A Theory of Experience Effects*

Ulrike Malmendier†
UC Berkeley, NBER, and CEPR

Demian Pouzo‡
UC Berkeley

Victoria Vanasco§
Stanford University

VERY PRELIMINARY AND INCOMPLETE - DO NOT DISTRIBUTE

July 6, 2015

Abstract

How does the experience of a financial crises and other macroeconomic shocks 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 and borrowing decisions. We propose a simple OLG model as a theoretical underpinning of experience-based learning. Risk averse investors invest in a 'Lucas tree' and a risk-free asset. They form beliefs based on data observed during their lifetime so far. We show that, in equilibrium, prices depend only on the dividends observed by the generations that are alive, and 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 prices, stock market participation, and trading volume. First, the more agents in an economy rely on recent observations, the more volatile are prices and the higher is the autocorrelation of prices. Second, the stronger the disagreement across generations (e.g. after a recent shock), the higher is the trade volume. Third, a recent crisis will increase the average age of stock market participants, while periods of stock-market boom have the opposite effect.

*We thank workshop participants at UC Berkeley and SITE (Psychology and Economics segment) for helpful comments.

†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 a better understanding of 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) argues 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 financial markets is consistent with growing empirical evidence on experience effects. This evidence suggests that individuals overweight personal experiences of macroeconomic shocks when forming beliefs, and that personal experiences appear to leave an imprint on individuals' 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 investigate the long-term effects of personal experiences on market participation and portfolio decisions of different cohorts in an economy. We derive implications for prices, volatility, and trade volume. Our theoretical framework illustrates that a deeper understanding of the influence of past experiences is important to improve not only the micro-modeling of financial risk-taking, but also our understanding of the aggregate dynamics of financial decision-making and the long-run effects of macro-shocks.

We propose a stylized overlapping generations (OLG) general equilibrium model in which agents form their beliefs by overweighting their own experiences. We consider CARA investors that live for a finite number of periods and, during their lifetimes, choose portfolios of risky and risk-free securities to maximize their final wealth. Consumption takes place at the end of their lifes for tractability, which is standard in much of the literature (see Vives (2008). Investors can invest either in a risky asset, which is in unit net supply and pays random dividends every period (a Lucas Tree), or in a risk-less asset that 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.

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 who (i) only use data observed during their lifetime, and (ii) *may* overweight more recent observations when forming beliefs. These agents do observe the entire history of dividends (no asymmetric information) but choose not to use it to form their beliefs. This assumption captures, in a stylized manner, the availability bias underlying experience-based learning, including the possibility of a recency bias. In empirical implementations (and generalizations of the theoretical model) investors are using more historical data, but 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 have been confronted with different experiences in their lives so far (i.e., different life-time sequences of dividends) have differential reactions to macroeconomic shocks. This differential reaction to "booms" and "recessions" will affect equilibrium quantities such as trade volume. Third, the different investment horizons (young vs. old) also affect anticipated future trading behavior. In our framework, agents fully understand everyone's belief formation process. Hence, they 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 to incorporate what we will refer to as a *hedging motive.* Note that the latter a mechanism is different than simply belief heterogeneity.

The model is stylized and it allows us to fully isolate the forces introduced by the presence of experience-based learners. We focus on affine equilibria, i.e., equilibria wherein prices are affine functions of current and past dividends. In the benchmark case where agents know the true mean of dividends, 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, in our model environment, any departure from constant equilibrium prices and trade levels can be cleanly attributed to experience-based learning.

Our first result characterizes of each generation's demand for risky assets. In rational expectations portfolio models where agents know the true mean of dividends (i.e. no learning), investors' wealth in the distant future is independent of next period returns. Hence, even though agents face a multi-period investment problem, their demands in each period coincide with those of a static problem. Thus, the multi-period investment problem can be partitioned into a sequence of one-period ones (Vives (2008)). Under experience-based learning, 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 periods returns and the simplification to a sequence of static problem no longer applies. However, exploiting the CARA-Gaussian setup, we 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, reflecting the fact that future wealth is correlated with next period payoff.

Second, we derive the model prediction about prices, which capture the intuition of Friedman and Schwartz (1963) that a recession (understood as a negative shock to dividends) can have a long-lasting effect on equilibrium outcomes. We show that, in our model, equilibrium prices are only a function of the dividends observed by the generations that are alive and actively trading in the market. That is, prices (and thus returns) are predictable; and only dividends observed by the generations that are alive are relevant for predictability. This observation proposes a novel link between factors predicting long-run prices (and returns) and investors' past experiences. Our second finding also has implications for price volatility. The resulting price volatility goes above and beyond the volatility of the assumed dividend process. Both features stem from the learning mechanism in our model: If agents know the true mean of dividends (or their beliefs converge to the truth), this setup yields constant prices.

Third, we characterize the heterogeneity in demand for risky assets and portfolio decisions across cohorts, focusing on the simple case where agents live two periods. We show that 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. However, the full mechanism is more complex, and we can decompose agents' demands for the risky asset into three components: a beliefs term, a hedging-motive term, and a horizon term. The *beliefs effect* captures that an increase (decrease) in dividends makes younger agents more optimistic (pessimistic) about the return of the risky asset than older agents. Therefore, they demand more (less) of the risky asset in response to a positive shock to dividends. The *hedging effect* captures that agents anticipate they will learn about the risky asset from future dividends. As a result, they distort their portfolio decisions to hedge their exposure to changes in beliefs. The *horizon effect* is the least interesting. It indicates that even when agents share beliefs, young agents react less aggressively to a change in dividends (in their beliefs) due to their longer remaining investment horizon. We show that the belief effect always dominates. As a result, the demand of young agents reacts positively more strongly to changes in dividends.

Finally, we derive the implications of experience-based on trade volume. We show that the presence of learning and disagreements generate positive trade volume in equilibrium through two channels. First, an increase (decrease) in dividends induces trade since young agents become more optimistic (pessimistic) than old agents, and disagreement generates gains from trade. Second, agents trade due to the hedging motive and is present even in the absence of disagreements. Note that the first channel unambiguously predicts that changes in dividends increase trade volume, while the direction of the second channel is ambiguous. However, we are able to show that trade volume always increases in response to large enough changes in dividends.

Our findings 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 literature on 'experience effects' rationalizes these long-run effects empricially 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 modelling 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.

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 Model

Consider an infinite horizon economy $t \in \{0, 1, 2...\}$ with overlapping generations of a continuum of risk averse agents. Each generation is born every period and lives for q periods with $q \in \{1, 2, 3, ...\}$; that is, one generation born at each $t \geq 0$, and at any time t, there are $q+1$ generations alive. Generation born at time $t = n$ is called generation n. Within each generation there is a mass of q^{-1} identical agents.

Agents only consume in their final period, and have CARA preferences with risk aversion γ. They are born with no endowment, but can accumulate wealth during their lifetime by investing in financial markets (i.e. trading). There is a single risky asset (a Lucas Tree), that pays random dividends $d_t \sim N\left(\theta, \sigma^2\right)$ at time $t, \forall t$, and that is in unit net supply, and a riskless asset that is in perfectly elastic supply and pays $r > 1$ at all times. See Figure 1 for the timeline of this economy for $q = 2$, two-period lived generations.

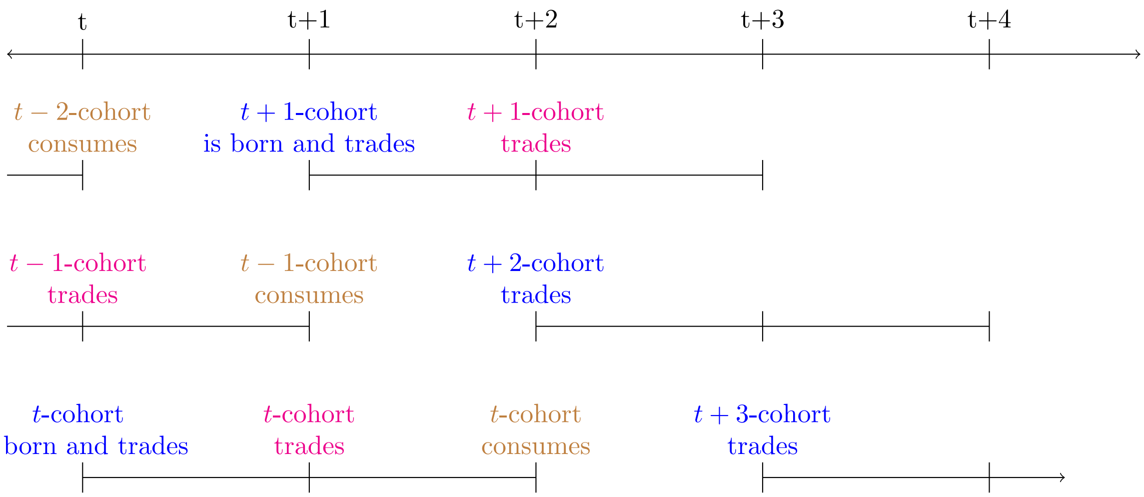

Figure 1: A timeline for an economy with two-period lived generations, $q = 2$.

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

$$W_t^n = x_t^n p_t + a_t^n \tag{1}$$

where W_t^n denotes the wealth of generation n at time t, x_t^n is the amount invested in the risky asset (number of shares on the Lucas Tree) and a_t^n is the amount invested in the riskless asset at time t by generation n, and p_t is the price of investing in one unit of the risky asset at time t. As a result, wealth next period is given by:

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

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 cost of investing in one unit of the risky asset at time t. Using this notation, $W_{t+1}^n = x_t^n s_{t+1} + rW_t^n$. For a given initial wealth level, W_n^n, the problem of a generation n is given by:

$$\max_{\{x_t^n\}_{t=nT}^{nT+q}} E_{nT}^n \left[-\exp(-\gamma W_{nT+q}^n) \right] \tag{3}$$

subject to 1-2, for all t. The operator $E_t^n[.]$ denotes the expectations computed with beliefs of generation n at period t.

2.1 Formation of Subjective Beliefs: 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 it and thus forecast dividends. To keep the model tractable, we assume that σ^2, the variance of dividends, is known at all times. It is important to note that we assume agents have *full information*, i.e., they observe the entire history of dividends.[1] However, they *choose not to use* observations outside their lifetime. Consistent with this, it is enough that agents learn only from dividends; prices do not add any additional information since the history of dividends is available to them. We make this assumption for simplicity, since all we need for our results to hold is that the history is heavily discounted when agents form their beliefs. In addition, we believe that adding private information and learning from prices to this framework would complicate matters without necessarily adding new intuition.

Experience-based learning (EBL) agents do not learn *about* the equilibrium, they learn *in* equilibrium. That is, agents understand the model and know all the primitives, except the mean of the dividend process. Also it is a *passive learning* problem, in the sense that actions of the players do not affect the information they receive. These two features make our problem different from reinforcement learning-type of problems. Note that if we have, say, participation, then that could be a link between action (e.g. participate or not) and learning/data. We consider this to be an interesting line to explore in the future.

