related Literature. Our findings are closely related to a growing literature arguing that financial crises and macroeconomic shocks have long-run effects. As alluded to earlier, [Friedman and Schwartz \(1963\)](#) discuss at length how the Great Depression created a long-lasting shift toward pessimism about economic conditions and economic stability. More recently, [Delong and Summers \(2012\)](#), argue that recessions such as the Great Recession of 2008-2009 leave scarring effects, or what they term ‘hysteresis effects.’

The empirical literature on ‘experience effects’ rationalizes these long-run effects by showing that personal experiences of macroeconomic shocks leave a lasting imprint and significantly affect individuals’ decision-making over lifetimes. For example, [Malmendier and Nagel \(2011\)](#) show that people who live through different stock-market histories differ in their level of risk taking in the stock market. They find that individuals who have experienced low stock market returns report lower willingness to take financial risk, are less likely to participate in the stock market, invest a lower fraction of their liquid assets in stocks if they participate, and are more pessimistic about future stock returns. [Malmendier and Shen \(2015\)](#) show that individual experiences of macroeconomic unemployment conditions strongly affect consumption behavior — households who have experienced higher unemployment conditions during their lifetime spend significantly less and are more likely to use coupons and allocate expenditure toward lower-end products. Moreover, [Malmendier and Nagel \(2013\)](#) show that experience effects work through the channel of beliefs. In the context of inflation expectations, they show that differences in life-time experiences of inflation strongly predict differences in individuals’ subjective inflation expectations. Empirical findings from these papers form the foundation for our model on learning from experience effects.

Our modeling approach builds on a large literature of learning models in asset pricing. For instance, [Barsky and DeLong \(1993\)](#), [Timmermann \(1993\)](#), [Timmermann \(1996\)](#), and [Adam, Marcet, and Nicolini \(2012\)](#) study the implications of learning for stock-return volatility and predictability. [Cecchetti, Lam, and Mark \(2000\)](#) construct a Lucas asset-pricing model with infinitely-lived agents where the representative agent’s subjective beliefs about endowment growth are distorted. More 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. The main difference to our paper is that, in our setup, agents are not Bayesian and live for a finite number of periods. Consequently, observations during the agents’ life-time have a non-negligible effect on their beliefs. We think that this feature provides an alternative modeling device that allow us to capture Friedman

and Schwartz’s idea that economic events, such as the Great Depression, shape the attitude of agents towards financial markets in the future.

Finally, [Ehling, Graniero, and Heyerdahl-Larsen \(2015\)](#) also explore the role of experience effects in portfolio decisions in a complete markets setting. Differently from our paper, they do not aim to capture ‘experience effects’ in the sense of the empirically observed pattern in [Malmendier and Nagel \(2011\)](#), which involves a declining weighting function and thus recency bias. Instead, they are interested in the pure effect of individuals restricting their use of data to their lifetimes. Similar to the ‘Bayesian Learners from Experience’ in our analysis, agents in their paper start from a given prior which they update only using lifetime observations.¹ The authors use this setting to carefully develop a theoretical underpinning for ‘trend-chasing’ and the negative relationship between beliefs about expected returns and realized future returns, as shown by [Greenwood and Shleifer \(2014\)](#). Our paper focuses on a somewhat different set of empirical predictions, including predictions about trade volume.

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

2 The Baseline Model

2.1 The Lucas-Tree Economy

Consider an infinite horizon economy $t \in \mathbb{Z}$ 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 $q \in \{1, 2, 3, \dots\}$. Hence, there are $q + 1$ generations alive at any t . The generation born at

¹[Ehling, Graniero, and Heyerdahl-Larsen \(2015\)](#) also choose a rather different, and complementary, set-up: complete markets in a continuous-time framework with log utility, as opposed to our discrete-time Lucas economy with CARA utility. As a result, their set-up is not designed to differentiate the hedging motive from other portfolio choice motives (due to the use of log utility). Naturally, their theoretical analysis of the portfolio problem of the agents does not revolve around the issue of characterizing a dynamic problem as a sequence one-period problems with a slanted distribution for dividends.

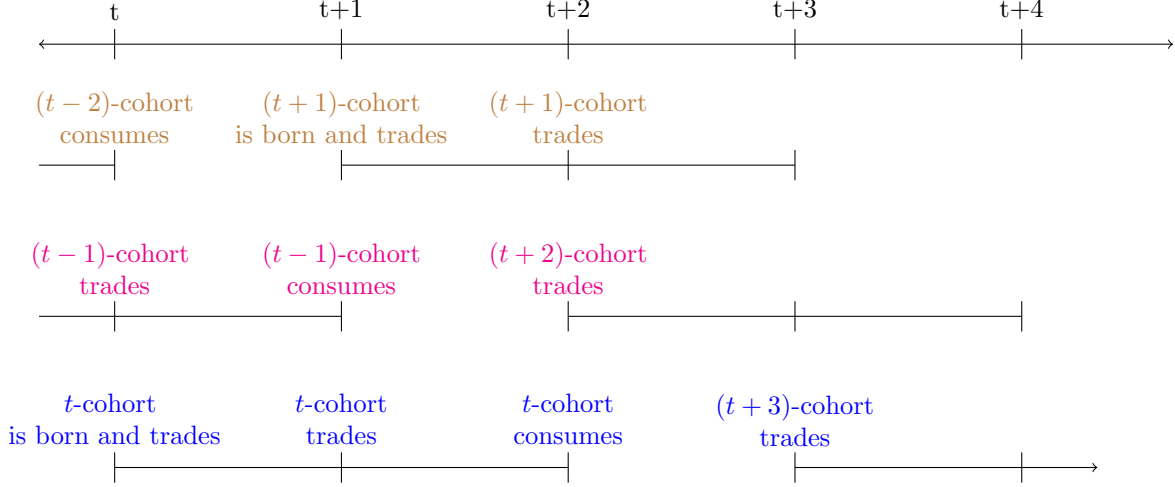


Figure 1: A timeline for an economy with two-period lived generations, $q = 2$.

time $t = n$ is called generation n . Each generation has a mass of q^{-1} identical agents.

Agents have CARA preferences with risk aversion γ . They are born with no endowment and can transfer resources across time by investing in financial markets (i.e., by trading). For the baseline model, we assume that agents are myopic and as a result maximize their per period utility. In Section 6, we remove this assumption and show that the main mechanisms continue to be present. Figure 1 illustrates the timeline of this economy for two-period lived generations ($q = 2$).

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 $t, \forall t \in \{0, 1, 2, \dots\}$, and there is a riskless asset that is in perfectly elastic supply and pays $r > 1$ at all times.

For each generation $n \in \{0, 1, 2, \dots\}$ and any $t \in \{n, \dots, n + q\}$, the budget constraint is given by

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

where W_t^n denotes the wealth of generation n at time t , x_t^n is the amount invested in the risky asset (units of Lucas Tree output), a_t^n is the amount invested in the riskless asset, and

p_t is the price of one unit of the risky asset at time t . As a result, wealth next period is:

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

To simplify on notation, we define $s_{t+1} \equiv p_{t+1} + d_{t+1} - p_t r$ as the net payoff received in $t + 1$ from investing in one unit of the risky asset at time t . Note that $p_{t+1} + d_{t+1}$ is the payoff of the risky asset in $t + 1$ and rp_t is the (opportunity) cost of investing in one unit of the risky asset at time t . Using this notation, $W_{t+1}^n = x_t^n s_{t+1} + rW_t^n$.

We assume that agents are myopic and, as a result, maximize their per-period utility. This assumption simplifies the maximization problem considerably, and highlights the main determinant of portfolio choice generated by experience-based learning. Later, in Section 6, we will remove this assumption and show that the main mechanism is present.

For a given initial wealth level, W_n^n , the myopic problem of a generation n is to choose $(x_t^n)_{t=n}^{n+q-1}$ such that, for each time $t \in \{n, \dots, n + q - 1\}$,

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

where $E_t^n [\cdot]$ denotes the expectation computed with beliefs of generation n at time t . Note that when x_n^t is negative, generation n is short selling units of the Lucas tree.

2.2 Experience-Based Learning

To model uncertainty about fundamentals, we assume that agents do not know the true mean of dividends θ and use past observations to estimate the mean. To keep the model tractable, we assume that σ^2 , the variance of dividends, is known at all times. In this framework, experienced-based learning (EBL) means that agents overweight observations received during their lifetimes when forecasting dividends. In our model, we assume that agents *only* use observations realized during their lifetimes. We make this assumption for simplicity; all we need for our results to hold is that the earlier history is heavily discounted when agents

form their beliefs. Note that agents have full information, i.e., observe the entire history of dividends. However, they *choose not to use* observations outside their lifetimes. This also implies that prices do not add any additional information since the history of dividends is available to the agents. While it is possible to add private information and learning from prices to our framework, adding these (realistic) feature would complicate matters without necessarily adding new intuition.

We also note that that 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 and to learning. We consider this to be an interesting line to explore in the future.

The parameter of interest that agents want to learn is θ , the mean of dividends. The belief of generation n at period t is given as follows:

$$E_t^n[\theta] = \sum_{k=0}^{age} w(k, \lambda, age) d_{t-k}, \text{ with } age = t - n \quad (4)$$

where $w(k, \lambda, age)$ denotes the weight an agent aged age assigns to the dividend observed k periods earlier, and λ parametrizes these weights, as described in detail below. Note that $\sum_{k=0}^{age} w(k, \lambda, age) = 1, \forall age \in \{0, 1, 2, \dots\}$.

We now characterize the probability measure implied by these weights (which would also allows us to extend the idea of experience-based learning to objects other than the mean). As in Malmendier and Nagel (2010), we consider the precision-weighted average of dividends

realized during the agent's lifetime. Given a realization of past dividends $(d_\tau)_{\tau=0}^t$, let²

$$\mathbb{P}_t^n(d) = \sum_{k=0}^{t-n} 1_{\{d_{t-k}\}}(d)w(k, \lambda, t-n), \quad \forall d \in \mathbb{R} \quad (5)$$

where, for any $k \leq \text{age} = t - n$,

$$w(k, \lambda, \text{age}) = \frac{(\text{age} + 1 - k)^\lambda}{\sum_{k'=0}^{\text{age}} (\text{age} + 1 - k')^\lambda} \quad (6)$$

be the *experience-based empirical probability measure* of generation n at period $t \in \{n, \dots, n + q\}$. Furthermore, $E_t^n[\cdot]$ denotes the expectation with respect to the discrete empirical probability measure \mathbb{P}_t^n .

That is, this probability measure assigns weights $(w(k, \lambda, t-n))_{k=0}^{\text{age}}$ to the observations during an agent's lifetime, and zero to all other observations. Within their lifetimes, agents put weight $\frac{(\text{age}+1)^\lambda}{\sum_{k'=0}^{\text{age}} (\text{age}+1-k')^\lambda}$ on the most recent observations, $\frac{(\text{age}+1-1)^\lambda}{\sum_{k'=0}^{\text{age}} (\text{age}+1-k')^\lambda}$ on the previous one, and so on. The parameter λ regulates the relative weight the most recent observations receive. For $\lambda > 0$, more recent observations receive relative more weight, whereas for $\lambda < 0$ the opposite holds. Note that the denominator in 6 is a normalizing constant that depends only on age and λ . Thus, one can rewrite $w(k, \lambda, \text{age})$ more concisely as

$$w(k, \lambda, \text{age}) = c_{\text{age}, \lambda} (\text{age} + 1 - k)^\lambda \quad (7)$$

where $c_{\text{age}, \lambda} = 1 / \sum_{n=1}^{\text{age}+1} n^\lambda$. Here are three examples:

Example 2.1 (Linearly Declining Weights, $\lambda = 1$). *For $\lambda = 1$, weights decay linearly, i.e., for any $0 \leq k, k + j \leq \text{age}$,*

$$w(k + j, 1, \text{age}) - w(k, 1, \text{age}) = -c_{\text{age}, 1} \cdot j = -\frac{2j}{(\text{age} + 1)(\text{age} + 2)}.$$

In particular, $c_{\text{age}, 1}$ is decreasing quadratically in age in this example.

²The function $x \mapsto 1_A(x)$ takes value 1 if $x \in A$, and 0 otherwise.

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 $\lambda \rightarrow \infty$, the weight assigned to the most recent observation converges to 1, all other weights converge to 0, i.e., for any $0 \leq k \leq \text{age}$

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

Applying the experience-based empirical probability measure, we can write the belief about θ of generation n at time $t \in \{n, \dots, n + q\}$, which we denote as θ_t^n , as

$$\theta_t^n = E_t^n[d] = \sum_{k=0}^{t-n} w(k, \lambda, t-n) d_{t-k}. \quad (8)$$

For example, for $q = 2$, we obtain

$$\theta_t^t = d_t \quad \text{and} \quad \theta_{t+1}^t = d_{t+1} \frac{2^\lambda}{1 + 2^\lambda} + d_t \frac{1}{1 + 2^\lambda} \quad \forall t.$$

Observe that by construction, $\theta_t^n \sim N(\theta, \sigma^2 \sum_{k=0}^{t-n} (w(k, \lambda, t-n))^2)$. Hence, θ_t^n does not necessarily converge to the truth as $t \rightarrow \infty$; it depends on whether $\sum_{k=0}^{t-n} (w(k, \lambda, t-n))^2 \rightarrow 0$, and this in turn depends how fast the weights for “old” observations decay to zero. When agents have finite lives, convergence will not occur. In addition, since separate cohorts weight different realizations differently, we should expect belief heterogeneity, driven by different experiences, at any point in time.

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

Lemma 2.1. [*Single-Crossing Property*] Let $a' < a$ and $\lambda > 0$. Then $w(j, \lambda, a) - w(j, \lambda, a')$

is an increasing function of j , for $j \in \{0, \dots, a'\}$.

Proof. See Appendix A. □

This result implies that, for $\lambda > 0$, the difference in weights, $w(\cdot, \lambda, a) - w(\cdot, \lambda, a')$ will change signs only once.

2.3 Comparison to Bayesian Learners

To better understand the experience-effect mechanism, we compare EBL agents to agents who update their beliefs using Bayes rule. We consider two sub-cases: the standard case, wherein agents use *all* the available observations ‘since the beginning of times’ (from period 0 on) to form their beliefs; and an alternative case where agents ‘learn from experience’ in the sense that they only use data realized during their lifetimes, but update their beliefs using Bayes rule. We call the former case *Full Bayesian Learning* (FBL), and in the latter case *Bayesian Learning from Experience* (BLE).

Full Bayesian Learners. All generations of FBL agents consider the whole set of observations since period 0 to form their belief. We assume that each generation t has a prior $N(m, \tau^2)$.³ The posterior mean of any generation alive at period $t + a$, γ_{t+a} , is given by

$$\gamma_{t+a} = \frac{\tau^{-2}}{\tau^{-2} + \sigma^{-2}(t+a)}m + \frac{(t+a)\sigma^{-2}}{\tau^{-2} + \sigma^{-2}(t+a)} \left(\frac{1}{t+a} \sum_{k=0}^{t+a} d_k \right)$$

That is, the belief of an FBL agent is a convex combination of the form $\alpha m + (1 - \alpha)\bar{d}$ of the prior, m , and the average of *all* observations, d_k , available to date. The key difference to EBL agents is that FBL agents do not differ in their beliefs. *All* generations alive in any given period will have the same belief about mean of dividends; different past experiences do not play a role, and hence there is no belief heterogeneity. Moreover, beliefs are non-stationary

³The analysis could be easily extended to allow heterogenous Gaussian priors across generations. The assumption of Gaussianity is also not needed but simplifies the exposition greatly.

(depend on the time period). And, as $t \rightarrow \infty$, the posterior mean eventually converges (almost surely) to the true mean.

Bayesian Learners from Experience. For Bayesian Learners from Experience the situation is different. We assume that each generation has a prior $N(m, \tau^2)$ when they are born (and not from $t = 0$). The posterior mean of generation t at period $t + a$, β_{t+a}^t , is given by

$$\beta_{t+a}^t = \frac{\tau^{-2}}{\tau^{-2} + \sigma^{-2}(a+1)}m + \frac{(a+1)\sigma^{-2}}{\tau^{-2} + \sigma^{-2}(a+1)}\theta_{t+a}^t.$$

That is, the belief of a BLE generation is a convex combination of the prior, m , and the weighted average of (only) the life-time observations, θ_{t+a}^t ; in turn this average coincides with the belief of our learners from experience with $\lambda = 0$. Thus, the only difference between a Bayesian Learner from experience and an EBL with $\lambda = 0$, is that the EBL is not born with a prior belief distribution. We see this as a strength of our framework, since we want to focus on how observations experienced by agents (as opposed to priors) shape their beliefs. If the prior is diffuse, i.e., $\tau \rightarrow \infty$, then β_{t+a}^t coincides with θ_{t+a}^t of EBL agents for $\lambda = 0$.

These two benchmark comparisons illustrate the importance and role of the main feature of our model, experience-based learning: the fact that agents only use data observed during their lifetimes and the resulting heterogeneity in beliefs. Under FBL, beliefs do not differ across agents and, eventually, will converge to the truth. Under BLE and our main approach, EBL, this is not true. Cohorts differ in their beliefs.

3 Equilibrium

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

Definition 3.1 (Equilibrium). *An equilibrium is a demand profile for the risky asset $\{x_t^n\}_{n \geq 0, t \geq 0}$, a demand profile for the riskless asset $\{a_t^n\}_{n \geq 0, t \geq 0}$ and a price schedule $\{p_t\}_{t \geq 0}$ such that:*

1. *given the price schedule, $\{(a_t^n, x_t^n) : t \in \{n, \dots, n+q-1\}\}$ solve the generation n problem,*

and

2. Market clears at all periods, i.e.,

$$1 = (q)^{-1} \sum_{n=t-q}^t x_t^n, \quad \forall t. \quad (9)$$

In this paper, we focus on the class of equilibria with *affine prices*

Definition 3.2 (Linear Equilibrium). *A linear equilibrium is an equilibrium wherein prices are an affine function of dividends. That is, there exists a $K \in \mathbb{N}$, $\alpha \in \mathbb{R}$, $\beta_k \in \mathbb{R}$ for $k \in \{0, \dots, K\}$ such that:*

$$p_t = \alpha + \sum_{k=0}^K \beta_k d_{t-k} \quad (10)$$

for all $t \geq K$. Thus, the price p_t is a linear function of the current and the last K dividends.

4 Characterization of the Linear Equilibrium

We begin by characterizing the portfolio choice and resulting demand for the risky asset of the different cohorts under affine prices. Given this, and using market clearing, we verify the affine prices guess and fully characterize demands and prices.

4.1 Characterization of Demands

For any $s, t \in \mathbb{N}$, let $d_{s:t} = (d_s, \dots, d_t)$ denote the history of dividends from time s up to time t . For simplicity and WLOG, we assume that the initial wealth of all generations is zero, i.e., $W_n^n = 0, \forall n \in \{0, 1, 2, \dots\}$. At time t , a n -generation agent solves the following problem:

$$\max_{x_t} E_t^n [-\exp(-\gamma x_t s_{t+1})] \quad (11)$$

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 trading in period t , demands for the risky asset are given by:*

$$x_t^n = \frac{E_t^n[s_{t+1}]}{\gamma V[s_{t+1}]} = \frac{E_t^n[s_{t+1}]}{\gamma(1 + \beta_0)^2 \sigma^2} \quad (12)$$

Proof of Proposition 4.1. The result follows by Lemma B.1 in Appendix B. □

From equation 27, it is not hard to show that demands at time t are affine in $d_{t-K:t}$. Moreover, it is also easy to see that beliefs about future dividends are linear functions of dividends observed by each generation. Therefore, one should expect that, in a linear equilibrium, prices *will not* depend on dividends which were *not* observed by the generations that are trading. This observation is the basis of Proposition 4.2 below.

4.2 Characterization of the Linear Equilibrium

We now establish that in a linear equilibrium prices (and demands) at any time t only depend on the dividends observed by the oldest generation in the market. Perhaps more importantly, the previous proposition provides a link between the factors influencing asset prices and demographic composition. In particular, in our model, only dividends observed by generations participating in the market predict prices.

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. For any $t \in \{0, 1, 2, \dots\}$*

$$p_t = \alpha + \sum_{k=0}^{q-1} \beta_k d_{t-k}. \quad (13)$$

with

$$\alpha = -\frac{1}{\left(1 - \sum_{j=0}^{q-1} \frac{w_j}{r^{j+1}}\right)^2} \frac{\gamma \sigma^2}{r-1} \quad (14)$$

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

$$\beta_q = \beta_{q+1} = \dots = 0 \quad (16)$$

where $w_k \equiv \frac{1}{q} \sum_{age=0}^{q-1} w(k, \lambda, age)$.⁴

Proof of Proposition 4.2. See Appendix B. □

For each $k = 0, 1, \dots, q-1$, one can interpret w_k as the average weight placed on the dividend observed at time $t-k$ by all trading generations at time t . The idea of the proof is as follows. By lemma C.4, demands at time t are affine in dividends $d_{t-K:t}$. However, from these dividends only (at most) $d_{t-q-1:t}$ 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 the dividends $d_{t-K:t-q}$ are zero.

This result captures the belief channel described by Friedman and Schwartz: prices are a function of past dividends solely due to the fact that generations form their beliefs using past data. By studying a general equilibrium model, however, we provide a more nuance view. Since observations of older generations affect current prices, they also affect the demand of younger generation, that did not necessarily experience those observations.

Observe that $\frac{\partial \beta_k}{\partial r} < 0$ and $\frac{\partial \alpha}{\partial r} > 0$ for any λ . Thus, the theory would predict that the equilibrium price of the risky asset is higher and less volatile if the interest rate is higher. Furthermore, higher risk aversion γ decreases the equilibrium price by lowering α .

The following proposition establishes that when agents form their beliefs by using non-decreasing weights (i.e., $\lambda \geq 0$) prices are more sensitive to more recent dividends.

⁴It is understood that $w(j, \lambda, age) = 0$ for all $j > age$.

Proposition 4.3. For $\lambda > 0$, $0 < \beta_{q-1} < \dots < \beta_1 < \beta_0 \leq \frac{1}{r-1}$.

Proof of Proposition 4.3. See Appendix B. □

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. Moreover, as $\lambda \rightarrow \infty$, it follows that w_k defined in Proposition 4.2 converges to $1_{\{k=0\}}$ for all $k = 0, 1, \dots, K$. Therefore, $\beta_k \rightarrow 0$ for all $k > 0$ and $\beta_0 \rightarrow \frac{1}{r-1}$. In other words, under extreme recency bias (i.e., $\lambda \rightarrow \infty$), only the current dividend affects prices in equilibrium, and at its maximal value, while the weights of all past dividends vanish.

Lemma 4.1. β_0 is increasing in λ .

Proof of Lemma 4.1. See Appendix B. □

4.3 Implications of EBL on Equilibrium Prices

PREDICTABILITY OF EXCESS RETURNS. We note that the equilibrium excess return at time $t + j$ is given by:

$$\frac{p_{t+j+1} + d_{t+j+1}}{p_{t+j}} - r = \frac{\alpha + (1 + \beta_0)d_{t+j+1} + \sum_{k=1}^{q-1} \beta_k d_{t+j+1-k}}{\alpha + \sum_{k=0}^{q-1} \beta_k d_{t+j-k}} - r.$$

Thus, at time t and for $j \leq q - 1$, the dividends $d_{t+j-(q-1)}, \dots, d_t$ can be used as factors for predicting the excess returns. For $j > q - 1$, our model predicts that excess returns are independent from dividends at time t . It is worth noting that the predictability of excess returns is an equilibrium phenomenon that stems solely from our learning mechanism and not from, say, a build-in dependence in dividends. In fact, our model provides a link between age profile of agents participating in the stock markets and factor for predicting stock returns. This theory provides a nuance mechanism that connects past realizations to future returns through the former's impact on the level of disagreements across market participants.

PRICE DYNAMICS. Our results imply that the variance of prices is given by $\sigma_P^2 = \left(\sum_{k=0}^{q-1} \beta_k^2\right) \sigma^2$ and also that the autocorrelation structure for prices is given by

$$\text{Cov}(p_{t+j}, p_t) = \begin{cases} \sigma^2 \left(\sum_{k=0}^{q-1-j} \beta_k \beta_{k+j}\right) & \text{for any } j \leq q-1 \\ 0 & \text{otherwise.} \end{cases}$$

A direct implication is that, as $\lambda \rightarrow \infty$, the autocorrelation of prices vanishes. That is, as the recency bias becomes stronger, the prices tend to be uncorrelated.

4.4 Characterization of the Equilibrium Demands for the Risky Asset.

The next proposition establishes that younger generations react more ‘optimistically’ (‘pessimistically’) than older generations to positive (negative) changes in current dividends.

Proposition 4.4. *For any $t \in \mathbb{Z}$ and any generation n alive at t , there is a threshold $j_0 \leq t - n - 1$ such that for dividends that date back up to j_0 periods, younger generations react stronger to changes than older generations, while for dividends that date back more than j_0 periods the opposite effect holds, i.e.,*

1. $\frac{\partial x_t^{n+1}}{\partial d_{t-j}} \geq \frac{\partial x_t^n}{\partial d_{t-j}}$ for $0 \leq j \leq j_0$ and
2. $\frac{\partial x_t^{n+1}}{\partial d_{t-j}} \leq \frac{\partial x_t^n}{\partial d_{t-j}}$ for $j_0 < j \leq t$.

Proof. See Appendix B. □

In our model, the younger generation puts more weight on current dividends when forming beliefs, so when d_t increase, they younger are “overly optimistic” relatively to the older generation. This reflects that an increase (decrease) in current dividends makes younger agents more optimistic (pessimistic) about the return of the risky asset than older agents because they put more weight on recent realizations. This term is only zero when both agents have the same belief formation (e.g. $w(0, \lambda, 1) = 1$). By Lemma 2.1, this intuition can be generalized to the more recent dividends as opposed to just the current one.

Moreover, let $\xi(n, k, t) \equiv x_t^n - x_t^{n+k}$ be the discrepancy between positions of generation n and $n + k$. By Proposition 4.1, and some algebra it follows that:

$$\xi(n, k, t) = \frac{1}{\gamma(1 + \beta_0)\sigma^2} \sum_{j=0}^{t-n} \left(w(j, \lambda, t - n) - 1_{\{j \leq t-n-k\}} \cdot w(j, \lambda, t - n - k) \right) d_{t-j} \quad (17)$$

$$= \frac{1}{\gamma(1 + \beta_0)\sigma^2} \left(E_t^n[\theta] - E_t^{n+k}[\theta] \right) \quad (18)$$

for any $k = 0, \dots, t - n$. So the discrepancy between positions of different generations is entirely explained by the discrepancy in beliefs. Note that as indicated by the indicator function, $1_{\{j \leq t-n-k\}}$, the younger generation does not form beliefs about dividends prior to their birth. For instance, if for some $a > 0$, $d_{n:t} \approx d_{n+a:t+a}$, then $\xi(n+a, k, t+a) \approx \xi(n, k, t)$.⁵

In addition, the next result shows that during "expansions", understood as periods with increasing dividends, for any two generations born during the expansions, the young generation has a relatively higher demand for risky assets than the old one.