We proceed to endogeneize the heterogeneity of beliefs across different cohorts by assuming

[1] Agents also fully understand the structure of the model.

that (1) agents do not know the mean of dividends (but know that dividends are Gaussian with variance σ^2), and (2) there is experience-based learning; that is, agents overweight observations received during their lifetime. 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 by the precision weighted average of dividends realized during the agent's lifetime, as in Malmendier and Nagel (2010). At any point in time t, generation n alive, with $age \equiv t - n$, forms its beliefs as follows:

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

where $w(k, \lambda, age)$ denotes the weight an agent aged age assigns to the $k - period$ before observation of dividends, and λ parametrizes these weights and will be described in what follows. Note that $\sum_{k=0}^{age} w(k, \lambda, age) = 1, \forall age$. We now characterize the probability measure implied by these weights.

2.1.1 The experience-based empirical probability measure

We now introduce the idea of experience-based empirical probability measure which allow us to extend the idea of experience-based learning to objects other than the mean. Given a realization of dividends $(d_\tau)_{\tau=0}^{t}$, let [2]

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

where, for any $k \leq a$

$$w(k, \lambda, a) = \frac{(a + 1 - k)^\lambda}{\sum_{k'=0}^{a} (a + 1 - k')^\lambda}, \ if \ a \geq 0 \tag{6}$$

be the *experience-based empirical probability measure* of generation n at period $t \in \{nT, ..., (n+q)T\}$. That is, this probability measure puts weight of $(w(k, \lambda, t - nT))_{k=0}^{t-nT}$ to the observa-

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

tions during the lifetime of the generation and zero to all other observations. This captures the assumptions that agents in this economy are experience-based learners. Within the observations during their lifetime, the generation puts weight $\frac{(a+1)^\lambda}{\sum_{k'=0}^{a}(a+1-k')^\lambda}$ to the most recent one, $\frac{(a+1-1)^\lambda}{\sum_{k'=0}^{a}(a+1-k')^\lambda}$ to the previous one, and so on. The parameter λ regulates the relative weight the most recent observations receive.

It is easy to see that for $\lambda > 0$, more recent observations receive relative more weights, whereas for $\lambda < 0$ the opposite holds. We now present some examples:

Example 1 (Linearly Declining Weights, $\lambda = 1$)**.** *With $\lambda = 1$, it is easy to see that, for any $0 \leq k, k+j \leq a$*

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

i.e., weights decay linearly. □

Example 2 (Equal Weights, $\lambda = 0$)**.** *With $\lambda = 0$, it is easy to see that, for any $0 \leq k \leq a$*

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

□

Example 3 (The Case with $\lambda = \infty$)**.** *If $\lambda \to \infty$, it follows that, for any $0 \leq k \leq a$*

$$w(k, \lambda, a) \to 1\{k = 0\}.$$

Hence, as λ diverges, the generation puts weight 1 to the most recent observation and 0 to all the rest. □

2.1.2 Beliefs about θ

The belief about θ of generation n at time $t \in \{n, ..., (n+q)\}$, from now on denoted as θ_t^n is thus the expectation of the dividends computed using the experience-based empirical probability measure, i.e.,

$$\theta_t^n = E_{\mathbb{P}_n}[d] = \sum_{k=0}^{t-n} w(k, \lambda, t-n) d_{t-k}. \tag{7}$$

Example 4. *For $q = 2$, it follows that*

$$\theta_t^t = d_t$$

$$\theta_{t+1}^t = d_{t+1} \frac{2^\lambda}{1+2^\lambda} + d_t \frac{1}{1+2^\lambda}$$

for all t. □

We conclude this section with a remark about the stochastic behavior of θ_t^n. By construction, $\theta_t^n \sim N(\theta, \sigma^2 \sum_{k=0}^{t-n} (w(k, \lambda, t-n))^2)$. Hence, whether θ_t^n converges to the truth as $t \to \infty$ will depend on whether $\sum_{k=0}^{t-n} (w(k, \lambda, t-n))^2 \to 0$; this in turn depends how fast the weights for "old" observations decay to zero. Note that when agents have finite lives, convergence will not occur. In addition, since separate cohorts weight different realizations differently, at any point in time we should expect belief heterogeneity driven by different experiences.

2.2 Connection to Bayesian Learners

For the sake of comparison, we compare our learning procedure with one where agents update their beliefs using Bayes rule. In principle, one can think of two sub-cases. The standard case, wherein agents use *all* the available observations from period 0 onwards to form their beliefs, and an alternative formulation where agents are learners from experience (in the sense they only use data observed during their lifetimes) but update their beliefs using Bayes rule. We explore both cases separately. We call the agents in the former case *Full Bayesian Learners* and in the latter case we call them *Bayesian Learners from Experience.*

This section shows that full Bayesian learners do not differ in their beliefs about the mean of the dividends and, eventually, it will converge to the truth. For EBL and Bayesian learners from experience this is not true (in fact we show that — for diffuse priors — Bayesian learners can be viewed as a particular case of EBL). These results illustrates the importance of the main feature of our experience-based learning: the fact that agents only use data observed during their lifetimes.

2.2.1 Full Bayesian Learners

In this case, all generations consider the whole set of observation from period 0 in order 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

$$\begin{aligned}
\gamma_{t+a} =& \frac{\tau^{-2}}{\tau^{-2}+\sigma^{-2}(t+a)} m + \frac{(t+a)\sigma^{-2}}{\tau^{-2}+\sigma^{-2}(t+a)} \sum_{k=0}^{t+a} d_{t+a-k} \frac{1}{t+a} \\
=& \frac{\tau^{-2}}{\tau^{-2}+\sigma^{-2}(t+a)} m + \frac{(t+a)\sigma^{-2}}{\tau^{-2}+\sigma^{-2}(t+a)} \left\{ \sum_{k=0}^{a} d_{t+a-k} \frac{1}{t+a} + \sum_{k=a+1}^{t+a} d_{t+a-k} \frac{1}{t+a} \right\} \\
=& \frac{\tau^{-2}}{\tau^{-2}+\sigma^{-2}(t+a)} m + \frac{(t+a)\sigma^{-2}}{\tau^{-2}+\sigma^{-2}(t+a)} \left\{ \theta^{t}_{t+a} \frac{a+1}{t+a} + \sum_{k=a+1}^{t+a} d_{t+a-k} \frac{1}{t+a} \right\}.
\end{aligned}$$

That is, the belief of a generation that is a full Bayesian learner is a convex combination of the prior and the average mean using *all* observations available to date. The key difference with our approach is that *all* generations alive in any given period will have the same belief; that is, the belief heterogeneity arising from different past experiences vanishes. Moreover, the beliefs are non-stationary (in the sense that depend on the time period) and as $t \to \infty$, the posterior mean converges (almost surely) to the true mean.

2.2.2 Bayesian Learners from Experience

For the Bayesian learner from experience the situation is different. First, we assume that the each generation t has a prior $N(m, \tau^2)$ when they are born (and not from $t = 0$). The

[3] The 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.

posterior mean of generation t at period $t+a$, β^t_{t+a}, is given by

$$\begin{aligned}\beta^t_{t+a} =& \frac{\tau^{-2}}{\tau^{-2}+\sigma^{-2}(a+1)}m + \frac{(a+1)\sigma^{-2}}{\tau^{-2}+\sigma^{-2}(a+1)}\sum_{k=0}^{a} d_{t+a-k}\frac{1}{a+1} \\ =& \frac{\tau^{-2}}{\tau^{-2}+\sigma^{-2}(a+1)}m + \frac{(a+1)\sigma^{-2}}{\tau^{-2}+\sigma^{-2}(a+1)}\theta^t_{t+a}.\end{aligned}$$

Now, the belief of a BL generation is a convex combination of the prior and the average mean using *only life-time* observations; in turn this average coincides with the belief of our learners from experience with $\lambda = 0$. That is, 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. Finally, we note that if the prior is diffuse, i.e., $\tau = \infty$, then β^t_{t+a} coincides with θ^t_{t+a} for $\lambda = 0$. [4]

2.3 Characterization of Demands for the Risky Assets under Affine Prices

In this section we characterize the portfolio choice and resulting demand for the risky asset of the different cohorts in a linear equilibrium. 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 learning and the fact that agents are sophisticated enough to understand how their beliefs evolve over their lifetime introduce a correlation between future returns and continuation values that distorts the portfolio decisions. This observation notwithstanding, we show that the agents dynamic portfolio problem can be expressed as a *adjusted static* problem where dividends follow a normal distribution with *adjusted* mean and variance. For the $q = 2$ case, we show that the adjusted distribution of dividends has a lower variance than the actual distribution of dividends. These adjustments result from learning making the value function of the agent *less* concave; that is, very high and very low realizations of future dividends are now associated

[4]The formal argument relies on looking at the limit of β^t_{t+a} as $\tau \to \infty$.

with higher continuation values.

We focus on the case where prices are an affine function of K past dividends, for some K such that $q \leq K < \infty$. Henceforth, we study equilibria wherein prices are an affine function of dividends. For some $\alpha_0 \in \mathbb{R}$ and $\beta_k \in \mathbb{R}$ for $k \in \{0, ..., K\}$ such that:

$$p_t = \alpha_0 + \sum_{k=0}^{K} \beta_k d_{t-k} \tag{8}$$

for all $t \geq K$.

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

$$\max_{x_{t:t+q-1}=(x_t,...,x_{t+q-1})\in\mathbb{R}^q} E_t^t \left[-\exp\left(-\gamma W_{t+q}^n(x_{t:t+q-1})\right)\right] \tag{9}$$

$$s.t. \quad W_{t+q}^t(x_{t:t+q-1}) = \sum_{\tau=t}^{t+q-1} r^{t+q-1-\tau} x_\tau s_{\tau+1} \tag{10}$$

For simplicity, we 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

$$V_{t+q-1}^t(d_{t+q-1-K:t+q-1}) = \max_{x\in\mathbb{R}} E_{t+q-1}^t \left[-\exp\left(-\gamma s_{t+q} x\right)\right] \ \textit{and} \tag{11}$$

$$V_\tau^t(d_{\tau-K:\tau}) = \max_{x\in\mathbb{R}} E_\tau^t \left[V_{\tau+1}^t(d_{\tau+1-K:\tau+1}) \exp\left(-\gamma s_{\tau+1} x\right)\right], \ \forall \ \tau \in \{t, ..., t+q-2\} \tag{12}$$

Remark 5. *Notice that V_τ^t does not include the wealth at time τ, that is, from equation 9, the optimization problem can be cast as $\max_{x\in\mathbb{R}} \exp\{-\gamma r W_{t+q-1}^t\} E_{t+q-1}^t \left[-\exp\left(-\gamma s_{t+q} x\right)\right]$. 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 $\tau \in \{t, ..., t + q - 2\}$, 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 $V^t_{\tau+1}(d_{\tau+1-K:\tau+1})$ by affecting the beliefs of the t-generation at $\tau+1$, and the resulting portfolio decision.

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, because of the presence of learning, this will not be the case. However, Proposition 6 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 $E_{N(\mu,\sigma^2)}[.]$ and $V_{N(\mu,\sigma^2)}[.]$ be the expectation and variance with respect to a Gaussian pdf with mean μ and σ^2.