Proposition 4.5. *Suppose $\lambda \geq 0$ and $t_0 \leq t_1$ are two points in time such that dividends are non-decreasing from t_0 up to t_1 . Then for any two generations n and $n + k$ born between t_0 and t_1 , the older generation has lower demand of the risky asset (x_t^n) than the younger generation (x_t^{n+k}) at any point $n \leq t \leq t_1$. In particular, $\xi(n, k, t) \leq 0$.*

On the other hand, if dividends are non-increasing then $\xi(n, k, t) \geq 0$.

Proof. See Appendix B. □

⁵This 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

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

4.5 Volume of Trade

We now study how learning and disagreements affect the volume of trade observed in the market. The resulting total volume of trade in the economy is defined as follows:

$$TR_t \equiv \left(\sum_{n=t-q}^t \frac{1}{q} (x_t^n - x_{t-1}^n)^2 \right)^{\frac{1}{2}} \quad (19)$$

with $x_{t-1}^t = 0$. The following Proposition characterizes trade volume for this economy.

Proposition 4.6. *Trade volume defined by (19) is given by:*

$$TR_t = \chi \left\{ \frac{1}{q} \sum_{n=0}^{q-1} \left[(\theta_t^n - \theta_{t-1}^n) - \frac{1}{q} \sum_{n=0}^{q-1} (\theta_t^n - \theta_{t-1}^n) \right]^2 \right\}^{\frac{1}{2}} \quad (20)$$

where $\chi = \frac{1}{\gamma\sigma^2(1+\beta_0)}$.

Proof. See Appendix B. □

The previous Proposition shows that the presence of learning and disagreements induces trade volume through changes in beliefs, that in our framework are driven by changes in the observed history of dividends. Note that the trade volume measure TR_t is similar to the variance of changes in beliefs of each cohort between t and $t+1$, so it proxies for the volatility of changes in beliefs. We can see that when the change in each cohorts beliefs is different from the average change in beliefs, trade volume is increased.

Therefore, to understand the drivers of trade volume, we need to understand the differential changes in beliefs across cohorts to a given shock. In our framework, an increase (decrease) in dividends impacts the belief of both generations in the market, but the effect on beliefs is stronger for the younger generations. Therefore, an increase (decrease) in dividends should induce trade if it makes young agents more optimistic (pessimistic) than old agents. This mechanism is solely due to the presence of experience-based learners, since it is essential that

each generation reacts differently to the same realization of dividends. We can see that if all agents adjust their beliefs equally, trade volume is zero.

Thought Experiment: Suppose $d_{t_0} = d_{t_0+1} = \dots = d_{t-1} = \bar{d}$ for $t - t_0 > q$ and that $d_t \neq \bar{d}$. This thought experiment is supposed to capture this economy's reaction to a shock after a long period of stability. First, note that all generations alive at time $t - 2$ and $t - 1$ have only observed a constant stream of dividends: \bar{d} . Therefore, $\theta_{t-2}^n = \theta_{t-1}^n = \bar{d}$ for all $n = \{0, \dots, q - 1\}$. As a result, we know that trade volume in $t = 1$ should be zero: $TR_{t-1} = 0$. What happens when dividend $d_t \neq \bar{d}$ is observed? Note that now, for each generation $n = \{0, \dots, q - 1\}$, beliefs are given by $\theta_t^n = w(0, \lambda, t - n)(d_t - \bar{d}) + \bar{d}$ which implies the following change in beliefs:

$$\theta_t^n - \theta_{t-1}^n = w(0, \lambda, t - n)(d_t - \bar{d}) \quad (21)$$

Trade volume in t is then given by:

$$TR_t = |d_t - \bar{d}| \chi \left[\frac{1}{q} \sum_{n=0}^{q-1} \left(w(0, \lambda, t - n)^2 - \left(\frac{1}{q} \sum_{n=0}^{q-1} w(0, \lambda, t - n) \right)^2 \right) \right]^{\frac{1}{2}} \quad (22)$$

First, note that trade volume increases proportionally to the change in dividends, independently of whether the latter is positive or negative, and to a function that reflects the dispersion of the weights agents assign to the most recent observation in their belief formation process. Second, the level of trade volume generated by a given change in dividends will depend on the level of recency bias of the economy. For example, as $\lambda \rightarrow \infty$, the dispersion in weights decreases as $w(0, \lambda, a) \rightarrow 1$ for all $a \in \{0, \dots, q - 1\}$. Thus, our results suggest that higher recency biases (reflected in higher λ), should generate lower trade volume responses for a given shock to dividends, and vice-versa.

5 Stylized Facts

Dividends in our model do not translate one-to-one to dividends in the real world. Firms might decide to retain earnings in the firm, for example in the form of cash or reinvestments, rather than distributing them to shareholders in form of dividends. In addition, the management might have incentives to smooth dividends such that they do not reflect currency profitability. We therefore turn to stock market returns rather than dividends as a measure of stock market performance.⁶

In this section, we examine whether our model is consistent with aggregate facts about equity holdings and stock turnover. While some predictions are harder to test, or generate predictions that can be attributed to several explanations (e.g., higher weights on the most recent dividend, relative to previous dividends, in determining prices), the model generates at least two predictions that are directly testable: first, the stronger response of the younger generations' risky asset demand to recent dividends; and second the relationship between differences in experience-based beliefs and trade volume.

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 stock turnover data from the Center for Research in Security Prices (CRSP).

We calculate the lifetime experiences of 'dividends' of the different generations as the weighted average of the performance over the lifetimes, using linearly declining weights and $\lambda = 3$. As described above, stock market performance is measured by annual returns of the SP500 Index. All performance measures are de-trended using the consumer price index (CPI). We construct two measures of disagreement in past experiences. First, we take the difference between the experienced performance by the old age-group (60 years and older)

⁶Further, dividends exhibit an increasing time trend which leads to overestimation of experienced dividends for younger cohorts. Earnings are another possible measure for dividends in the model but similar to dividends, earnings might not perfectly reflect profitability, plus, they exhibit a similar time trend. If dividends are included into the return calculation, the results are very similar. The appendix contains figures based on returns including dividends.

and a younger cohort (40 years and younger). For each cohort, the experienced performance is calculated as the weighted average of each single-age group within that age range.⁷ Each single-age group is weighted by the total number of people of that age and year. Second, disagreement is approximated by the standard deviation of experienced performance between the different single-age groups in each year. Again, we use the total number of individuals in each single-age group to weight the observations.

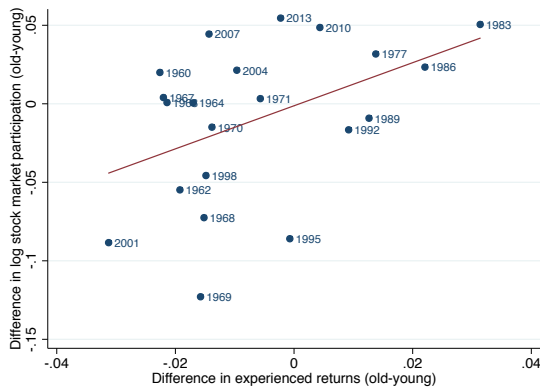
The SCF includes data on dollar stock holdings and liquid by individuals from 1960 to 2013 on a household level.⁸ With these two variables to code a measure for the extensive margin of stock holding, i.e., whether the household has invested a positive amount into stock, and a measure for the intensive margin, i.e., how much of the liquid assets are invested into stocks. For the intensive margin, we drop all households that have no money in stocks. Since the age of the head of the household is known, we can match our experienced performance measures to each household. We then aggregate the household into the aforementioned age groups by taking the unweighted average of the intensive and extensive margin over all households whose head falls into that particular age-group.⁹

Figures 7(a) to 7(d) depict the relationship between the extensive and intensive margin of stock holdings and difference in experienced returns between the above-60 age-group and the below-40 age-group for weights of past returns. In figures 7(a) and 7(c), experienced returns are calculated from equation 6, where λ is set to 1, which corresponding to linearly declining weights. For figures 7(b) and 7(d), λ is set to three. A λ corresponds closer to our model but the results are in line with the predictions for either calibration. If the older age-group experienced higher stock returns, they are more likely to hold stock compared to the younger age-group (See figure 7(a)). Likewise, the above-60 age-group invest a relatively higher share of the their liquid assets into stocks vis--vis the younger age-group if their experienced returns

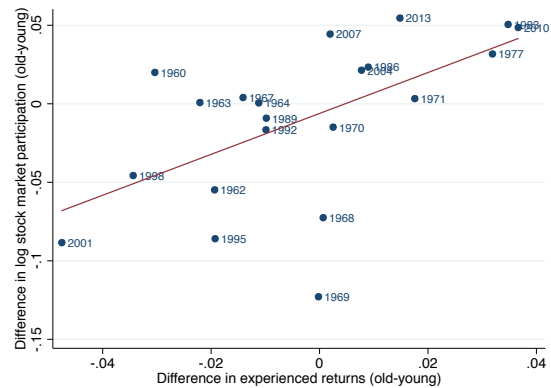
⁷We assume that every individual in the survey is born on January 1 and experiences those returns. We further assume that the SCF is also conducted on January 1. The results are unchanged if either, birthday or date of survey, or both are assumed to be December 31.

⁸Appendix D provides a more detailed description

⁹Note that the extent of the survey has changed over time. Hence, both, the intensive and extensive margin is weighted by sample weights included in the SCF.



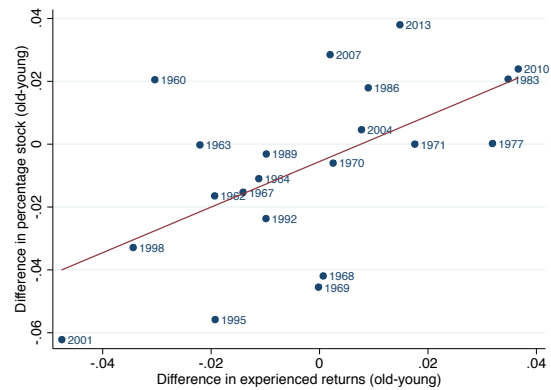
(a) Stock Market Participation. ($\lambda = 1$)



(b) Stock Market Participation. ($\lambda = 3$)



(c) Fraction of Stock of Financial Assets. ($\lambda = 1$)



(d) Fraction of Stock of Financial Assets. ($\lambda = 3$)

Figure 2: Experienced Returns and Stock Holding

are higher than those of the younger age-group. (See figure 7(c) and 7(d)).¹⁰

Our model suggests that trade volume is high when disagreement among investors is high. To validate this postulate, we examine the co-movement of trade volume and the evolution of the standard deviation of experienced performance, the aforementioned measure of disagreement. We obtain monthly data on the number of traded shares, number of outstanding shares and stock price for every ordinary common share in CRSP from 1960 to 2007. For each month the traded value is calculate as the number of traded shares times the average

¹⁰The results are less clear-cut for the other performance measure, such as experienced dividends and earnings; see Appendix XXX.

share price.¹¹ To account for changes in capitalization, we scale the traded volume by the market capitalization of the firm. The market capitalization is calculated as the number of outstanding shares times the average stock price. We refer to the scaled traded volume as turnover ratio. To align the turnover ratio with the frequency of our disagreement variable, we calculate the annual turnover ratio as the average of the monthly turnover ratios. Since the trade volume is likely driven by a number of factors not related to disagreement, for instance technological progress, we de-trend the turnover ratio as follows. First, we take the log of the turnover ratio. Second, the logged turnover series are regressed on a linear time trend. The residuals are averaged for each year to obtain a measure of deviation of the trading activity from the trend.

The co-movement of trading volume and disagreement are generally in line with our prediction (See Figure 3). If disagreement among investors, approximated by the standard deviation of experienced performance, is higher, the actual turnover ratio is higher than the trend turnover ratio.¹²

6 Extension: The model with non-myopic agents.

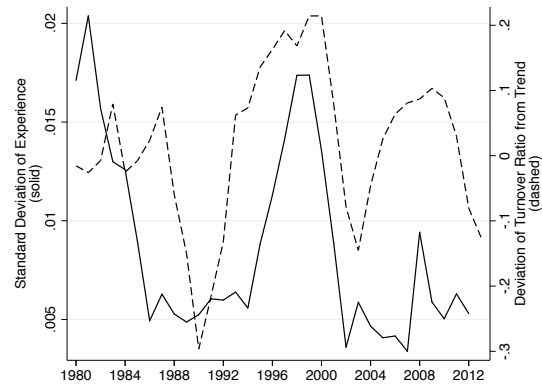
In this section, we consider non-myopic agents who choose their portfolios by looking at their entire lifetime. We assume that agents consume only in their final period, i.e., they consume their final wealth. We first characterize the demands for risky assets. Even though demands for risky assets are linear (in accordance with Proposition 4.1), the functional form contains an additional term that accounts for the dynamic nature of the non-myopic problem. Due to the aforementioned linearity in risky demands, we are able to show that the result in Proposition 4.2 continue to hold: Prices are still affine functions of past dividends only observed by generations that are trading.

¹¹In particular, we obtain the items "prc" for the monthly stock price, "vol" for the number of traded shares in a month and "shout" for the number of shares outstanding from CRSP.

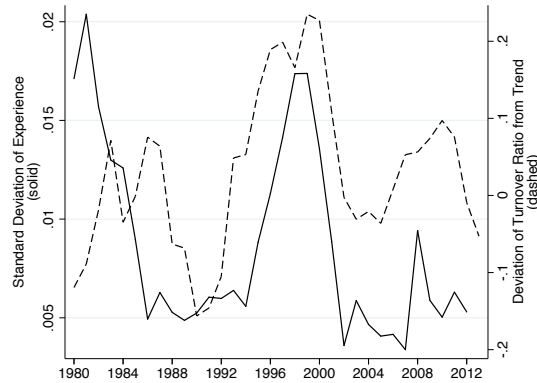
¹²Again, the results are less clear for experienced dividends and earnings.



(a) Moving Average with 1 Lag



(b) Moving Average with 3 Lags



(c) Moving Average with 5 Lags

Figure 3: Turnover Ratio and Disagreement in Experienced Returns

Figure 4: The dashed line depicts the turnover ratio from its trend. The solid line shows the standard deviation of experienced stock returns for a given year. The turnover ratio is smoothed by taking the moving average with 1, 3 and 5 lags.

6.1 Characterization of risky demands for non-myopic agents.

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

$$\max_{\mathbf{x} \in \mathbb{R}^q} E_t^t [-\exp(-\gamma W_{t+q}^n(\mathbf{x}))] \quad (23)$$

$$\text{s.t. } W_{t+q}^t(\mathbf{x}) = \sum_{\tau=t}^{t+q-1} r^{t+q-1-\tau} x_\tau s_{\tau+1} \quad (24)$$

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

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 $t+q-1$ backwards — as

$$V_{t+q-1}^t(d_{t+q-1-K:t+q-1}) = \max_{x \in \mathbb{R}} E_{t+q-1}^t [-\exp(-\gamma s_{t+q} x)] \quad \text{and} \quad (25)$$

$$V_\tau^t(d_{\tau-K:\tau}) = \max_{x \in \mathbb{R}} E_\tau^t [V_{\tau+1}^t(d_{\tau+1-K:\tau+1}) \exp(-\gamma s_{\tau+1} x)], \quad \forall \tau \in \{t, \dots, t+q-2\} \quad (26)$$

Remark 6.1. Notice that V_τ^t does not include the wealth at time τ , that is, from equation 11, the optimization problem can be cast as $\max_{x \in \mathbb{R}} \exp\{-\gamma r W_{t+q-1}^t\} E_{t+q-1}^t [-\exp(-\gamma s_{t+q} x)]$. However, our definition of V_{t+q-1}^t omits the term $\exp\{-\gamma r W_{t+q-1}^t\}$ since it does not affect the maximization.

This shows that, although the t -generation's problem at $t+q-1$ is a static portfolio problem, for any other $\tau \in \{t, \dots, t+q-2\}$, it is not because $V_{\tau+1}^t$ is correlated with $s_{\tau+1}$ through dividends. That is, dividend realization $d_{\tau+1}$ impacts (i) the net payoff obtained from investing x_τ in the risky asset at time τ , and (ii) the continuation value $V_{\tau+1}^t(d_{\tau+1-K:\tau+1})$ by affecting the beliefs of the t -generation at $\tau+1$, and the resulting portfolio decision.

First, we characterize the portfolio choice and resulting demand for the risky asset of the different cohorts under affine prices. We begin by highlighting that the dynamic portfolio problem of agents in this economy cannot be expressed as a succession of static problems, as

is standard in the literature (see [Vives \(2008\)](#)). This is because of learning and the fact that agents are sophisticated enough to understand how their beliefs evolve over their lifetime. These features introduce a correlation between future returns and continuation values that distorts the portfolio decisions. We will be able to show, however, that the agents dynamic portfolio problem can be expressed as an *adjusted static* problem where dividends follow a normal distribution with *adjusted* mean and variance. Intuitively, agents recognize that very high and very low realizations of future dividends will lead to more disagreement, which they will exploit in their future trades. As a result, extreme realizations are now associated with higher continuation values, leading to a downward adjustment of the variance.

In the CARA-Gaussian framework with *no learning*, continuation values are constant and thus uncorrelated with returns $s_{\tau+1}$. Therefore, the dynamic problem becomes a sequence of static ones with a risk-aversion coefficient adjusted by the horizon of the agent. In our setup, with non-myopic agents, because of the presence of learning, this will not be the case. However, [Proposition 6.1](#) below shows that at each time t , can be expressed as an *adjusted static* portfolio problem where dividends follow a normal distribution with *adjusted* mean and variance.

Let $E_{N(\mu, \sigma^2)}[\cdot]$ and $V_{N(\mu, \sigma^2)}[\cdot]$ be the expectation and variance with respect to a Gaussian pdf with mean μ and σ^2 instead of the empirical probability measure \mathbb{P}_t^n .

Proposition 6.1 (pro: demands). *Suppose $p_t = \alpha + \sum_{k=0}^K \beta_k d_{t-k}$ with $\beta_0 \neq -1$. Then, for any generation t in period $t+j$ for $j \in \{0, \dots, q-1\}$ (the age of the generation), demands for the risky asset are given by:*

$$x_{t+j}^t = \frac{E_{N(m_j, \sigma_j^2)}[s_{t+j+1}]}{\gamma r^{q-1-j} V_{N(m_j, \sigma_j^2)}[s_{t+j+1}]} \quad (27)$$

where:

$$m_j \equiv \frac{\theta_{t+j}^t - \sigma^2 \left(b_j + \sum_{k=1}^K b_j(k) d_{t+j-k} \right)}{2c_j \sigma^2 + 1} \quad (28)$$

$$\sigma_j^2 \equiv \frac{\sigma^2}{2c_j \sigma^2 + 1} \quad (29)$$

for $\{\{b_j(k)\}_{k=1}^{q-1}, b_j, c_j\}$ constants that change with the agent's age (j) (for exact expressions see the proof).

Proof of Proposition 6.1. See Appendix C. □

The intuition of the proof is as follows. By solving the problem backwards we note that at time $t+q-1$ the problem is in fact a static one (see equation 25). In particular we show that V_{t+q-1}^t is of the form exponential-quadratic in d_{t+q-1} (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 $t+q-2$ can be viewed as a static problem with a modified Gaussian distribution, and consequently (a) demands are of the form of 27 and V_{t+q-2}^t is also of the exponential-quadratic form. The process thus continues until time t .

After straightforward algebra, we can cast equation 27, as ¹³

$$\begin{aligned} x_{t+j}^t &= \frac{1}{r^{q-1-j}} \frac{E_{N(\theta_{t+j}^t, \sigma^2)}[s_{t+j+1}]}{\gamma V_{N(\theta_{t+j}^t, \sigma^2)}[s_{t+j+1}]} - \frac{(b_j + \sum_{k=1}^K b_j(k) d_{t+j-k})}{\gamma r^{q-1-j} (1 + \beta_0)} \\ &\equiv \frac{1}{r^{q-1-j}} \tilde{x}_{t+j}^t + \Delta_{t+j}^t \end{aligned} \quad (30)$$

The term \tilde{x}_{t+j}^t coincides with the demand of a *static* portfolio problem for an agent with beliefs θ_{t+j}^t ; see proposition 4.1. We coin this term the *myopic component* of the demand for risky assets. The scaling by $1/r^{q-1-j}$ arise because agents discount the future by r . The second term $\Delta_{t+j}^t \equiv -\frac{(b_j + \sum_{k=1}^K b_j(k) d_{t-k})}{\gamma r^{q-1-j} (1 + \beta_0)}$, is an adjustment which accounts for the dynamic

¹³Note that $E_{N(b+a,s)}[s_{t+1}] = E_{N(a,s)}[s_{t+1}] + (1 + \beta_0)b$.

nature of the problem, and thus, we call it the *dynamic component*. It arises because agents understand that they are learning about the risky asset, and thus understand that the value function is correlated with the one-period-ahead returns.

6.2 Characterization of equilibrium prices for non-myopic agents

The following proposition shows that in a linear equilibrium prices at any time t only depend on the dividends observed by the generations trading at time t . This result shows that the insights in Proposition 4.2 continue to hold in this setup with non-myopic agents.

Proposition 6.2. *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 $t \geq 0, q \geq 1$,*¹⁴

$$p_t = \alpha + \sum_{k=0}^{q-1} \beta_k d_{t-k}. \quad (31)$$

Proof of Proposition 6.2. See Appendix C. □

The idea of the proof is as the one discussed for the myopic case.

6.3 The $q = 2$ Case.

We now specialize our results to the case with $q = 2$. By doing so, we are able to sharpen our previous results regarding the behavior of prices and risky demands in equilibrium.

The next lemma shows that $\{\alpha, \beta_0, \beta_1\}$ solve a complicated system of non-linear equations

Lemma 6.1. *For $r > 1$ in any linear equilibrium prices are given by:*

$$p_t = \alpha + \beta_0 d_t + \beta_1 d_{t-1} \quad \forall t \geq 1 \quad (32)$$

¹⁴Heuristically, 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 depending on $d_{t+q-1} + p_{t+q-1} - r p_{t+q-2}$.

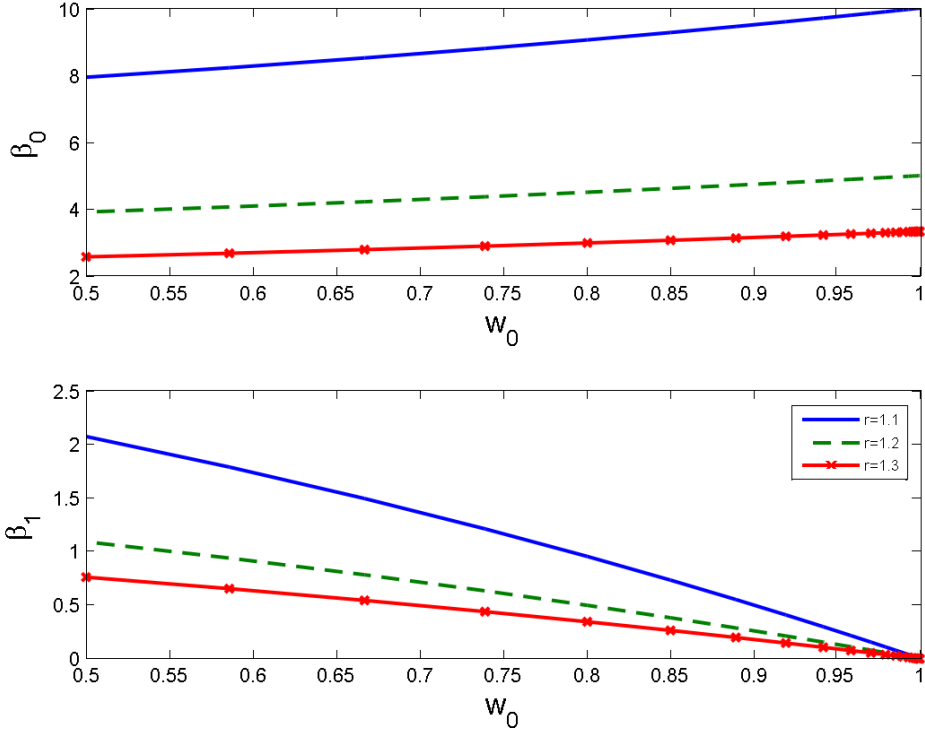


Figure 5: Comparative Statics: Sensitivity of Prices to Dividends for the $q=2$ Case.

where the coefficients $\{\alpha, \beta_0, \beta_1\}$ are uniquely determined by a set of three non-linear equations specified in the appendix.

Proof of Lemma 6.1. See Appendix C. □

Although the equations in the lemma form a complicated system of non-linear equations, we are able to establish that prices react positively to dividends d_t and d_{t-1} . Formally,

Proposition 6.3. For $\lambda > 0$, $\alpha \leq 0$ and $0 < \beta_1 < r\beta_0$.

Proof of Proposition 6.3. See Appendix C. □

This proposition is analogous to Proposition 4.3 and establishes that when agents form their beliefs by using non-decreasing weights (i.e., $\lambda \geq 0$) $\beta_0 r$ is larger than β_1 . This result

reflects the fact that the dividends at time t are observed by both generations whereas d_{t-1} is only observed by the old generation; in fact it is not hard to see from the equations that in the case $w(1, \lambda, 0) = 0$ –agents do not put any weight on the previous dividend,– then $\beta_1 = 0$.

Figure 5 depicts the behavior of $\{\beta_0, \beta_1\}$ for different values of (λ, r) . Note that the values of $\{\beta_0, \beta_1\}$ are independent of the process for dividends, σ^2 , and of the coefficient of risk aversion, γ . Thus, the results shown in the figure do not depend on parameter values other than the ones used for comparative statics: (λ, r) .

The next proposition establishes that, as before, the demand of the young generation (decreases) increases, while the one of the old generation (increases) decreases, when current dividends (decrease) increase; and the opposite holds for the dividends last period.

Proposition 6.4. *For $\lambda > 0$: (1) $\frac{\partial x_t^t}{\partial d_t} > 0 > \frac{\partial x_t^{t-1}}{\partial d_t}$, and (2) $\frac{\partial x_t^t}{\partial d_{t-1}} < 0 < \frac{\partial x_t^{t-1}}{\partial d_{t-1}}$.*

Proof of Proposition 6.4. See Appendix C. □

In our model, the young generation puts more weight on current dividends when forming beliefs, so when d_t increase, they young are "overly optimistic" relatively to the old generation. This effect contributes to the result (1) (and similar reasoning contributes to results (2)); however, this is not the only effect to consider. There additional effects due to the fact that the young are confronted with a different horizon investment.