Proposition 6 (pro: demands)**.** *Suppose $p_t = \alpha_0 + \sum_{k=0}^{K} \beta_k d_{t-k}$. For any generation t in period $t+j$ for $j \in \{0, ..., q-1\}$ (the age of the generation), demands for the risky asset are given by:*

$$x^t_{t+j} = \frac{E_{N(m_j,\sigma_j^2)}[s_{t+j+1}]}{\gamma r^{q-1-j} V_{N(m_j,\sigma_j^2)}[s_{t+j+1}]} \tag{13}$$

where:

$$m_j \equiv \frac{\theta^t_{t+j} - \sigma^2 \left(b_j + \sum_{k=1}^{K} b_j(k) d_{t+j-k} \right)}{2c_j\sigma^2 + 1} \tag{14}$$

$$\sigma_j^2 \equiv \frac{\sigma^2}{2c_j\sigma^2 + 1} \tag{15}$$

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. See Appendix A.1. □

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 11). In particular we show that V^t_{t+q-1} is of the form exponential-quadratic in d_{t+q-1} (see Lemma 6 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 3 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 13 and V^t_{t+q-2} is also of the exponential-quadratic form. The process thus continues until time t.

Remark 7. *From equation 13, we can cast the optimal demand at time $t+j$ as* [5]

$$x^t_{t+j} = \frac{E_{N(\theta^t_{t+j},\sigma^2)}[s_{t+j+1}]}{\gamma r^{q-1-j} V_{N(\theta^t_{t+j},\sigma^2)}[s_{t+j+1}]} - \frac{(b_j + \sum_{k=1}^{K} b_j(k) d_{t+j-k})}{\gamma r^{q-1-j}(1+\beta_0)}. \tag{16}$$

The first term coincides with the demand of a static *portfolio problem for an agent with beliefs θ^t_{t+j}. The second term $\frac{(b_j+\sum_{k=1}^{K} b_j(k) d_{t-k})}{\gamma r^{q-1-j}(1+\beta_0)}$, is an adjustment which accounts for the dynamic nature of the problem.*

Remark 8. *From equation 16, it is not hard to show (which we still show in lemma 7 in the Appendix) that demands at time t are affine in $d_{t-K:t}$. From the derivations, it can be seen that this is because prices are stated as a function of dividends from t to $t-K$ (while beliefs about future dividends depend on the history observed by a given generation).*

This observation is the basis of Proposition 9 below where we show that, in a linear equilibrium, prices will only depend on the history of dividends observed by the oldest generation in the market. This result in turn, also implies that demands at time t will also only depend on $d_{t-q:t}$.

2.4 Characterization of the Linear Equilibrium

We now establish that in a linear equilibrium prices at any time t only depend on the dividends observed by the generations trading at time t.

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

Proposition 9. *For $r > 1$, the price in any linear equilibrium is affine in the history of dividends observed by the oldest generation participating in the market. For any $t \geq 0, q \geq 1$,*

$$p_t = \alpha_0 + \sum_{k=0}^{q-1} \beta_k d_{t-k}. \tag{17}$$

Proof of Proposition 9. See Appendix A.3. □

The idea of the proof is as follows. By lemma 7, 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 are definition of linear equilibrium. The proof shows that under market clearing, the coefficients accompanying the dividends $d_{t-K:t-q}$ are zero.

As a result, prices and demands depend only on the history of 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.

This result captures the belief channel described by Friedman and Schawrtz: 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. This suggests that agents may use different sets of past data to predict dividends (fundamentals) and prices. To form expectations about future dividends agents use dividends observed during their lifetime, while to form expectations about future prices agents look at the history of dividends observed by the cohort in the market.

We briefly discuss the implications of proposition 9 for the price dynamics.

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

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

The presence of learning introduces volatility and correlation to the price process, that would otherwise be constant. This can be seen in Figure 2. In addition, the numerical results show that the standard deviation of prices is increasing in w_0 (parametrized by λ), while decreasing in the riskfree rate. This is because higher λ is associated with a stronger response of old agent's beliefs to present dividends, that do follow a random process. In contrast, higher rates reduce the response of demands for the risky asset to changes in beliefs, and thus reduce the volatility of prices. The bottom panel shows that the correlation of prices is decreasing in w_0 (prices react more strongly to recent dividends relative to past dividends), and increasing in the riskfree rate. It is also important to highlight that prices are only positively correlated to the past prices observed by generations that are present in the market.

Predictability of Excess Returns. We note that the equilibrium excess return at time $t+j$ is given by $\frac{p_{t+j+1}+d_{t+j+1}}{p_{t+j}} - r = \frac{(1+\beta_0)d_{t+j+1}+\sum_{k=1}^{q-1}\beta_k d_{t+j+1-k}}{\sum_{k=0}^{q-1}\beta_k d_{t+j-k}} - r$. Thus, at time t and for $j \leq q-1$, the dividends $d_t, ..., d_{t+j-(q-1)}$ can be used as factors for predicting the excess returns. For $q > j-1$ 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 latter's impact on the level of disagreements across market participants. [Connect to the literature on disagreements and trade volume and return predictability.]

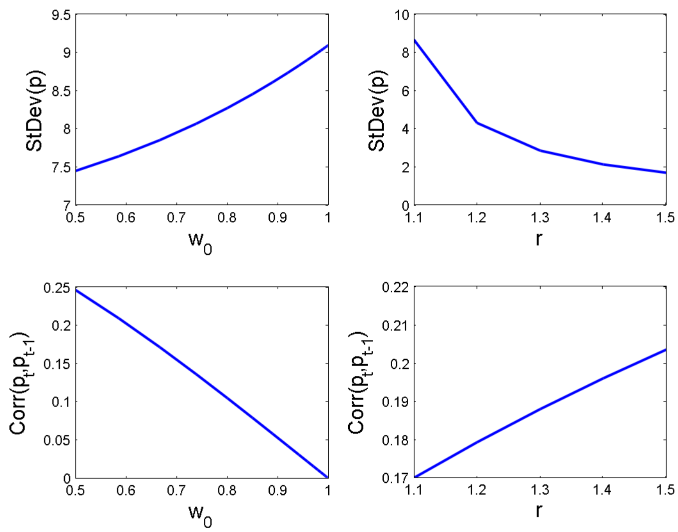

Figure 2: Comparative Statics on Standard Deviations and Correlations. The q=2 Case. $r = 1.11, \gamma = 10, \lambda = 1 (w_0 = \frac{2}{3}), \sigma^2 = 1$, unless otherwise noted.

3 Characterization of the Demands for Risky Assets for $q = 2$.

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 $\{\alpha_0, \beta_0, \beta_1\}$ solve a complicated system of non-linear equations

Lemma 1 (lem:prices-q2). *For $r > 1$ in any linear equilibrium prices are given by:*

$$p_t = \alpha_0 + \beta_0 d_t + \beta_1 d_{t-1} \qquad t \geq 1 \tag{18}$$

where $\{\alpha_0, \beta_0, \beta_1\}$ *solve the following system of equations:*

$$0 = \alpha_0 (1-r)\left[r + \frac{\sigma^2}{s^2} - \frac{[(1+\beta_0)w(0,\lambda,0) + \beta_1 - r\beta_0]}{1+\beta_0}\right] - 2r\gamma (1+\beta_0)^2 \sigma^2 \tag{19}$$

$$0 = [(1+\beta_0)w(0,\lambda,0) + \beta_1 - r\beta_0] + \frac{1}{r}\frac{\sigma^2}{s^2}(\beta_1 - r\beta_0) \tag{20}$$

$$+ \frac{1}{r}(1+\beta_0)\left(1 - \frac{\frac{\sigma^2}{s^2}\beta_1[(1+\beta_0)w(0,\lambda,0) + \beta_1 - r\beta_0]}{(1+\beta_0)^2}\right) \tag{21}$$

$$0 = [(1+\beta_0)w(1,\lambda,0) - r\beta_1] - \frac{\sigma^2}{s^2}\beta_1 \tag{22}$$

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

Proof of Lemma 1. See Appendix A.4. □

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 10. *For* $\lambda > 0$, $\alpha_0 \leq 0$ *and* $0 < \beta_1 < r\beta_0$.

Proof of Proposition 10. See Appendix A.4. □

This proposition establishes that when agents form their beliefs by using non-decreasing weights (i.e. $w_0 \geq 0.5$) $\beta_0 r$ 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 $\beta_1 = 0$. In the Appendix, we also show that when agents use increasing weights, i.e. $w_0 < 0.5$, there is a lower bound on the risk-free rate that guarantees that the main result in Proposition 10 holds.

Figure 3 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).

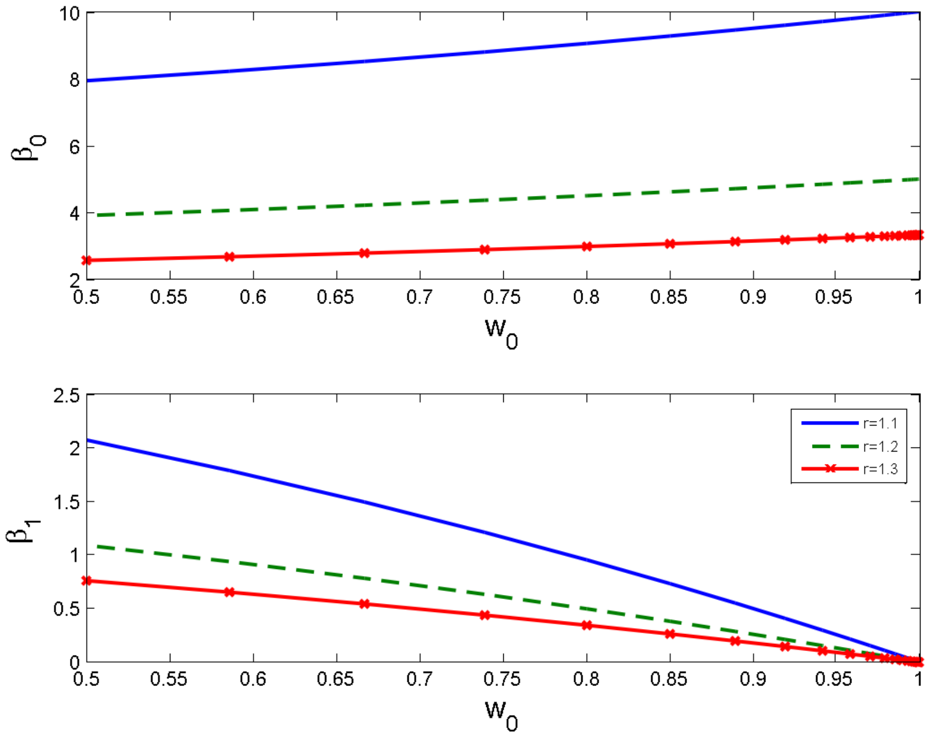

Figure 3: Comparative Statics: Sensitivity of Prices to Dividends for the q=2 Case.