In order to shed some light on the different effects, recall that in equation 30 we decomposed the demand for risky asset into two components: The myopic one and the dynamic one. For the particular case of $q = 2$ these decomposition yields $x_t^{t-1} = \tilde{x}_t^{t-1}$ and $x_t^t = \tilde{x}_t^t + \Delta_t^t$, where

$$\tilde{x}_t^{t-1} = \frac{\alpha(1-r)}{\gamma(1+\beta_0)^2\sigma^2} + \frac{(1+\beta_0)w(0, \lambda, 1) + \beta_1 - r\beta_0}{\gamma(1+\beta_0)^2\sigma^2}d_t + \frac{(1+\beta_0)(1-w(0, \lambda, 1)) - r\beta_1}{\gamma(1+\beta_0)^2\sigma^2}d_{t-1},$$

$$\tilde{x}_t^t = \frac{\alpha(1-r)}{\gamma r(1+\beta_0)^2\sigma^2} + \frac{1+\beta_0 + \beta_1 - r\beta_0}{\gamma r(1+\beta_0)^2\sigma^2}d_t + \frac{-r\beta_1}{\gamma r(1+\beta_0)^2\sigma^2}d_{t-1}$$

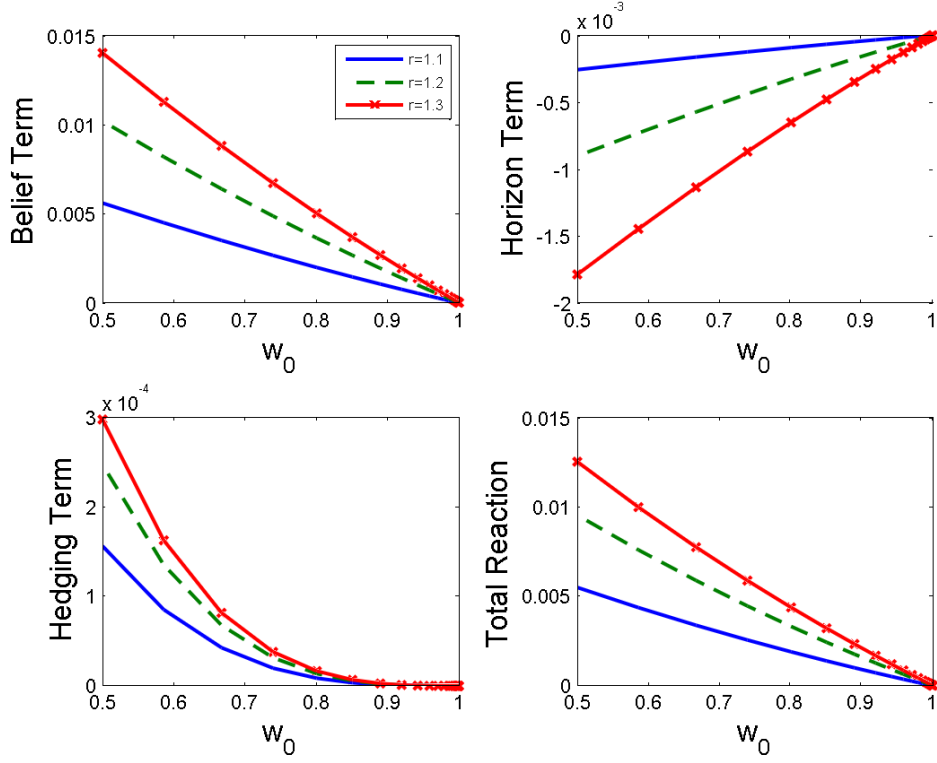


Figure 6: Comparative Statics: Sensitivity of Demands to Dividends for the $q=2$ Case.

Decomposition of $\frac{\partial(x_t^t - x_t^{t-1})}{\partial d_t}$ into the Belief, Horizon, and Hedging Terms.

and

$$\Delta_t^t \equiv \frac{\alpha(1-r) + (\beta_1 - r\beta_0)d_t - r\beta_1 d_{t-1}}{\gamma r(1+\beta_0)^2} \left(\frac{1}{s^2} - \frac{1}{\sigma^2} \right) + \frac{1}{\gamma r(1+\beta_0)} \left(\frac{m}{s^2} - \frac{d_t}{\sigma^2} \right)$$

where $s^2 = \sigma^2 \frac{(1+\beta_0)^2}{(1+\beta_0)^2 + ((1+\beta_0)w(0,\lambda,0) + \beta_1 - r\beta_0)^2}$.

We focus first on understanding the changes in the myopic term. Let

$$\frac{\partial(\tilde{x}_t^t - \tilde{x}_t^{t-1})}{\partial d_t} = \underbrace{\frac{(1+\beta_0)(1-w(0,\lambda,1))}{\gamma(1+\beta_0)^2\sigma^2}}_{\text{Beliefs Term}} + \underbrace{\frac{1+\beta_0+\beta_1-r\beta_0}{\gamma(1+\beta_0)^2\sigma^2} \left(-\frac{r-1}{r} \right)}_{\text{Discount Term}}$$

We refer to the first term as the *Beliefs Term*. This term is positive, and it reflects that an

increase (decrease) in dividends makes young agents more optimistic (pessimistic) about the return of the risky asset than adult agents because they put more weight on recent realizations. This term is zero when both agents have the same belief formation (e.g. $w(0, \lambda, 1) = 1$). The second term is the *Discount Term*, which is negative (see Lemma C.5 in Appendix). Even when agents share beliefs, young agents react less aggressively to a change in dividends (in their beliefs) because they discount the future more than old agents since $r > 1$.

Regarding the myopic term, observe that

$$\frac{\partial \Delta_t^t}{\partial d_t} = \frac{\beta_1 - r\beta_0}{\gamma(1 + \beta_0)^2 \sigma^2} \frac{1}{r} \left(\frac{\sigma^2}{s^2} - 1 \right) - \frac{(1 + \beta_0)}{\gamma(1 + \beta_0)^2 \sigma^2} \left(\frac{l(1, 1)l(0, 1)}{(1 + \beta_0)^2 r} \right)$$

Because the first term in the RHS is negative but the second term is positive (see the proof of Proposition 6.4), we can not pin down the sign of $\frac{\partial \Delta_t^t}{\partial d_t}$.

Therefore, even though the behavior $\frac{\partial(x_t^t - x_t^{t-1})}{\partial d_t}$ is affected by all these terms, we are able to show that the belief term dominates and thus the overall sign of $\frac{\partial(x_t^t - x_t^{t-1})}{\partial d_t}$ is positive.

In figure 6 we show the behavior of each of the terms for different values of (r, λ) . Importantly, as λ increases, the "old" generation puts less weight to past dividends, and thus the discrepancy between the beliefs of the old and the young vanishes.

7 Conclusion

To be completed.

A Proofs of Section 2

Proof of Lemma 2.1. First, note that $\sum_{j=0}^a w(j, \lambda, a) = \sum_{j=0}^{a'} w(j, \lambda, a') = 1$, and thus $j \mapsto w(j, \lambda, a) - w(j, \lambda, a')$ change signs at least once for $j \in \{0, \dots, a'\}$ with $a \neq a'$. The sign of the difference can be characterized as follows:

$$w(j, \lambda, a) - w(j, \lambda, a') \geq 0 \iff Q(j) := \frac{w(j, \lambda, a)}{w(j, \lambda, a')} \geq 1 \quad (33)$$

Thus, we are done if we can show that Q is monotonic regardless of λ . Note that the normalizing constants used in the weights $w(j, \lambda, a)$ are independent of j . Hence, we can absorb them in a constant $c \in \mathbb{R}_+$ and write

$$Q(j) = c \cdot \left[\frac{(a+1-j)^\lambda}{(a'+1-j)^\lambda} \right] = c \cdot \left[\frac{(a+1-j)}{(a'+1-j)} \right]^\lambda = c \cdot \alpha(j)^\lambda \quad (34)$$

where $\alpha(j) = \frac{a+1-j}{a'+1-j}$. By the chain rule, Q is monotonic if α is monotonic. This is the case since $\alpha'(j) = \frac{a-a'}{(a'+1-j)^2}$ never changes sign. Thus, there is exactly one crossing point. \square

B Proofs of Section 4

Proposition 4.1 directly follows from the following Lemma.

Lemma B.1 (l: static). *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 $\max_x E[-\exp\{-axz\}] = -\exp\{-0.5(\sigma ax^*)^2\} = -\exp\left(-0.5\frac{\mu^2}{\sigma^2}\right)$

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

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

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

Proof of Proposition 4.2. Give our guess, prices are given by $p_t = \alpha + \beta_0d_t + \dots + \beta_qd_{t-q}$. The solution to the agents problem 3 is given by $x_t^n = \frac{E_t^n[s_{t+1}]}{\gamma V_t^n[s_{t+1}]}$ from Lemma B.1. Plugging in our guess for prices:

$$x_t^n = \frac{(1 + \beta_0)\theta_t^n + \alpha + \beta_1d_t + \dots + \beta_qd_{t-q+1} - rp_t}{\gamma(1 + \beta_0)^2\sigma^2} \quad (35)$$

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

$$\frac{(1 + \beta_0)\frac{1}{q} \sum_{n=t}^{t-q+1} \theta_t^n}{\gamma(1 + \beta_0)^2\sigma^2} + \frac{\alpha + \beta_1d_t + \dots + \beta_qd_{t-q+1} - rp_t}{\gamma(1 + \beta_0)^2\sigma^2} = 1.$$

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

$$(1 + \beta_0)\frac{1}{q} \sum_{n=t}^{t-q+1} \left[\sum_{k=0}^{t-n} w(k, \lambda, t-n) d_{t-k} \right] + \left[\alpha - \gamma(1 + \beta_0)^2\sigma^2 \right] + \beta_1d_t + \dots + \beta_qd_{t-q+1} = rp_t.$$

Using the method of undetermined coefficients we find the expressions for α and the β s:

$$\begin{aligned}
-\frac{\gamma(1+\beta_0)^2\sigma^2}{r-1} &= \alpha \\
(1+\beta_0)\frac{1}{q}\sum_{n=t}^{t-q+1} w(0,\lambda,t-n) + \beta_1 &= r\beta_0 \\
(1+\beta_0)\frac{1}{q}\sum_{n=t-k}^{t-q+1} w(k,\lambda,t-n) + \beta_{k+1} &= r\beta_k \quad k = \{2, \dots, q-1\} \\
(1+\beta_0)\frac{1}{q}w(q-1,\lambda,q-1) + \beta_q &= r\beta_{q-1} \\
0 &= r\beta_q
\end{aligned}$$

Let $w_k = \frac{1}{q}\sum_{n=t}^{t-q+1} w(k,\lambda,t-n)$ be the weight assigned to dividend d_{t-k} given by the average of weights assigned by each generation in the market (where note that generations that did not observe it (that is, $k > t-n$) assign a weight of zero. Then:

$$(1+\beta_0)w_{q-1} = r\beta_{q-1} \quad (36)$$

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

By solving this system of equations we obtain the expressions in the proposition. In particular, $(1+\beta_0)(w_{q-2} + r^{-1}w_{q-1}) = r\beta_{q-2}$, $(1+\beta_0)(w_{q-3} + r^{-1}w_{q-2} + r^{-2}w_{q-1}) = r\beta_{q-3}$ and so on to obtain

$$(1+\beta_0)\sum_{j=0}^{k-1} r^{-j}w_{q-(k-j)} = r\beta_{q-k}, \text{ for } k = \{2, \dots, q\}$$

In particular, this implies that $\beta_0 = \frac{\sum_{j=0}^{q-1} r^{-j}w_j}{r - \sum_{j=0}^{q-1} r^{-j}w_j} = \frac{\sum_{j=0}^{q-1} r^{-(j+1)}w_j}{1 - \sum_{j=0}^{q-1} r^{-(j+1)}w_j}$ and $\beta_k = \frac{\sum_{j=0}^{q-1-k} r^{-(j+1)}w_{k+j}}{1 - \sum_{j=0}^{q-1} r^{-(j+1)}w_j}$. \square

Proof of Proposition 4.3. For this proof, we will use equations 36 and 37. In addition, note that by construction, $w_k < w_{k-1}$ for $\lambda > 0$ since for all generations, $w(k,\lambda,age)$ is decreasing

in k and more agents observe the realization of $d_{t-(k-1)}$ than d_{t-k} . Given this, it follows that since $\beta_0 > 0$ then $\beta_{q-1} > 0$ and

$$\beta_{q-1} = \frac{1}{r}[(1 + \beta_0) w_{q-1}] < \frac{1}{r}[(1 + \beta_0) w_{q-2} + \beta_{q-1}] = \beta_{q-2} \quad (38)$$

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

$$\beta_{k-1} = \frac{1}{r}[(1 + \beta_0) w_{k-1} + \beta_k] < \frac{1}{r}[(1 + \beta_0) w_{k-2} + \beta_{k-1}] = \beta_{k-2} \quad (39)$$

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

Proof of Lemma 4.1. A. β_0 is increasing in λ . Let $G_q(\lambda) = \sum_{j=0}^{q-1} c^{j+1} w_j$ where $c = \frac{1}{r}$. Thus $\beta_0 = \frac{G_q(\lambda)}{1 - G_q(\lambda)}$, and it suffices to show that $G'_q(\lambda) > 0$, $\forall q > 0$, $\forall \lambda > 0$. After some algebra, the terms in $G_q(\cdot)$ can be re-organized as follows:

$$G_q(\lambda) = \sum_{a=0}^{q-1} \frac{1}{q} \sum_{j=0}^a c^{j+1} w(j, \lambda, a) \quad (40)$$

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

Proof of Proposition 4.4. From Proposition 4.1, for any k and t ,

$$\frac{\partial x_t^k}{\partial d_t} = \frac{1}{V[s_{t+1}]} \left((1 + \beta_0) \frac{\partial \theta_t^k}{\partial d_{t-j}} - r \beta_0 \right).$$

It follows that $\frac{\partial \theta_t^k}{\partial d_{t-j}} = w(j, \lambda, k)$. Hence, it suffices to show that $w(j, \lambda, t-n) < w(j, \lambda, t -$

$(n + 1)$). First lets consider the case for $j = 0$. For any age ,

$$w(0, \lambda, age) = \frac{(age + 1)^\lambda}{\sum_{k=0}^{age} (age + 1 - k)^\lambda} = \left(\sum_{k=0}^{age} \left(\frac{age + 1 - k}{age + 1} \right)^\lambda \right)^{-1} = \left(1 + \sum_{k=0}^{age-1} \left(\frac{age - k}{age + 1} \right)^\lambda \right)^{-1},$$

and

$$w(0, \lambda, age + 1) = \frac{(age + 2)^\lambda}{\sum_{k=0}^{age} (age + 2 - k)^\lambda} = \left(\sum_{k=0}^{age+1} \left(\frac{age + 2 - k}{age + 2} \right)^\lambda \right)^{-1} = \left(1 + \sum_{k=0}^{age} \left(\frac{age + 1 - k}{age + 2} \right)^\lambda \right)^{-1},$$

So to establish $w(0, \lambda, age + 1) < w(0, \lambda, age)$ with $age = t - n - 1$, it suffices to show that

$$\sum_{k=0}^{age-1} \left(\frac{age - k}{age + 1} \right)^\lambda < \sum_{k=0}^{age} \left(\frac{age + 1 - k}{age + 2} \right)^\lambda$$

We note that in the second expression, there are $age + 1$ terms whereas in the first one there are age . We show that $\frac{age-l}{age+1} < \frac{age+1-l}{age+2}$ for any $l = 0, \dots, age$. The inequality holds iff $(age - l)(age + 2) < (age + 1 - l)(age + 1)$ iff $age^2 + 2age - l(age + 2) < age^2 + (1 - l)age + age + (1 - l)$ iff $-2l < 1 - l$ for holds for any $l \geq 0$.

Therefore, $w(0, \lambda, t - n) < w(0, \lambda, t - (n + 1))$. From Lemma 2.1, there exists a j_0 such that $w(j, \lambda, t - n) < w(j, \lambda, t - (n + 1))$ for all $j \in \{0, \dots, j_0\}$ and $w(j, \lambda, t - n) \geq w(j, \lambda, t - (n + 1))$ for the rest of the j 's.

□

The proof of Proposition 4.5 relies on the following First order stochastic dominance result.

Lemma B.2. *Let $F(m, a) \equiv \sum_{j=0}^m w(j, \lambda, a)$. If $j \mapsto w(j, \lambda, a) - w(j, \lambda, a')$ for $a' < a$ is increasing in j , then $F(m, a) \leq F(m, a')$ for all $m \in \{0, \dots, a\}$.*

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

□

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

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

It follows that the quantities above, as functions of m , are non-decreasing and $F(t - n, t - n - k) = F(t - n, t - n) = 1$ and $F(-1, t - n - k) = F(-1, t - n) = 0$. Moreover, $F(m + 1, t - n - k) - F(m, t - n - k) = w(m + 1, \lambda, t - n - k)$ and $F(m + 1, t - n) - F(m, t - n) = w(m + 1, \lambda, t - n)$.

By these observations and our previous derivations for $\xi(n, k, t)$, it follows that,

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

By assumption $d_{t-j} - d_{t-j-1} \geq 0$ for all $j = 0, \dots, t - n + 1$ in the case that dividends are non-decreasing. Thus, it suffices to show that $F(j, t - n) \leq F(j, t - n - k)$ for all $j = 0, \dots, t - n + 1$. To show this, we note that by Lemma 2.1 the hypothesis in Lemma B.2 holds. Thus, the result follows from applying the latter lemma with $a = t - n > t - n - k = a'$ and $j \in \{0, \dots, t - n\}$.

Note that if the weights are non-increasing, then $d_{t-j} - d_{t-j-1} \leq 0$. Therefore, the sign of

$\xi(n, k, t)$ changes accordingly. □

Proof of Proposition 4.6. By Proposition 4.1 and 4.2, it follows that

$$\begin{aligned} x_t^n &= \frac{1}{\gamma\sigma^2(1+\beta_0)^2} \left(\alpha_0 + (1+\beta_0)\theta_t^n + \sum_{k=1}^{q-1} \beta_k d_{t+1-k} - r \left(\alpha_0 + \sum_{k=0}^{q-1} \beta_k d_{t-k} \right) \right) \\ &= \frac{1}{\gamma\sigma^2(1+\beta_0)^2} \left(\alpha_0(1-r) + (1+\beta_0)\theta_t^n - r\beta_0 d_t + \sum_{k=1}^{q-1} \beta_k (d_{t+1-k} - r d_{t-k}) \right) \end{aligned}$$

and thus

$$x_t^n - x_{t-1}^n = \frac{1}{\gamma\sigma^2(1+\beta_0)^2} \left((1+\beta_0)(\theta_t^n - \theta_{t-1}^n) - r\beta_{q-1}(d_{t-q+1} - d_{t-q}) + \sum_{k=1}^{q-1} (\beta_k - r\beta_{k-1})(d_{t+1-k} - r d_{t-k}) \right)$$

and

$$\begin{aligned} \frac{1}{q} \left(\sum_{n=0}^{q-1} x_t^n - x_{t-1}^n \right) &= 0 = \frac{1}{q} \sum_{n=0}^{q-1} \frac{(\theta_t^n - \theta_{t-1}^n)}{\gamma\sigma^2(1+\beta_0)} \\ &\quad + \frac{1}{q} \sum_{n=0}^{q-1} \left(\frac{-r\beta_{q-1}(d_{t-q+1} - d_{t-q}) + \sum_{k=1}^{q-1} (\beta_k - r\beta_{k-1})(d_{t+1-k} - r d_{t-k})}{\gamma\sigma^2(1+\beta_0)^2} \right) \end{aligned}$$

Then, we can express the change in individual demands as follows:

$$x_t^n - x_{t-1}^n = \chi \left[(\theta_t^n - \theta_{t-1}^n) - \frac{1}{q} \sum_{a=0}^{q-1} (\theta_t^a - \theta_{t-1}^a) \right]$$

where $\chi \equiv \frac{1}{\gamma\sigma^2(1+\beta_0)}$. From our TV formula, we focus on $\frac{TV^2}{\chi}$:

$$\frac{TV^2}{\chi} = \frac{1}{q} \sum_{n=0}^{q-1} \left[(\theta_t^n - \theta_{t-1}^n) - \frac{1}{q} \sum_{a=0}^{q-1} (\theta_t^a - \theta_{t-1}^a) \right]^2$$

$$\begin{aligned}
&= \frac{1}{q} \sum_{n=0}^{q-1} \left[(\theta_t^n - \theta_{t-1}^n)^2 + \left(\frac{1}{q} \sum_{a=0}^{q-1} (\theta_t^a - \theta_{t-1}^a) \right)^2 - 2 (\theta_t^n - \theta_{t-1}^n) \frac{1}{q} \sum_{a=0}^{q-1} (\theta_t^a - \theta_{t-1}^a) \right] \\
&= \frac{1}{q} \sum_{n=0}^{q-1} (\theta_t^n - \theta_{t-1}^n)^2 + \left(\frac{1}{q} \sum_{n=0}^{q-1} (\theta_t^n - \theta_{t-1}^n) \right)^2 - 2 \frac{1}{q^2} \sum_{n=0}^{q-1} (\theta_t^n - \theta_{t-1}^n) \sum_{a=0}^{q-1} (\theta_t^a - \theta_{t-1}^a) \\
&= \left(\frac{1}{q} \sum_{n=0}^{q-1} (\theta_t^n - \theta_{t-1}^n) \right)^2 + \frac{1}{q} \sum_{n=0}^{q-1} (\theta_t^n - \theta_{t-1}^n)^2 - 2 \left(\frac{1}{q} \sum_{n=0}^{q-1} (\theta_t^n - \theta_{t-1}^n) \right)^2 \\
&= \frac{1}{q} \sum_{n=0}^{q-1} (\theta_t^n - \theta_{t-1}^n)^2 - \left(\frac{1}{q} \sum_{n=0}^{q-1} (\theta_t^n - \theta_{t-1}^n) \right)^2
\end{aligned}$$

□

C Proofs for Section 6

To establish the results in this section we need the following lemmas (the proofs are relegated to the end of this section).

Lemma C.1 (l: AdjustedGaussian). *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 (l: TwoLastPeriods). *Demands for the risky asset in the last two period of an agent's life are given by: $x_t^{t-q} = 0$ and $x_t^{t-q+1} = \frac{E_t^{t-q+1}[s_{t+1}]}{\gamma \sigma_z^2}$, $\forall t \geq 0, q \geq 1$.*

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

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

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

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

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

$$x_t^{t-age} = \delta(age) + \sum_{k=0}^K \delta_k(age) d_{t-k}, \text{ for } age \in \{0, \dots, q\} \quad (41)$$

with

$$\delta(q) = \delta_k(q) = 0, \quad \forall k \in \{0, \dots, K\} \quad (42)$$

$$\delta(q-1) = \frac{\alpha(1-r)}{\gamma((1+\beta_0)\sigma)^2}, \quad \delta_k(q-1) = \frac{(1+\beta_0)w(k, \lambda, q-1) + \beta(k)}{\gamma((1+\beta_0)\sigma)^2}, \quad \forall k \in \{0, \dots, q-1\} \quad (43)$$

$$\delta_k(q-1) = \frac{\beta(k)}{\gamma((1+\beta_0)\sigma)^2}, \quad \forall k \in \{q, \dots, K\}, \quad (44)$$

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

$$\delta(age) = \frac{\alpha(1-r) - s_{age}^2(1+\beta_0)\delta_0(age+1)\delta(age+1)(r^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2}{r^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2}, \quad (45)$$

$$\delta_k(age) = \frac{(1+\beta_0)s_{age}^2(\sigma^{-2}w(k, \lambda, age) - [(r^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2\delta_{k+1}(age+1)\delta_0(age+1)]) + \beta(k)}{r^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2} \quad (46)$$

$$k \in \{0, \dots, q-1\}, \quad (47)$$

$$\delta_k(age) = \frac{-(1+\beta_0)s_{age}^2[(r^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2\delta_{k+1}(age+1)\delta_0(age+1)] + \beta(k)}{r^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2}, \quad k \in \{q, \dots, K-1\} \quad (48)$$

$$\delta_K(age) = \frac{\beta(K)}{r^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2}, \quad (49)$$

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

The expressions for $b_j, b_j(k)$ and c_j for $j \in \{0, \dots, q-1\}$ are:

$$b_j \equiv (r^{q-1-j}\gamma)^2((1+\beta_0)\sigma_j)^2\delta(j)\delta_0(j)$$

$$b_j(k) \equiv \delta_k(j)\delta_0(j)(r^{q-1-j}\gamma)^2((1+\beta_0)\sigma_j)^2$$

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

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

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

Proof of Proposition 6.1. By lemma B.1,

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

with $m_{q-1} = \theta_{t+q-1}^t$ and $\sigma_{q-1} = \sigma$, and

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

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 $d_{t+q-1-K:t+q-2}$ and C is constant with respect to $d_{t+q-1-K:t+q-1}$:

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

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

At time $t + q - 2$, by equation 26,

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

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

can be cast as

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

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

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

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

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

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

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

Proof of Proposition 6.2. Market Clearing and Lemma C.4 imply that, for all $k \in \{0, \dots, K\}$,

$$\sum_{age=0}^{q-1} \delta_k(age) = 0 \tag{50}$$

and

$$\sum_{age=0}^{q-1} \delta(age) = q.$$

For $k = K$, it follows from equations 44 and 49

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

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

For $k = K-1$, by equations 44 and 48

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

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

By induction, for any $k \in \{q, \dots, K-2\}$, taking $\beta_{k+1} = 0$, it follows by equations 44 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 $age \in \{q, \dots, K\}$.

□

Proof of Lemma 6.1. By Proposition 6.1, we have the following demands:

$$x_t^{t-2} = 0 \tag{51}$$

$$x_t^{t-1} = \frac{E_t^{t-1} [s_{t+1}]}{\gamma r (1 + \beta_0) \sigma^2} = \frac{\alpha (1 - r) + l(0, 1) d_t + l(1, 1) d_{t-1}}{\gamma (1 + \beta_0)^2 \sigma^2} \tag{52}$$

$$x_t^t = \frac{E_{\Phi(m, s^2)} [s_{t+1}]}{\gamma r (1 + \beta_0) s^2} = \frac{\alpha (1 - r) + (\beta_1 - r\beta_0) d_t - r\beta_1 d_{t-1} + (1 + \beta_0) m}{\gamma r (1 + \beta_0)^2 s^2} \tag{53}$$

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

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

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

and

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

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

Therefore:

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

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

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

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

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

By Market clearing:

$$\begin{aligned}
1 &= \frac{1}{2} \left(\frac{\alpha(1-r) + l(0,1)d_t + l(1,1)d_{t-1}}{\gamma(1+\beta_0)^2\sigma^2} \right) \\
&\quad + \frac{1}{2} \left(\frac{\alpha(1-r) \left[1 - \frac{s^2}{\sigma^2} \frac{l(0,1)}{(1+\beta_0)} \right] + \left[\beta_1 - r\beta_0 + \frac{s^2}{\sigma^2} (1+\beta_0) \left(1 - \frac{l(1,1)l(0,1)}{(1+\beta_0)^2} \right) \right] d_t - r\beta_1 d_{t-1}}{\gamma r (1+\beta_0)^2 s^2} \right) \\
&= \frac{1}{2} \left(\frac{\alpha(1-r) + l(0,1)d_t + l(1,1)d_{t-1}}{\gamma(1+\beta_0)^2\sigma^2} \right) \\
&\quad + \frac{1}{2} \left(\frac{\alpha(1-r) \frac{\sigma^2}{s^2} \left[1 - \frac{s^2}{\sigma^2} \frac{l(0,1)}{(1+\beta_0)} \right] + \left[\frac{\sigma^2}{s^2} (\beta_1 - r\beta_0) + (1+\beta_0) \left(1 - \frac{l(1,1)l(0,1)}{(1+\beta_0)^2} \right) \right] d_t - \frac{\sigma^2}{s^2} r\beta_1 d_{t-1}}{\gamma r (1+\beta_0)^2 \sigma^2} \right),
\end{aligned}$$

which implies

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

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

$$0 = \alpha(1-r) \left[r + \frac{\sigma^2}{s^2} - \frac{l(0,1)}{1+\beta_0} \right] - 2r\gamma(1+\beta_0)^2\sigma^2 \quad (54)$$

$$0 = l(0,1) + \frac{1}{r} \frac{\sigma^2}{s^2} (\beta_1 - r\beta_0) + \frac{1}{r} (1+\beta_0) \left(1 - \frac{l(1,1)l(0,1)}{(1+\beta_0)^2} \right) \quad (55)$$

$$0 = l(1,1) - \frac{\sigma^2}{s^2} \beta_1 \quad (56)$$

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

Proof of Proposition 6.3. Throughout the proof, let $w_0 \equiv w(0,\lambda,0)$.