3.1 Characterization of the Demands for the Risky Asset

The next proposition establishes that 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 11. *For $r > \bar{r}$: (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 11. See Appendix A.4. □

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, it is useful to re-write the demand of agents as follows:

$$x_t^{t-1} = \frac{\alpha_0 (1-r)}{\gamma (1+\beta_0)^2 \sigma^2} + \frac{(1+\beta_0)w(0,\lambda,1) + \beta_1 - r\beta_0}{\gamma (1+\beta_0)^2 \sigma^2} d_t + \frac{(1+\beta_0)(1-w(0,\lambda,1)) - r\beta_1}{\gamma (1+\beta_0)^2 \sigma^2} d_{t-1}$$

$$x_t^t = \underbrace{\frac{\alpha_0 (1-r)}{\gamma r (1+\beta_0)^2 \sigma^2} + \frac{1+\beta_0+\beta_1 - r\beta_0}{\gamma r (1+\beta_0)^2 \sigma^2} d_t - \frac{r\beta_1}{\gamma r (1+\beta_0)^2 \sigma^2} d_{t-1}}_{\tilde{x}_t^t} + \Delta_t$$

with

$$\Delta_t \equiv \frac{\alpha_0 (1-r) + (\beta_1 - r\beta_0) d_t - r\beta_1 d_{t-1}}{\gamma r (1+\beta_0)^2} \left(\frac{1}{s^2} - \frac{1}{\sigma^2} \right) + \frac{1}{\gamma r (1+\beta_0)} \left(\frac{m}{s^2} - \frac{d_t}{\sigma^2} \right). \quad (23)$$

In particular, the demand of the young agent can be expressed as a term reflecting the demand of a static agent with risk aversion $r\gamma$ and the beliefs of the young agent (we denote this term by $\tilde{x}_t^t$) and a second term reflecting the adjustment in demand of the risky asset that arises in a learning framework, which we will refer to as a *hedging motive*, Δ_t.

To understand how the demand of young and adult agents react to changes in dividends, we need to find:

$$\frac{\partial (x_t^t - x_t^{t-1})}{\partial d_t} = \frac{\partial (\tilde{x}_t^t - x_t^{t-1})}{\partial d_t} + \frac{\partial \Delta_t}{\partial d_t}. \quad (24)$$

We focus first on understanding the changes in demands when we abstract from the hedging motive. Let

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

The 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

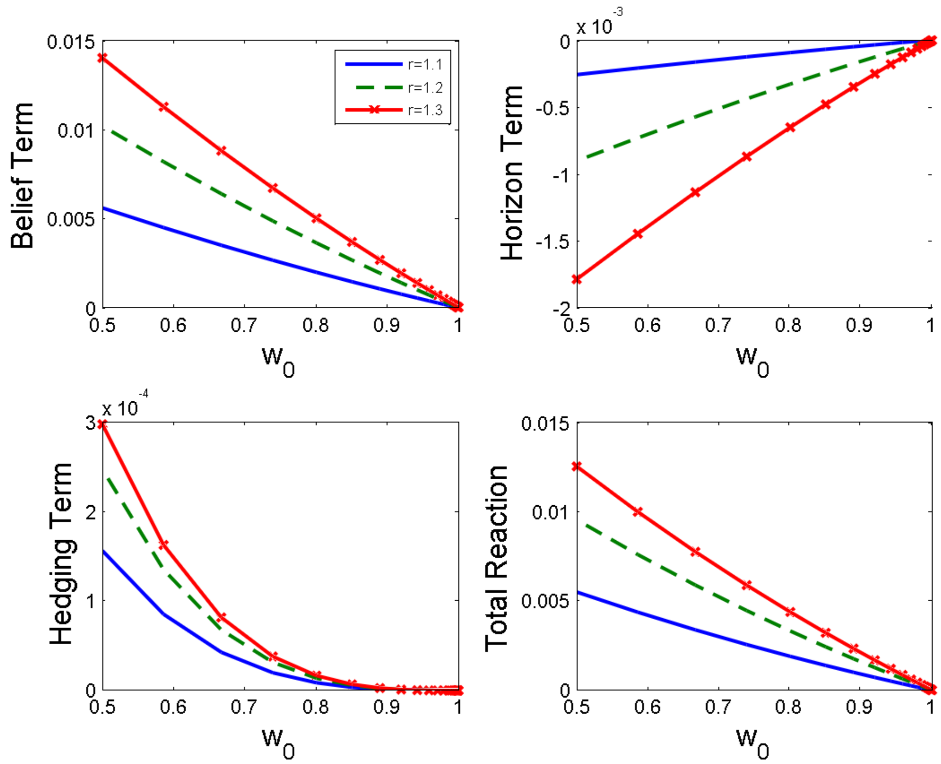

Figure 4: Comparative Statics: Sensitivity of Demands to Dividends for the q=2 Case.

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

the return of the risky asset than adult agents, who are also weighing past realizations of dividends in their belief formation. This term is zero when both agents have the same belief formation (e.g. $w(0, \lambda, 1) = 1$). The second term is the *Horizon Term*. It is negative (see Lemma 8 in Appendix for a proof). It reflects the fact that even when agents share beliefs, young agents react less aggressively to a change in dividends (in their beliefs) due to their longer horizon. These terms fully characterize the differential response across different cohorts when agents *do not* internalize the learning in their portfolio decisions.

When agents understand that they are learning about the risky asset, they distort their portfolio decisions accordingly, giving rise to a *hedging* motive. Note that when $m = \theta_t^t = d_t$, and $s^2 = \sigma^2$; that is, there is no adjustment in the distribution, the hedging motive disappears.

However, for other cases, we are interested in how this term reacts to changes in present dividends. Observe that

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

since $\frac{\partial m}{\partial d_t} = \frac{s^2}{\sigma^2}\left(1 - \frac{l(1,1)l(0,1)}{(1+\beta_0)^2}\right)$. Even though we can not formally pin down the sign of $\frac{\partial \Delta}{\partial d_t}$, since the first term is negative, while the second term is positive (see the proof of Proposition 11, in the numerical solutions we find that it is always positive (see Figure (4)). Most importantly, we are able to show that the overall sign of $\frac{\partial (x_t^t - x_t^{t-1})}{\partial d_t}$ is positive.

In figure 4 we show the behavior of each of the terms for different values of (r, λ).

3.2 Volume of Trade

We now study how learning and disagreements affect the volume of trade observed in the market. In our OLG framework, there is always trade due to the agents' changing horizon. We focus our analysis on the trade levels driven by the presence of learning and disagreements. Therefore, we are interested in the difference in trade levels between our economy with experience-based learners and our economy with full information. Let $x^n_{FI,t}$ be the demand for the risky asset of generation n at time t in the full information economy.[6] We define the trade of generation n at time t that is driven by learning as $tr^n_t = x^n_t - x^n_{FI,t}$. If agents are not learning, and there are no disagreements, our trade measure is zero. In this sense, we are measuring the trade volume that is solely generated by the presence of learning and disagreements and not to our specific OLG framework. We focus on characterizing trade levels for the $q = 2$ case. The resulting total volume of trade in the economy is defined as follows:

[6] Trade levels would be the same in an economy with no learning and no disagreements, even if we were not in the full information case. As long as all agents agree on the distribution of dividends, the demand functions are the same as the ones in the full information case where all agents know the true mean of dividends.

$$TR_t \equiv \sum_{n=t-2}^{t} \frac{1}{2} \left(tr_t^n\right)^2 \tag{25}$$

By market clearing, $tr_t^t + tr_t^{t-1} = 0$ and thus trade volume can be written as: $TR_t = \left(tr_t^n\right)^2$. In the full information economy, the demand of the young generation is given by $x_{FI,t}^t \equiv \frac{E[s_{t+1}]}{\gamma r(1+\beta_0)^2\sigma^2}$, that of the adult generation by $x_{FI,t}^{t-1} \equiv \frac{E[s_{t+1}]}{\gamma(1+\beta_0)^2\sigma^2}$, and that of the old is zero. By market clearing: $x_{FI,t}^t = \frac{2}{1+r}$, $x_{FI,t}^{t-1} = \frac{2r}{1+r}$. The following Lemma characterizes trade volume for this economy.

Lemma 2. *For the $q = 2$ case, trade volume defined by (25) is given by:*

$$TR_t = \frac{1}{1+r} \left(\frac{1-w_0}{\gamma\left(1+\beta_0\right)\sigma^2} \psi_t + r\Delta\left(d_{t-1}, \psi\right) \right)^2 \tag{26}$$

where $\psi = d_t - d_{t-1}$. In addition, trade volume increases in response to large changes in dividends, i.e.

$$\frac{\partial TR_t}{\partial \psi} = \frac{2}{1+r} \times tr_t^t \times \frac{\partial tr_t^t}{\partial \psi} \tag{27}$$

and there exists an interval $\left[\underline{\psi}, \bar{\psi}\right]$ with $\underline{\psi} \leq 0 \leq \bar{\psi}$ such that for $\psi < \underline{\psi}$ and $\psi > \bar{\psi}$, $\frac{\Delta TR_t}{\Delta \psi} > 0$ [It is more general than this, think about how to write it].

Proof. See Appendix A.5. □

The previous Lemma shows that the presence of learning and disagreements induces trade volume through two channels. The first is the belief, or disagreements, channel, which is captured by the first term and is proportional to ψ, the change in dividends. Remember that an increase (decrease) in dividends impacts the belief of both generations in the market, but the effect on beliefs is stronger for the younger generation. Therefore, an increase (decrease) in dividends induces trade since young agents become more optimistic (pessimistic) than old agents, and disagreements generate gains from trade. This mechanism is solely due to the presence ef experience-based learners, since it is essential that each generation reacts differently to the same realization of dividends.

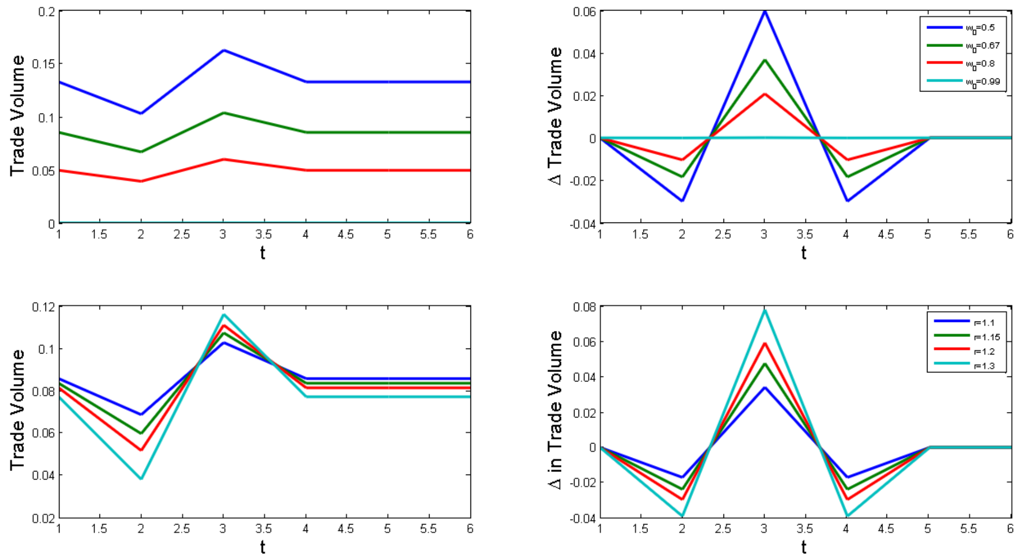

Figure 5: Comparative Statics: Trade Volume Levels and Changes.

We assume that $d_t = 20$ for all t but $d_2 = 10$. That is, $\psi = -10$ in $t = 2$ and zero otherwise. Parameters are as the ones described in other plots.

The second is the learning channel, and it is captured by the second term: the hedging motive. Young cohorts have an incentive to distort their portfolios due to learning (since returns are now correlated with their continuation values through beliefs), while old agents do not have continuation values and thus invest as static agents. This difference in the way they form optimal portfolios also induces trade. The response of this hedging motive to changes in dividends, however, could be positive or negative. What we show, however, is that for any initial level of dividends d_{t-1}, there always exists a large enough change in dividends (positive or negative), that will increase the hedging motive, and thus trade volume.

Figure 5 shows the response of trade levels to a negative shock to dividends. It is clear from the simulations that the change in trade volume induced by the change in dividends is larger when disagreements among cohorts are larger (i.e. when w_0 is smaller). This is because changes in dividends induce more disagreements when the recency bias in agent's belief formation is not that strong. It is also important to highlight that a one-time change in dividends induces traded volume not only at the time of the shock, but also in the future, since there will be disagreements between the generation that experienced the shock and the newly born. This suggests that there is a persistent component to changes in trade volume.

4 Simulations

We solve the model numerically for different parameter values, and simulate the economy for the $q = 2$ case to highlight the main results discussed in the paper. For the following numerical exercise, we assume (unless otherwise noted), that the belief parameter is $\lambda = 1$ (implying a belief weight on most recent dividends of the adult generation of $w_0 = 0.67$), volatility of dividends $\sigma^2 = 1$, risk-free rate $r = 1.1$, risk aversion $\gamma = 10$, and the mean of dividends $\theta = 10$. We simulate the following scenario: dividends are constant at their mean level: $d_t = 20$, but in $t = 2$ there is a one time negative shock implying that $d_2 = 10$.

We study the reaction to a negative shock in dividends on the demands for the risky asset of different cohorts, the price of the risky asset, and trade volume. We do so for several weigh functions in the agent's beliefs formation, and for various interest rates.

5 Conclusion

To be completed.

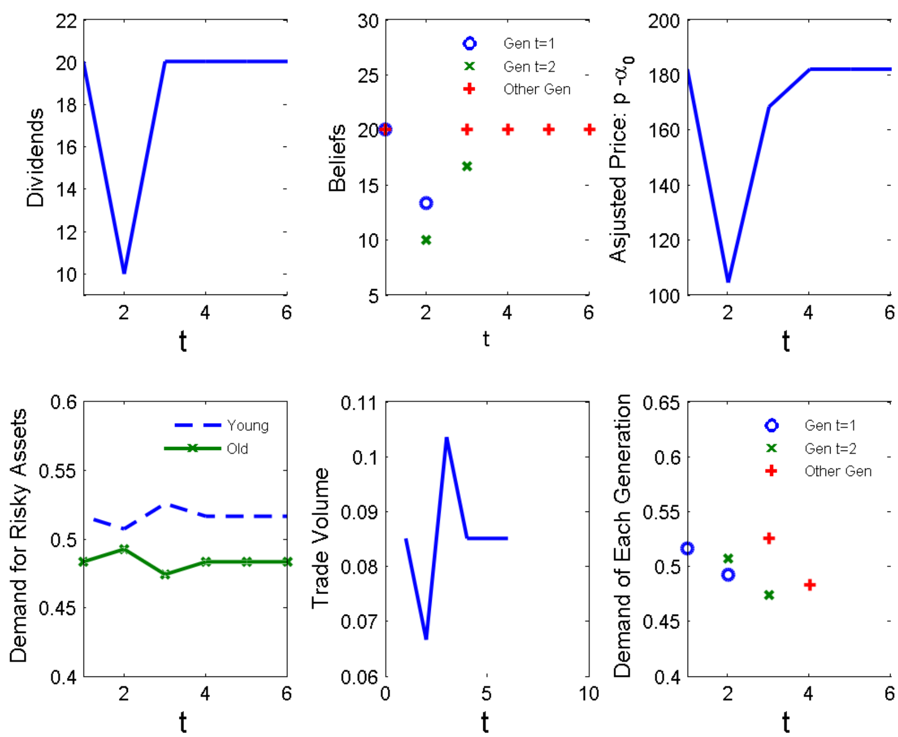

Figure 6: Simulating the q=2 Case.

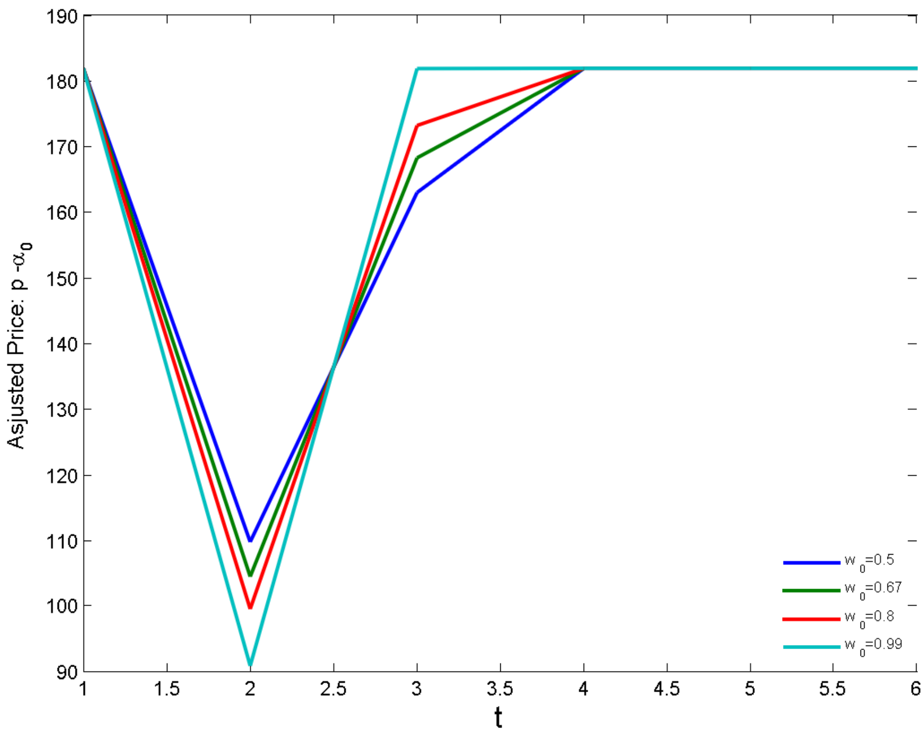

Figure 7: Simulating the q=2 Case.

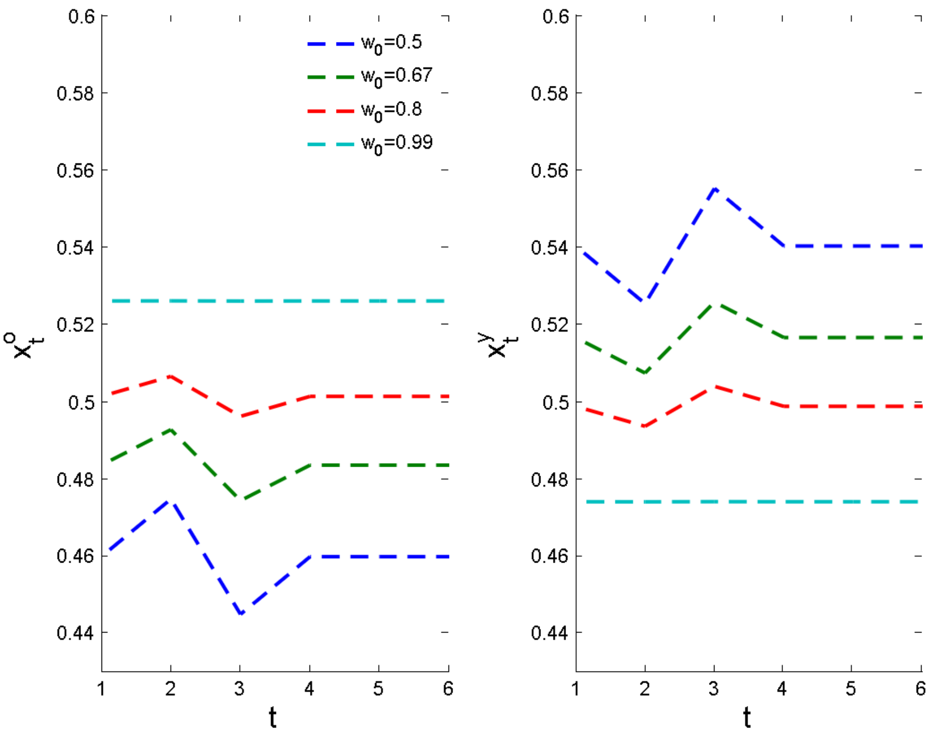

Figure 8: Simulating the q=2 Case.

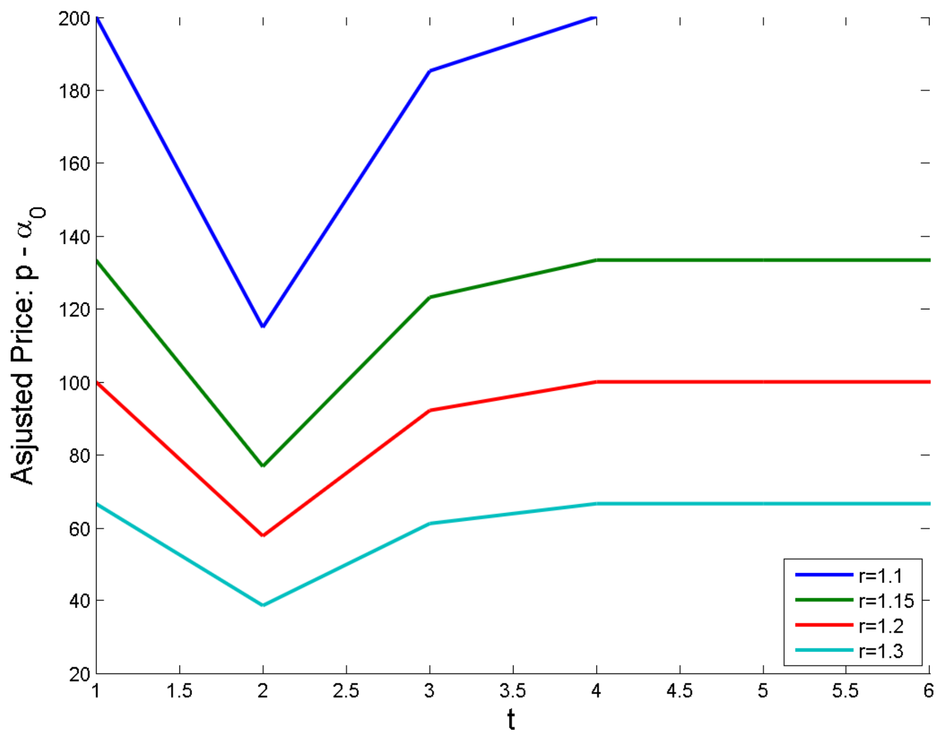

Figure 9: Simulating the q=2 Case.