We know from Lemma 6.1 that $\{\alpha, \beta_0, \beta_1\}$ solve the system of equations given by (??) and (??) and ??.

STEP 1. By equation ??,

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

We note that $r > 1 \geq w(0, \lambda, 0)$, thus, if $0 < \beta_1 < r\beta_0$ and $1+\beta_0 > 0$, then $\left[r + \frac{\sigma^2}{s^2} - \frac{[(1+\beta_0)w(0,\lambda,0) + \beta_1 - r\beta_0]}{1+\beta_0} \right] > 0$ and $\alpha \leq 0$.

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

For $1+\beta_0 > 0$, equation (??) 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 (??) to hold it must be that $1 - \frac{1}{r} \frac{l(1,1)}{1+\beta_0} < 0$.

$$1 - \frac{1}{r} \frac{l(1,1)}{(1+\beta_0)^2} = 1 - \frac{1}{r} \frac{(1+\beta_0)(1-w_0) - r\beta_1}{(1+\beta_0)} \quad (57)$$

$$= 1 - \frac{1}{r} (1-w_0) + \frac{\beta_1}{1+\beta_0} > 0 \quad (58)$$

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 (??):

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

We plug this into equation (??) and we obtain:

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

Note that this is a linear equation on β_0 , i.e.,

$$\beta_0 \left\{ \phi \left(\frac{1-w_0}{\phi+r} - r \right) + r \left[\frac{1-w_0}{\phi+r} + w_0 - r \right] + 1 - \frac{\phi(1-w_0)(1-\phi r - r^2 + (\phi+r-1)w_0)}{(\phi+r)^2} \right\} \\ + \phi \left(\frac{1-w_0}{\phi+r} \right) + r \left[\frac{(1-w_0)}{\phi+r} + w_0 \right] + \left[1 - \frac{\phi(1-w_0)(1+(\phi+r-1)w_0)}{(\phi+r)^2} \right].$$

Therefore,

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

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

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

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

□

In order to show Proposition 6.4, we need the following Lemmas (their proofs are relegated to the end of the section).

Lemma C.5 (l: sign1). *For $\lambda \geq 0, 1 + \beta_0 + \beta_1 - r\beta_0 > 0$.*

Lemma C.6 (lem:risk-demands-q2). *Given our linear guess for prices (10), when $q = 2$, at time t :*

$$x_t^{t-1} = \frac{E_t^{t-1}[s_{t+1}]}{\gamma r (1 + \beta_0) \sigma^2} = \frac{\alpha (1-r)}{\gamma (1 + \beta_0)^2 \sigma^2} + \frac{l(0, 1)}{\gamma (1 + \beta_0)^2 \sigma^2} d_t + \frac{l(1, 1)}{\gamma (1 + \beta_0)^2 \sigma^2} d_{t-1} \quad (59)$$

$$x_t^t = \frac{E_{\Phi(m, s^2)}[s_{t+1}]}{r(1 + \beta_0) s^2} = \delta(0) + \delta_0(0) d_t + \delta_1(0) d_{t-1} \quad (60)$$

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

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

Proof of Proposition 6.4. 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

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

And $\frac{\partial x_t^t}{\partial d_t} = \delta_0(0) = -\frac{\partial x_t^{t-1}}{\partial d_t}$, $\frac{\partial x_t^t}{\partial d_{t-1}} = \delta_1(0)$, and $\frac{\partial x_t^{t-1}}{\partial d_t} = \frac{l(0,1)}{\gamma(1+\beta_0)^2 \sigma^2}$ and $\frac{\partial x_t^{t-1}}{\partial d_{t-1}} = \frac{l(1,1)}{\gamma(1+\beta_0)^2 \sigma^2} = -\frac{\partial x_t^t}{\partial d_{t-1}}$.

Therefore, it suffices to show that $l(0,1) < 0$ and $\delta_1(0) < 0$.

By proposition 6.3, $\beta_1 > 0$ and $\beta_0 > 0$ and thus $\delta_1(0) = -\frac{r\beta_1}{r\gamma(1+\beta_0)s^2} < 0$. So it only remains to show that $l(0,1) < 0$.

We now show that $l(0,1) < 0$. From the equilibrium condition (??) we have:

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

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

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

$$F(x) \equiv ax^2 + bx + c$$

with $a = \beta_1 - r\beta_0 < 0$ (by Proposition 6.3), $b = r(1+\beta_0) - l(1,1) = r(1+\beta_0) - (1+\beta_0)w(1, \lambda, 0) + r\beta_1 > 0$ (by Proposition 6.3) and $c = 1 + \beta_0 + \beta_1 - r\beta_0 > 0$ (by Lemma C.5). Thus: F is convex and $F(0) = c > 0$. From FOC $a2x^* + b = 0 \Rightarrow x^* = -\frac{b}{2a} > 0$. Let's focus on x_2 . Therefore, $F(x)$ has two roots x_1, x_2 with $x_1 < 0 < x^* < x_2$, where

$x^* = \arg \max_{x \in \mathbb{R}} F(x)$.

We now show that $x_2 = \frac{l(0,1)}{1+\beta_0}$ cannot be a solution. Suppose not, that is assume that our solution is the positive root $\frac{l(0,1)}{1+\beta_0} = x_2$, then:

$$-\frac{b}{2a} < \frac{l(0,1)}{1+\beta_0} \quad (61)$$

$$\frac{r(1+\beta_0) - l(1,1)}{2[-(\beta_1 - r\beta_0)]} < \frac{l(0,1)}{1+\beta_0} \quad (62)$$

$$\frac{r(1+\beta_0) - l(1,1)}{2} < l(0,1) \frac{r\beta_0 - \beta_1}{1+\beta_0} \quad (63)$$

Let $Z \equiv -\frac{\beta_1 - r\beta_0}{1+\beta_0}$

$$r(1+\beta_0) - (1+\beta_0)(1-w_0) + r\beta_1 < 2l(0,1)Z$$

$$r(1+\beta_0) - (1+\beta_0)(1-w_0) + r\beta_1 < 2Z[(1+\beta_0)w_0 + \beta_1 - r\beta_0]$$

$$r - 1 + w_0 + r\frac{\beta_1}{(1+\beta_0)} < 2Z\left[w_0 + \frac{\beta_1 - r\beta_0}{(1+\beta_0)}\right]$$

$$Z(w_0 - Z) > 0.5w_0 + \frac{1}{2}\left[r - 1 + r\frac{\beta_1}{1+\beta_0}\right]$$

$$\frac{w_0}{4} > 0.5w_0 + \frac{1}{2}\left[r - 1 + r\frac{\beta_1}{1+\beta_0}\right].$$

Observe that $\frac{1}{2}\left[r - 1 + r\frac{\beta_1}{1+\beta_0}\right] > 0$ and thus a contradiction follows. The solution must be the negative root.

□

Proof of Lemma ??. Let $\psi = d_t - d_{t-1}$ denote the change in dividends. The demand of young and adult agents expressed in terms of d_{t-1} and ψ are given by:

$$x_t^t = \frac{\alpha(1-r)}{\gamma r(1+\beta_0)^2 \sigma^2} + \frac{1+\beta_0 + \beta_1 - r\beta_0 - r\beta_1}{\gamma r(1+\beta_0)^2 \sigma^2} d_{t-1} + \frac{1+\beta_0 + \beta_1 - r\beta_0}{\gamma r(1+\beta_0)^2 \sigma^2} \psi + \Delta(d_{t-1}, \psi) \quad (64)$$

$$x_t^{t-1} = \frac{\alpha(1-r)}{\gamma(1+\beta_0)^2 \sigma^2} + \frac{1+\beta_0 + \beta_1 - r\beta_0 - r\beta_1}{\gamma(1+\beta_0)^2 \sigma^2} d_{t-1} + \frac{(1+\beta_0)w_0 + \beta_1 - r\beta_0}{\gamma(1+\beta_0)^2 \sigma^2} \psi \quad (65)$$

Let $x \equiv \frac{\alpha(1-r)}{\gamma(1+\beta_0)^2\sigma^2} + \frac{1+\beta_0+\beta_1-r\beta_0-r\beta_1}{\gamma(1+\beta_0)^2\sigma^2}d_{t-1}$. Then from Market Clearing $:x_t^{t-1} + x_t^t = 2$:

$$x = \frac{2r}{1+r} - \frac{r}{1+r}\Delta(d_t) - \left[\frac{(1+\beta_0+\beta_1-r\beta_0)}{\gamma(1+\beta_0)^2\sigma^2} - \frac{r}{1+r} \frac{(1+\beta_0)(1-w_0)}{\gamma(1+\beta_0)^2\sigma^2} \right] \psi$$

Plugging this back into x_t^t , and comuting tr_t^t we obtain:

$$tr_t^t = x_t^t - x_{FI}^t = \frac{1}{1+r} \frac{(1+\beta_0)(1-w_0)}{\gamma(1+\beta_0)^2\sigma^2} \psi + \frac{r}{1+r}\Delta(d_t)$$

The formula for trade volume follows. Now we're interested in understanding $\frac{\partial TR_t}{\partial \psi} = \left(\frac{1}{1+r}\right)^2 \times tr_t^t \times \frac{\partial tr_t^t}{\partial \psi} \propto tr_t^t \times \frac{\partial tr_t^t}{\partial \psi}$ near $\psi = 0$.

$$\Delta(d_{t-1}, \psi) = \frac{1}{\gamma r (1+\beta_0)^2 \sigma^2} \times \left[\alpha(1-r) \left(\frac{l(0,1)^2}{(1+\beta_0)^2} - \frac{l(0,1)}{(1+\beta_0)} \right) + \left((\beta_1 - r\beta_0 - r\beta_1) \frac{l(0,1)^2}{(1+\beta_0)^2} - l(1,1)l(0,1) \right) d_{t-1} - l(1,1)l(0,1)\psi \right]$$

Note that $\frac{\partial tr_t^t}{\partial \psi} = \frac{1}{1+r} \left[\frac{(1+\beta_0)(1-w_0)}{\gamma(1+\beta_0)^2\sigma^2} - \frac{l(1,1)l(0,1)}{\gamma(1+\beta_0)^2\sigma^2} \right] > 0$. Thus, we need to pin down the sign of tr_t^t to understand how trade volume would react to changes in dividends. This will depend on initial level of dividends d_{t-1} .

Case A: $(d_{t-1}, \psi = 0)$ is such that $tr_t^t(d_{t-1}, 0) \leq 0$. Therefore, since $\frac{\Delta tr_t^t}{\Delta \psi} > 0$ for all d_{t-1} , there exists shock large enough $\bar{\psi} > 0$ such that $tr_t^t(d_{t-1}, \bar{\psi}) = 0$ and for all $\psi \geq \bar{\psi}$, $tr_t^t(d_{t-1}, \psi) \geq 0$. Therefore, from a region in which trade volume was negative, there exists a positive shock to dividends large enough that it increases trade volume. It is clear that in this scenario, negative shocks to dividends $\psi \leq 0$, increase trade volume.

Case B: $(d_{t-1}, \psi = 0)$ is such that $tr_t^t(d_{t-1}, 0) \geq 0$. By the same argument, there exists $\underline{\psi} \leq 0$ such that $tr_t^t(d_{t-1}, \underline{\psi}) = 0$ and thus for all $\psi \leq \underline{\psi}$, $tr_t^t(d_{t-1}, \psi) \leq 0$, thus, trade volume increases when ψ falls below $\underline{\psi}$. Again, it is straightforward in this case that any positive ψ would increase trade levels.