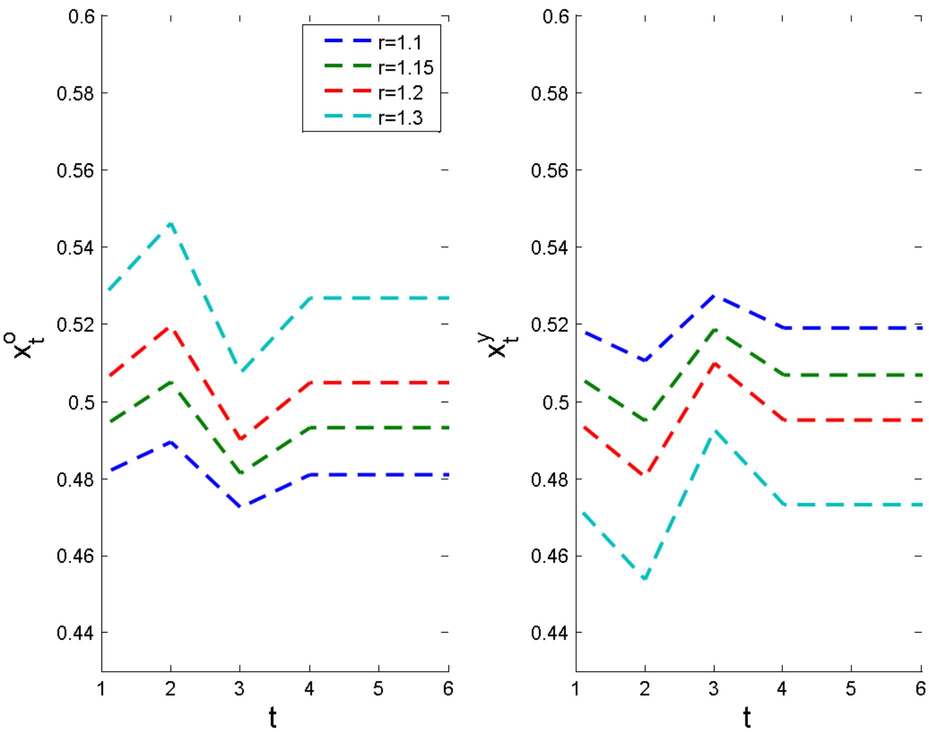

Figure 10: Simulating the q=2 Case.

A Appendix

A.1 Proofs of Section 2.3

For the proof of Proposition 6 we need the following technical lemmas (their proofs are relegated to section A.2)

Lemma 3 (l: AdjustedGaussian)**.** *Suppose $z \sim N(\mu, \sigma^2)$, then for any $A, B, C \in \mathbb{R}$, $z \mapsto K^{-1}\exp\{-A - Bz - Cz^2\}\phi(z; \mu, \sigma^2)$ 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*

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

Lemma 4 (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 $x_t^{t-q+1} = \frac{E_t^{t-q+1}[s_{t+1}]}{\gamma\sigma_*^2}$, $\forall t \geq 0, q \geq 1$.*

Lemma 5 (l: GralMax)**.** *Let $z \sim \Phi(\mu, \sigma^2)$. Let $A, B, C \in \mathbb{R}$, and $z \mapsto h(z) \equiv f + ez$ for any $e, f \in \mathbb{R}$. Then*

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

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

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

Lemma 6 (l: static)**.** *Let $z \sim \Phi(\mu, \sigma^2)$, then for any $a > 0$,*

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

$$\text{and } \max_x E[-\exp\{-axz\}] = -\exp\{-0.5(\sigma a x^*)^2\} = -\exp\left(-0.5\frac{\mu^2}{\sigma^2}\right)$$

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

Lemma 7. *Suppose $p_t = \alpha_0 + \sum_{k=0}^{K} \beta_k d_{t-k}$, 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,*

$$x_t^{t-age} = \delta(age) + \sum_{k=0}^{K} \delta_k(age) d_{t-k}, \ \ for\ age \in \{0, ..., q\} \tag{28}$$

with

$$\delta(q) = \delta_k(q) = 0, \ \ \forall k \in \{0, ..., K\} \tag{29}$$

$$\delta(q-1) = \frac{\alpha_0(1-r)}{\gamma((1+\beta_0)\sigma)^2}, \ \ \delta_k(q-1) = \frac{(1+\beta_0)w(k,\lambda,q-1) + \beta(k)}{\gamma((1+\beta_0)\sigma)^2}, \ \ \forall k \in \{0, ..., q-1\} \tag{30}$$

$$\delta_k(q-1) = \frac{\beta(k)}{\gamma((1+\beta_0)\sigma)^2}, \ \ \forall k \in \{q, ..., K\}, \tag{31}$$

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

$$\delta(age) = \frac{\alpha_0(1-r) - s_{age}^2(1+\beta_0)\delta_0(age+1)\delta(age+1)(r^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2}{r^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2}, \tag{32}$$

$$\delta_k(age) = \frac{(1+\beta_0)s_{age}^2(\sigma^{-2}w(k,\lambda,age) - [(r^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2\delta_{k+1}(age+1)\delta_0(age+1)]) + \beta(k)}{r^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2} \tag{33}$$

$$k \in \{0, ..., q-1\}, \tag{34}$$

$$\delta_k(age) = \frac{-(1+\beta_0)s_{age}^2[(r^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2\delta_{k+1}(age+1)\delta_0(age+1)] + \beta(k)}{r^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2}, \ \ k \in \{q, ..., K-1\} \tag{35}$$

$$\delta_K(age) = \frac{\beta(K)}{r^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2}, \tag{36}$$

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 $j \in \{0, ..., q-1\}$ are:

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

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

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

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

Proof of Proposition 6. By lemma 6,

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

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

$$V^t_{t+q-1} = -\exp\{-0.5\left((1+\beta_0)\sigma\gamma x^t_{t+q-1}\right)^2\}.$$

By lemma 7, x^t_{t+q-1} is affine in $d_{t+q-1-K:t+q-1}$ and thus $V^t_{t+q-1} = -\exp\{-A - Bd_{t+q-1} - C(d_{t+q-1})^2\}$ 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 $d_{t+q-1-K:t+q-1}$:

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

(see Lemma 7 step 1 for the expressions for $\delta(q-1)$ and $(\delta_k(q-1))_{k=1}^{K}$).

At time $t+q-2$, by equation 12,

$$x^t_{t+q-2} = \arg\max_{x\in\mathbb{R}} E^t_{t+q-2}\left[V^t_{t+q-1}(d_{t+q-1-K:t+q-1})\exp\left(-\gamma s_{t+q-1}x\right)\right]$$

where the expectation is taken with respect $N(\theta^t_{t+q-2},\sigma^2)$. Hence, by lemma 3, this problem can be cast as

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

where $m_{q-2} = \sigma_{q-2}(\frac{\theta^t_{t+q-2}}{\sigma^2} - B)$ and $\sigma^2_{q-2} = \frac{\sigma^2}{2C\sigma^2+1}$. Hence, by lemma 6

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

Also, by lemma 6, $V^t_{t+q-2} = -\exp\{-0.5\left(V_{N(m_{q-2},\sigma^2_{q-2})}[s_{t+q-1}]r\gamma x^t_{t+q-2}\right)^2\}$. By lemma 7, x^t_{t+q-2} is affine and thus $V^t_{t+q-2} = -\exp\{-A - Bd_{t+q-2} - C(d_{t+q-2})^2\}$ 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 $d_{t+q-1-K:t+q-1}$:

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

(observe that the A and B and C are not the same as the previous ones; the expressions for $\delta(q-2)$ and $(\delta_k(q-2))^K_{k=1}$ can be found in the proof of lemma 7 step 2).

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

□

A.2 Proof of Lemmas 3, 4, 5, 6 and 7

Proof of Lemma 3. 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 $K = \frac{1}{\sqrt{2\sigma^2 C+1}}\exp\{-(A+0.5\sigma^{-2}\mu^2)+\frac{m^2}{2\Sigma^2}\}$:

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

□

Proof of Lemma 4. 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: $x^t_{t+q} = 0, a^t_{t+q} = 0$. 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_{t+q-1}:

$$\max_x E^t_{t+q-1}\left[-\exp\left(-\gamma\left(W^t_{t+q-1}+xs_{t+q}\right)\right)\right] = \max_x E^t_{t+q-1}\left[-\exp\left(-\gamma x s_{t+q}\right)\right] \tag{37}$$

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

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

$$V^{t-q}_{t-1} \equiv \max_x \left[-\exp\left(-\gamma x E^{t-q}_{t-1}\left[s_t\right] + \frac{1}{2}\gamma^2 x^2 \left(1+\beta_0\right)\sigma^2\right)\right] \tag{38}$$

$$\max_x \quad x E^{t-q}_{t-1}\left[s_t\right] - \frac{1}{2}\gamma x^2 \left(1+\beta_0\right)^2 \sigma^2 \tag{39}$$

And therefore, by FOC:

$$x^t_{t+q-1} = \frac{E^t_{t+q-1}\left[s_{t+q}\right]}{\gamma\sigma^2_*} \tag{40}$$

□

Proof of Lemma 5. Note that $E[-\exp\{-A-Bz-Cz^2\}\exp\{-axh\left(z\right)\}]$ can be written as:

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

By Lemma 3, we know that his can be re-written as:

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

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

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

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

$V_{\Phi(m,s^2)}\left[h\left(z\right)\right]$, by Lemma 6, we know that

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

□

Proof of Lemma 6. Since $z\sim\Phi(\mu,\sigma^2)$, we can re-write the problem as follows:

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

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

□

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

$$\Lambda(d_{t-K},...,d_t)\equiv\alpha_0(1-r)+\sum_{k=1}^{K}\beta_k d_{t+1-k}-r\sum_{k=0}^{K}\beta_k d_{t-k}$$
$$=\alpha_0(1-r)+\sum_{j=0}^{K-1}\beta_{j+1}d_{t-j}-r\sum_{k=0}^{K}\beta_k d_{t-k}=\alpha_0(1-r)+\sum_{k=0}^{K}\beta(k)d_{t-k}$$

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

Proof of Lemma 7. 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 4, 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,

$$\delta(q) = \delta_k(q) = 0, \;\; \forall k \in \{0, ..., K\} \tag{41}$$

$$\delta(q-1) = \frac{\alpha_0(1-r)}{\gamma((1+\beta_0)\sigma)^2}, \;\; \delta_k(q-1) = \frac{(1+\beta_0)w(k,\lambda,q-1) + \beta(k)}{\gamma((1+\beta_0)\sigma)^2}, \;\; \forall k \in \{0, ..., q-1\} \tag{42}$$

$$\delta_k(q-1) = \frac{\beta(k)}{\gamma((1+\beta_0)\sigma)^2}, \;\; \forall k \in \{q, ..., K\}. \tag{43}$$

We also know from Lemma 6 that

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

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

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,

$$\max_x E_t^{t-q+2}\left[V_{t+1}^{t-q+2}\exp\left(-\gamma r x s_{t+1}\right)\right]. \tag{44}$$

By the calculations in step 1, and using Λ_t as defined in (41), this problem becomes:

$$V^{q-2}(d_{t-K}, ..., d_t) \tag{45}$$

$$= \max_x E_t^{t-q+2}\left[-\exp\left(-\frac{1}{2}\left(x_t^{q-1}\right)^2\gamma^2((1+\beta_0)s_{q-1})^2 - \gamma r x((1+\beta_0)d_{t+1} + \Lambda_t)\right)\right]. \tag{46}$$

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

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

and that future dividends are the only random variable, with $d_{t+1}\sim N\left(\theta_t^{t-q+2},\sigma^2\right)$. Therefore, by Lemma 5, and with:

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

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,

$$\begin{aligned}
\delta(q-2) =& \frac{\alpha_0(1-r) - s_{q-2}^2(1+\beta_0)\delta_0(q-1)\delta(q-1)\gamma^2((1+\beta_0)s_{q-1})^2}{r\gamma((1+\beta_0)s_{q-2})^2} \\
\delta_k(q-2) =& \frac{(1+\beta_0)s_{q-2}^2(\sigma^{-2}w(k,\lambda,q-2) - [\gamma^2((1+\beta_0)s_{q-1})^2\delta_{k+1}(q-1)\delta_0(q-1)]) + \beta(k)}{r\gamma((1+\beta_0)s_{q-2})^2}, \; k \in \{0,...,q-1\} \\
\delta_k(q-2) =& \frac{-(1+\beta_0)s_{q-2}^2[\gamma^2((1+\beta_0)s_{q-1})^2\delta_{k+1}(q-1)\delta_0(q-1)] + \beta(k)}{r\gamma((1+\beta_0)s_{q-2})^2}, \; k \in \{q,...,K-1\} \\
\delta_K(q-2) =& \frac{\beta(K)}{r\gamma((1+\beta_0)s_{q-2})^2}.
\end{aligned}$$

By lemma 3, $d_{t+1} \sim N(m_t, s_{q-2}^2)$ with $m_t \equiv -s_{q-2}^2 B + s_{q-2}^2\sigma^{-2}\theta_t^{t-q+2}$. Thus, invoking lemma 6 for this distribution for dividends and $a = r\gamma(1+\beta_0)$ implies that

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

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

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

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

The maximization problem is given by:

$$V^{age}(d_{t-K},...,d_t) \equiv \max_x E_t^{t-age}\left[V_{t+1}^{t-age-1}\exp\left(-\gamma r^{q-1-age}x((1+\beta_0)d_{t+1} + \Lambda_t)\right)\right]. \quad (47)$$

By similar calculations to step 2 and Lemma 5,

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

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

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

Therefore

$$\begin{aligned}
\delta(age) =& \frac{\alpha_0(1-r) - s_{age}^2(1+\beta_0)\delta_0(age+1)\delta(age+1)(r^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2}{r^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2}, \\
\delta_k(age) =& \frac{(1+\beta_0)s_{age}^2(\sigma^{-2}w(k,\lambda,age) - [(r^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2\delta_{k+1}(age+1)\delta_0(age+1)]) + \beta(k)}{r^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2} \\
& k \in \{0, ..., q-1\}, \\
\delta_k(age) =& \frac{-(1+\beta_0)s_{age}^2[(r^{q-1-(age+1)}\gamma)^2((1+\beta_0)s_{age+1})^2\delta_{k+1}(age+1)\delta_0(age+1)] + \beta(k)}{r^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2}, \ k \in \{q, ..., K-1\} \\
\delta_K(age) =& \frac{\beta(K)}{r^{q-1-(age)}\gamma((1+\beta_0)s_{age})^2}.
\end{aligned}$$

By lemma 3, $d_{t+1} \sim N(m_t, s_{age}^2)$ with $m_t \equiv -s_{age}^2 B + s_{age}^2\sigma^{-2}\theta_t^{t-q+2}$. Thus, invoking lemma 6 for this distribution for dividends and $a = r^{q-1-age}\gamma(1+\beta_0)$ implies that

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

□

A.3 Proof for Section 2.4

Proof of Proposition 9. [HERE AND BEFORE, $\beta_0 \neq -1$] Market Clearing and Lemma 7 imply that, for all $k \in \{0, ..., K\}$,

$$\sum_{age=0}^{q-1} \delta_k(age) = 0 \tag{48}$$

and

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

For $k = K$, it follows from equations 31 and 36

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

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 31 and 35

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

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

By induction, for any $k \in \{q, ..., K-2\}$, taking $\beta_{k+1} = 0$, it follows by equations 31 and 35, 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 $age \in \{q, ..., K\}$.

For $k = q-1$, it follows by equations 30 and 33

$$\begin{aligned}
\sum_{age=0}^{q-1} \delta_{q-1}(age) =& \frac{(1+\beta_0)w(q-1,\lambda,q-1) - r\beta_{q-1}}{\gamma((1+\beta_0)\sigma)^2} \\
&+ \sum_{age=0}^{q-2} \frac{(1+\beta_0)s_{age}^2\sigma^{-2}w(q-1,\lambda,age) - r\beta_{q-1}}{r^{q-1-age}\gamma((1+\beta_0)s_{age})^2} \\
=& \frac{(1+\beta_0)w(q-1,\lambda,q-1) - r\beta_{q-1}}{\gamma((1+\beta_0)\sigma)^2} - \sum_{age=0}^{q-2} \frac{r\beta_{q-1}}{r^{q-1-age}\gamma((1+\beta_0)s_{age})^2} \\
=& \frac{1}{\gamma(1+\beta_0)^2}\left(\frac{(1+\beta_0)w(q-1,\lambda,q-1)}{\sigma^2} - r\beta_{q-1}\sum_{age=0}^{q-1}\frac{1}{r^{q-1-age}(s_{age})^2}\right)
\end{aligned}$$

where the second line follows from the fact that $w(q-1,\lambda,age) = 0$ for $age \in [0,...q-2]$. Thus

$$\beta_{q-1} = \frac{(1+\beta_0)w(q-1,\lambda,q-1)}{r\sigma^2\Gamma},$$

with $\Gamma = \sum_{age=0}^{q-1} \frac{1}{r^{q-1-age}(s_{age})^2}$.

For any $k \in \{0,...,q-2\}$, by equations 30 and 33

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

and finally, by equations 30 and 32

$$\sum_{age=0}^{q-2} \delta(age) + \delta(q-1) = q^{-1}$$

□

A.4 Proofs for Section 3

A.4.1 Proof of Lemma 1

Proof of Lemma 1. By Proposition 6, we have the following demands:

$$x_t^{t-2} = 0 \tag{49}$$

$$x_t^{t-1} = \frac{E_t^{t-1}\left[s_{t+1}\right]}{\gamma r\left(1+\beta_0\right)\sigma^2} = \frac{\alpha_0\left(1-r\right)+l\left(0,1\right)d_t+l\left(1,1\right)d_{t-1}}{\gamma\left(1+\beta_0\right)^2\sigma^2} \tag{50}$$

$$x_t^{t} = \frac{E_{\Phi(m,s^2)}\left[s_{t+1}\right]}{\gamma r\left(1+\beta_0\right)s^2} = \frac{\alpha_0\left(1-r\right)+\left(\beta_1-r\beta_0\right)d_t-r\beta_1 d_{t-1}+\left(1+\beta_0\right)m}{\gamma r\left(1+\beta_0\right)^2 s^2} \tag{51}$$

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

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

and

$$B_{t+1}\left(1\right) = \frac{\alpha_0\left(1-r\right)l\left(0,1\right)}{\left(1+\beta_0\right)^2\sigma^2}+\frac{l(1,1)l\left(0,1\right)}{\left(1+\beta_0\right)^2\sigma^2}d_t$$
$$C\left(1\right) = \frac{l\left(0,1\right)^2}{\left(1+\beta_0\right)^2\sigma^2}$$

Therefore:

$$m = \frac{s^2}{\sigma^2}\left[d_t-\frac{\alpha_0\left(1-r\right)l\left(0,1\right)}{\left(1+\beta_0\right)^2}-\frac{l(1,1)l\left(0,1\right)}{\left(1+\beta_0\right)^2}d_t\right] = \frac{s^2}{\sigma^2}\left[-\frac{\alpha_0\left(1-r\right)l\left(0,1\right)}{\left(1+\beta_0\right)^2}+\left(1-\frac{l(1,1)l\left(0,1\right)}{\left(1+\beta_0\right)^2}\right)d_t\right]$$
$$s^2 = \frac{\sigma^2}{2\frac{l(0,1)^2}{(1+\beta_0)^2\sigma^2}\sigma^2+1} = \frac{\left(1+\beta_0\right)^2}{l\left(0,1\right)^2+\left(1+\beta_0\right)^2}\sigma^2.$$

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

$$\begin{aligned}
x_t^t =& \frac{\alpha_0 (1-r) + (\beta_1 - r\beta_0) d_t - r\beta_1 d_{t-1} + (1+\beta_0) \frac{s^2}{\sigma^2} \left[-\frac{\alpha_0 (1-r) l(0,1)}{(1+\beta_0)^2} + \left(1 - \frac{l(1,1) l(0,1)}{(1+\beta_0)^2}\right) d_t \right]}{\gamma r (1+\beta_0)^2 s^2} \\
=& \frac{\alpha_0 (1-r) \left[1 - \frac{s^2}{\sigma^2} \frac{l(0,1)}{(1+\beta_0)}\right] + \left[\beta_1 - r\beta_0 + (1+\beta_0) \frac{s^2}{\sigma^2} \left(1 - \frac{l(1,1) l(0,1)}{(1+\beta_0)^2}\right)\right] d_t - r\beta_1 d_{t-1}}{\gamma r (1+\beta_0)^2 s^2}.
\end{aligned}$$

By Market clearing:

$$\begin{aligned}
1 =& \frac{1}{2} \left(\frac{\alpha_0 (1-r) + l(0,1) d_t + l(1,1) d_{t-1}}{\gamma (1+\beta_0)^2 \sigma^2} \right) \\
&+ \frac{1}{2} \left(\frac{\alpha_0 (1-r) \left[1 - \frac{s^2}{\sigma^2} \frac{l(0,1)}{(1+\beta_0)}\right] + \left[\beta_1 - r\beta_0 + \frac{s^2}{\sigma^2} (1+\beta_0) \left(1 - \frac{l(1,1) l(0,1)}{(1+\beta_0)^2}\right)\right] d_t - r\beta_1 d_{t-1}}{\gamma r (1+\beta_0)^2 s^2} \right) \\
=& \frac{1}{2} \left(\frac{\alpha_0 (1-r) + l(0,1) d_t + l(1,1) d_{t-1}}{\gamma (1+\beta_0)^2 \sigma^2} \right) \\
&+ \frac{1}{2} \left(\frac{\alpha_0 (1-r) \frac{\sigma^2}{s^2} \left[1 - \frac{s^2}{\sigma^2} \frac{l(0,1)}{(1+\beta_0)}\right] + \left[\frac{\sigma^2}{s^2} (\beta_1 - r\beta_0) + (1+\beta_0) \left(1 - \frac{l(1,1) l(0,1)}{(1+\beta_0)^2}\right)\right] d_t - \frac{\sigma^2}{s^2} r\beta_1 d_{t-1}}{\gamma r (1+\beta_0)^2 \sigma^2} \right),
\end{aligned}$$

which implies

$$\begin{aligned}
2\gamma (1+\beta_0)^2 \sigma^2 =& (\alpha_0 (1-r) + l(0,1) d_t + l(1,1) d_{t-1}) \\
&+ \frac{1}{r} \left[\alpha_0 (1-r) \frac{\sigma^2}{s^2} \left[1 - \frac{s^2}{\sigma^2} \frac{l(0,1)}{(1+\beta_0)} \right] \right] \\
&+ \frac{1}{r} \left[\left[\frac{\sigma^2}{s^2} (\beta_1 - r\beta_0) + (1+\beta_0) \left(1 - \frac{l(1,1) l(0,1)}{(1+\beta_0)^2} \right) \right] d_t - \frac{\sigma^2}{s^2} r\beta_1 d_{t-1} \right] \\
=& \alpha_0 (1-r) \frac{1}{r} \left[r + \frac{\sigma^2}{s^2} - \frac{l(0,1)}{(1+\beta_0)} \right] \\
&+ \left[l(0,1) + \frac{1}{r} \frac{\sigma^2}{s^2} (\beta_1 - r\beta_0) + \frac{1}{r} (1+\beta_0) \left(1 - \frac{l(1,1) l(0,1)}{(1+\beta_0)^2} \right) \right] d_t + \left[l(1,1) - \frac{\sigma^2}{s^2} \beta_1 \right] d_{t-1}.
\end{aligned}$$

Therefore $\{\alpha_0, \beta_0, \beta_1\}$ solve the following system of equations:

$$0 = \alpha_0 (1-r)\left[r + \frac{\sigma^2}{s^2} - \frac{l(0,1)}{1+\beta_0}\right] - 2r\gamma (1+\beta_0)^2 \sigma^2 \tag{52}$$

$$0 = l(0,1) + \frac{1}{r}\frac{\sigma^2}{s^2}(\beta_1 - r\beta_0) + \frac{1}{r}(1+\beta_0)\left(1 - \frac{l(1,1)l(0,1)}{(1+\beta_0)^2}\right) \tag{53}$$

$$0 = l(1,1) - \frac{\sigma^2}{s^2}\beta_1 \tag{54}$$

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

A.4.2 Proof of Proposition 10

Proof of Proposition 10. Throughout the proof, let $w_0 \equiv w(0,\lambda,0)$.