□

C.1 Proof of Lemmas C.1, C.2, C.3, C.4 and C.5 and C.6

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

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

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

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

□

Proof of Lemma C.2. At time $t + q$, an agent born in t is in the last period of his life, consuming all of its wealth. Therefore, he will sell all of its claims to the assets it 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: $x_{t+q}^t = 0, a_{t+q}^t = 0$. Given this, we can compute the portfolio choice of an agent with age $q - 1$, who does 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_{t+q-1}^t :

$$\max_x E_{t+q-1}^t \left[-\exp(-\gamma(W_{t+q-1}^t + x s_{t+q})) \right] = \max_x E_{t+q-1}^t \left[-\exp(-\gamma x s_{t+q}) \right] \quad (66)$$

At time $t + q - 1$, the only random variable is d_{t+q} , which is normally distributed, and thus

$s_{t+q} \sim N(E_{t+q-1}^t[s_{t+q}]; (1 + \beta_0)\sigma^2)$. Given this, the agent's problem becomes:

$$V_{t-1}^{t-q} \equiv \max_x \left[-\exp\left(-\gamma x E_{t-1}^{t-q}[s_t] + \frac{1}{2}\gamma^2 x^2 (1 + \beta_0)\sigma^2\right) \right] \quad (67)$$

$$\max_x x E_{t-1}^{t-q}[s_t] - \frac{1}{2}\gamma x^2 (1 + \beta_0)^2 \sigma^2 \quad (68)$$

And therefore, by FOC:

$$x_{t+q-1}^t = \frac{E_{t+q-1}^t[s_{t+q}]}{\gamma \sigma_*^2} \quad (69)$$

□

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

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

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

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

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

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

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

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

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

□

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

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

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

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

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

$$\delta(q) = \delta_k(q) = 0, \quad \forall k \in \{0, \dots, K\} \quad (70)$$

$$\delta(q-1) = \frac{\alpha(1-r)}{\gamma((1+\beta_0)\sigma)^2}, \quad \delta_k(q-1) = \frac{(1+\beta_0)w(k, \lambda, q-1) + \beta(k)}{\gamma((1+\beta_0)\sigma)^2}, \quad \forall k \in \{0, \dots, q-1\} \quad (71)$$

$$\delta_k(q-1) = \frac{\beta(k)}{\gamma((1+\beta_0)\sigma)^2}, \quad \forall k \in \{q, \dots, K\}. \quad (72)$$

We also know from Lemma B.1 that

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

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

STEP 2. We now derive the risky demand and continuation value for generation aged $q-2$. The problem of generation aged $q-2$ at time t is given by,

$$\max_x E_t^{t-q+2} \left[V_{t+1}^{t-q+2} \exp(-\gamma r x s_{t+1}) \right]. \quad (73)$$

By the calculations in step 1, and using Λ_t as defined in (70), this problem becomes:

$$V^{q-2}(d_{t-K}, \dots, d_t) \quad (74)$$

$$= \max_x E_t^{t-q+2} \left[-\exp\left(-\frac{1}{2}\left(x_t^{q-1}\right)^2 \gamma^2((1+\beta_0)s_{q-1})^2 - \gamma r x((1+\beta_0)d_{t+1} + \Lambda_t)\right) \right]. \quad (75)$$

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

Observe that

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

and that future dividends are the only random variable, with $d_{t+1} \sim N\left(\theta_t^{t-q+2}, \sigma^2\right)$. There-

fore, by Lemma C.3, and with:

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

we obtain:

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

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

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

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

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

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

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

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

The maximization problem is given by:

$$V^{age}(d_{t-K}, \dots, d_t) \equiv \max_x E_t^{t-age} \left[V_{t+1}^{t-age-1} \exp(-\gamma r^{q-1-age} x ((1 + \beta_0) d_{t+1} + \Lambda_t)) \right]. \quad (76)$$

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

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

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

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

Therefore

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

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

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

□

Proof of Lemma C.5. Assume it is not: $1 + \beta_0 + \beta_1 - r\beta_0 \leq 0$. This implies that $l(0, 1) = (1 + \beta_0)w_0 + \beta_1 - r\beta_0 \leq 0$ From condition (??) 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 6.3, for the previous equation to hold it must be that $\left[r - \frac{l(1, 1)}{(1 + \beta_0)}\right] \leq 0$.

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

Thus, $[1 + \beta_0 + \beta_1 - r\beta_0] > 0$. □

Proof of Lemma C.6. From Lemma B.1, we know that $x_t^{t-1} = \frac{E_t^{t-1}[s_{t+1}]}{\gamma(1 + \beta_0)\sigma^2}$. Therefore, given our guess for prices and Lemma 6.2, we have:

$$x_t^{t-1} = \frac{E_t^{t-1}[d_{t+1} + p_{t+1} - rp_t]}{\gamma(1 + \beta_0)\sigma^2} \quad (77)$$

$$= \frac{(1 + \beta_0)\theta_t^{t-1} + \alpha(1 - r) + (\beta_1 - r\beta_0)d_t - r\beta_1 d_{t-1}}{\gamma(1 + \beta_0)\sigma^2} \quad (78)$$

since $\theta_t^{t-1} = w_0 d_t + (1 - w_0) d_{t-1}$, we obtain equation (59), where $l(0, 1) = (1 + \beta_0)w_0 + \beta_1 - r\beta_0$

and $l(1, 1) = (1 + \beta_0)(1 - w_0) - r\beta_1$. We also know from Lemma C.2 that

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

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

$$V_t^{t-1} = -\exp(-A_t - B_t d_t - C d_t^2) \quad (79)$$

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:

$$\max_x E_t^t [V_{t+1}^t \exp(-\gamma r x s_{t+1})] \quad (80)$$

From Lemma C.3:

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

Where,

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

with $m = \frac{\theta_t^t - \sigma^2 B_{t+1}}{2C\sigma^2 + 1}$, $s^2 = \frac{\sigma^2}{2C\sigma^2 + 1}$. Incorporating the fact that $B_{t+1} = \frac{(\alpha(r-1) + l(1, 1)d_t)l(0, 1)}{(1 + \beta_0)^2\sigma^2}$ and $\theta_t^t = d_t$ we obtain equation (60) and the respective δ_s .

□

D Data Appendix

Our source of household-level microdata is the Survey of Consumer Finances (SCF), which provides repeated cross-section observations on asset holdings and various household background characteristics. Our sample has two parts. The first one is the standard SCF from 1983 to 2013, obtained from the Board of Governors of the Federal Reserve System and available every three years. The second source is the precursor of the "modern" SCF, obtained from the Inter-university Consortium for Political and Social Research at the University of Michigan. The precursor surveys start in 1947, partly annually, but with some gaps. The data before 1960 contains information on stock holdings in some years, but age is measured in 5 or 10-year brackets, which would make our measurement of experienced returns imprecise, particularly for younger individuals. For this reason, we start in 1960 and use all survey waves that offer stock-market participation information, i.e., the 1960, 1962, 1963, 1964, 1967, 1968, 1969, 1970, 1971, and 1977 surveys.

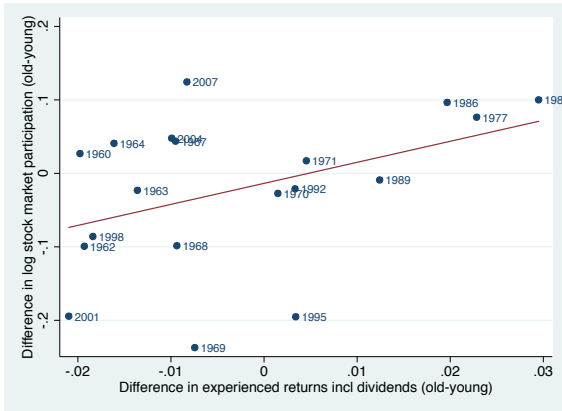
The first measure is a binary variable for stock-market participation, available in each survey wave from 1960-2013. It indicates whether a household holds more than zero dollars worth of stocks. We define stock holdings as the sum of directly held stocks (including stock held through investment clubs) and the equity portion of mutual fund holdings. In our main tests, we include stocks held in retirement accounts (e.g., IRA, Keogh, and 401(k) plans). 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. The Appendix provides the details. Online Appendix F reports robustness checks that exclude retirement account holdings from the analysis.

Our second measure of stock market participation is the fraction of liquid assets invested in stocks. The share of directly held stocks plus the equity share of mutual funds can be

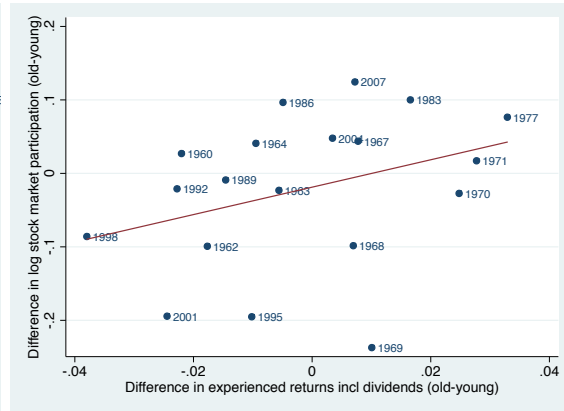
calculated in all surveys from 1960-2013 other than 1971. Liquid assets are defined as stock holdings plus bonds plus cash and short-term instruments (checking and savings accounts, money market mutual funds, certificates of deposit).

The 1983-2013 waves of the SCF oversample high-income households. The oversampling provides a substantial number of observations on households with significant stock holdings, which is helpful for our analysis of asset allocation, but could also induce selection bias. In our main tests, we weight the data using SCF sample weights, which undo the overweighting of high-income households and which also adjust for non-response bias. The weighted estimates are representative of the U.S. population.

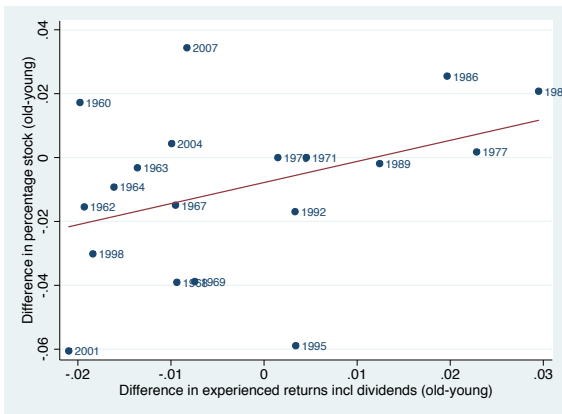
E Figures



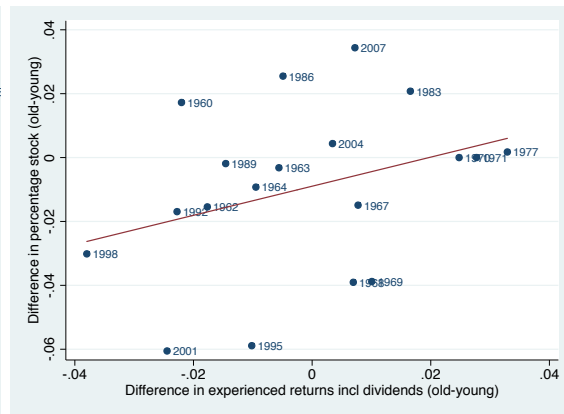
(a) Stock Market Participation. ($\lambda = 1$)



(b) Stock Market Participation. ($\lambda = 3$)



(c) Fraction of Stock of Financial Assets. ($\lambda = 1$)



(d) Fraction of Stock of Financial Assets. ($\lambda = 3$)

Figure 7: Experienced Returns (including dividends) and Stock Holding

References

- Adam, K., A. Marcet, and J. P. Nicolini (2012). Stock Market Volatility and Learning. *Journal of Finance*, forthcoming.
- Alesina, A. and N. Fuchs-Schundeln (2007). Good-bye Lenin (or Not?): The Effect of Communism on People's Preferences. *American Economic Review* 97, 1507–1528.
- Barsky, R. and J. B. DeLong (1993). Why Does the Stock Market Fluctuate? *Quarterly Journal of Economics* 108(2), 291–311.
- Blanchard, O. (2012). Sustaining a Global Recovery. *Finance and Development* 46(3), 9–12.
- Cecchetti, S. G., P. Lam, and N. C. Mark (2000). Asset Pricing with Distorted Beliefs: Are Equity Returns Too Good to Be True? *American Economic Review* 90(4), 787–805.
- Cogley, T. and T. J. Sargent (2008). The Market Price of Risk and the Equity Premium: A Legacy of the Great Depression? *Journal of Monetary Economics* 55(3), 454–476.
- DeLong, B. and L. Summers (2012). Fiscal Policy in a Depressed Economy. *Brookings Papers on Economic Activity* 44(1), 233–297.
- Ehling, P., A. Graniero, and C. Heyerdahl-Larsen (2015). Asset Prices and Portfolio Choice with Learning from Experience. *working paper*, NYU.
- Friedman, M. and A. J. Schwartz (1963). *A Monetary History of the United States*. Princeton, NJ: Princeton University Press.
- Greenwood, R. and A. Shleifer (2014). Expectations of Returns and Expected Returns. *Review of Financial Studies*, 714–746.
- Hertwig, R., G. Barron, E. U. Weber, and I. Erev (2004). Decisions from Experience and the Effect of Rare Events in Risky Choice. *Psychological Science* 15, 534–539.
- Kaustia, M. and S. Knüpfer (2008). Do Investors Overweight Personal Experience? Evidence from IPO Subscriptions. *Journal of Finance* 63, 2679–2702.
- Malmendier, U. and S. Nagel (2011). Depression Babies: Do Macroeconomic Experiences Affect Risk-Taking? *Quarterly Journal of Economics* 126, 373–416.
- Malmendier, U. and S. Nagel (2013). Learning from Inflation Experiences. *Quarterly Journal of Economics*, forthcoming.
- Malmendier, U. and L. S. Shen (2015). Experience Effects in Consumption. Working paper, UC-Berkeley.
- Timmermann, A. G. (1993). How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices. *Quarterly Journal of Economics* 108(4), 1135–1145.
- Timmermann, A. G. (1996). Excess Volatility and Predictability of Stock Prices in Autoregressive Dividend Models with Learning. *The Review of Economic Studies* 63(4), 523–557.

- Tversky, A. and D. Kahneman (1974). Judgment under uncertainty: Heuristics and biases. *Science* 4157, 1124–1131.
- Vives, X. (2008). *Information and Learning in markets: The Impact of Market Microstructure*.
- Weber, E. U., U. Bockenholt, D. J. Hilton, and B. Wallace (1993). Determinants of Diagnostic Hypothesis Generation: Effects of Information, Base Rates, and Experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19, 1151 C1164.