We know from Lemma 1 that $\{\alpha_0, \beta_0, \beta_1\}$ solve the system of equations given by (21) and (22) and 19.

Step 1. By equation 19,

$$2r\gamma (1+\beta_0)^2 \sigma^2 = \alpha_0 (1-r)\left[r + \frac{\sigma^2}{s^2} - \frac{[(1+\beta_0)w(0,\lambda,0) + \beta_1 - r\beta_0]}{1+\beta_0}\right].$$

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 $\alpha_0 \leq 0$.

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

For $1+\beta_0 > 0$, equation (22) 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 (21) to hold it must be that $1 - \frac{1}{r}\frac{l(1,1)}{1+\beta_0} < 0$.

$$1 - \frac{1}{r}\frac{l(1,1)}{(1+\beta_0)^2} = 1 - \frac{1}{r}\frac{(1+\beta_0)(1-w_0) - r\beta_1}{(1+\beta_0)} \tag{55}$$

$$= 1 - \frac{1}{r}(1-w_0) + \frac{\beta_1}{1+\beta_0} > 0 \tag{56}$$

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

STEP 3. We now show that $1+\beta_0>0$. Let $\phi \equiv \frac{\sigma^2}{s^2} > 1$. From equation (22):

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

We plug this into equation (21) and we obtain:

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

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

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

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

□

A.4.3 Proof of Proposition 11

In order to show Proposition 11, we need the following Lemmas (their proofs are relegated to the end of the section).

Lemma 8 (l: sign1)**.** *For* $\lambda \geq 0, 1+\beta_0+\beta_1-r\beta_0>0$.

Lemma 9 (lem:risk-demands-q2)**.** *Given our linear guess for prices (8), when* $q=2$*, at time* t*:*

$$x_t^{t-1}=\frac{E_t^{t-1}\left[s_{t+1}\right]}{\gamma r\left(1+\beta_0\right)\sigma^2}=\frac{\alpha_0\left(1-r\right)}{\gamma\left(1+\beta_0\right)^2\sigma^2}+\frac{l\left(0,1\right)}{\gamma\left(1+\beta_0\right)^2\sigma^2}d_t+\frac{l\left(1,1\right)}{\gamma\left(1+\beta_0\right)^2\sigma^2}d_{t-1} \tag{57}$$

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

with $l(0,1)\equiv[(1+\beta_0)w(0,\lambda,0)+\beta_1-r\beta_0]$ *and* $l(1,1)\equiv[(1+\beta_0)w(1,\lambda,0)-r\beta_1]$*, and* $\delta(0)=\frac{\alpha_0(1-r)\left[1-\frac{s^2}{\sigma^2}\frac{l(0,1)}{(1+\beta_0)}\right]}{\gamma r(1+\beta_0)^2s^2}$*,* $\delta_0(0)=\frac{\beta_1-r\beta_0+(1+\beta_0)\frac{s^2}{\sigma^2}\left(1-\frac{l(1,1)l(0,1)}{(1+\beta_0)^2}\right)}{\gamma r(1+\beta_0)^2s^2}$*, and* $\delta_1(0)=-\frac{r\beta_1}{r\gamma(1+\beta_0)s^2}$.

Proof of Proposition 11. By lemma 9 and Market Clearing, it follows that

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

and

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

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 10, $\beta_1>0$ 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 (21) 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 8, $1 + \beta_0 + \beta_1 - r\beta_0 > 0$. 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 10), $b = r\,(1+\beta_0) - l(1,1) = r\,(1+\beta_0) - (1+\beta_0)w(1,\lambda,0) + r\beta_1 > 0$ (by Proposition 10) and $c = 1 + \beta_0 + \beta_1 - r\beta_0 > 0$ (by Lemma 8). 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:

$$-\frac{b}{2a} < \frac{l\,(0,1)}{1+\beta_0} \tag{59}$$

$$\frac{r\,(1+\beta_0) - l(1,1)}{2\,[-(\beta_1 - r\beta_0)]} < \frac{l\,(0,1)}{1+\beta_0} \tag{60}$$

$$\frac{r\,(1+\beta_0) - l(1,1)}{2} < l\,(0,1)\,\frac{r\beta_0 - \beta_1}{1+\beta_0} \tag{61}$$

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

$$\begin{aligned}
r\left(1+\beta_0\right)-\left(1+\beta_0\right)\left(1-w_0\right)+r\beta_1 <& 2l\left(0,1\right)Z\\
r\left(1+\beta_0\right)-\left(1+\beta_0\right)\left(1-w_0\right)+r\beta_1 <& 2Z\left[\left(1+\beta_0\right)w_0+\beta_1-r\beta_0\right]\\
r-1+w_0+r\frac{\beta_1}{\left(1+\beta_0\right)} <& 2Z\left[w_0+\frac{\beta_1-r\beta_0}{\left(1+\beta_0\right)}\right]\\
Z\left(w_0-Z\right) >& 0.5w_0+\frac{1}{2}\left[r-1+r\frac{\beta_1}{1+\beta_0}\right]\\
\frac{w_0}{4} >& 0.5w_0+\frac{1}{2}\left[r-1+r\frac{\beta_1}{1+\beta_0}\right].
\end{aligned}$$

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.

□

A.4.4 Proofs for Supplementary Lemmas

Proof of Lemma 8. 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 (21) we have:

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

Then, since $\beta_1 - r\beta_0 \leq 0$ by proposition 10, for the previous equation to hold it must be that $\left[r-\frac{l(1,1)}{(1+\beta_0)}\right] \leq 0$.

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

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

Proof of Lemma 9. From Lemma 6, 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 9, we have:

$$x_t^{t-1} = \frac{E_t^{t-1}[d_{t+1} + p_{t+1} - rp_t]}{\gamma(1+\beta_0)\sigma^2} \tag{62}$$

$$= \frac{(1+\beta_0)\theta_t^{t-1} + \alpha_0(1-r) + (\beta_1 - r\beta_0)d_t - r\beta_1 d_{t-1}}{\gamma(1+\beta_0)\sigma^2} \tag{63}$$

since $\theta_t^{t-1} = w_0 d_t + (1-w_0)d_{t-1}$, we obtain equation (57), where $l(0,1) = (1+\beta_0)w_0 + \beta_1 - r\beta_0$ and $l(1,1) = (1+\beta_0)(1-w_0) - r\beta_1$. We also know from Lemma 4 that

$$V_t^{t-1} = -\exp\left(-\frac{1}{2}\frac{E_t^{t-q+1}[s_{t+1}]^2}{\gamma(1+\beta_0)\sigma^2}\right)$$
$$= -\exp\left(-\frac{1}{2}\frac{(\alpha_0(1-r) + l(1,1)d_{t-1} + l(0,1)d_t)^2}{\gamma(1+\beta_0)\sigma^2}\right)$$
$$= -\exp\left(-\frac{1}{2}\frac{(L_t(1,1) + l(0,1)d_t)^2}{\gamma(1+\beta_0)\sigma^2}\right)$$

where $L_t(1,1) \equiv \alpha_0(1-r) + l(1,1)d_{t-1}$. Thus, we can write the value function of the generation who is investing for the last time on the market as follows:

$$V_t^{t-1} = -\exp(-A_t - B_t d_t - C d_t^2) \tag{64}$$

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:

$$\max_x \quad E_t^t\left[V_{t+1}^t \exp\left(-\gamma r x s_{t+1}\right)\right] \tag{65}$$

From Lemma 5:

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

Where,

$$\tilde{\mu}\left(m,s^2\right) = E_{\Phi(m,s^2)}\left[h\left(z\right)\right] = \alpha_0(1-r) + (\beta_1 - r\beta_0)d_t - r\beta_1 d_{t-1} + (1+\beta_0)m$$
$$\tilde{\sigma}\left(m,s^2\right)^2 = V_{\Phi(m,s^2)}\left[h\left(d_{t+1}\right)\right] = (1+\beta_0)^2 s^2$$

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 $B_{t+1} = \frac{(\alpha_0(r-1)+l(1,1)d_t)l(0,1)}{(1+\beta_0)^2\sigma^2}$ and $\theta_t^t = d_t$ we obtain equation (58) and the respective δs.

□

A.5 Proofs of Section 3.2

Proof of Lemma 2. 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:

$$x_t^t = \frac{\alpha_0\left(1-r\right)}{\gamma r\left(1+\beta_0\right)^2\sigma^2} + \frac{1+\beta_0+\beta_1-r\beta_0-r\beta_1}{\gamma r\left(1+\beta_0\right)^2\sigma^2}d_{t-1} + \frac{1+\beta_0+\beta_1-r\beta_0}{\gamma r\left(1+\beta_0\right)^2\sigma^2}\psi + \Delta\left(d_{t-1},\psi\right) \tag{66}$$

$$x_t^{t-1} = \frac{\alpha_0\left(1-r\right)}{\gamma\left(1+\beta_0\right)^2\sigma^2} + \frac{1+\beta_0+\beta_1-r\beta_0-r\beta_1}{\gamma\left(1+\beta_0\right)^2\sigma^2}d_{t-1} + \frac{(1+\beta_0)w_0+\beta_1-r\beta_0}{\gamma\left(1+\beta_0\right)^2\sigma^2}\psi \tag{67}$$

Let $x \equiv \frac{\alpha_0(1-r)}{\gamma(1+\beta_0)^2\sigma^2} + \frac{1+\beta_0+\beta_1-r\beta_0-r\beta_1}{\gamma(1+\beta_0)^2\sigma^2}d_{t-1}$. Then from Market Clearing :$x_t^{t-1} + x_t^t = 2$:

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

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

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

The formula for trade volume follows. Now we're interested in understanding $\frac{\partial TR_t}{\partial\psi}$ =

$\left(\frac{1}{1+r}\right)^2 \times tr_t^t \times \frac{\partial tr_t}{\partial \psi} \propto tr_t^t \times \frac{\partial tr_t^t}{\partial \psi}$ near $\psi = 0$.

$$\Delta\left(d_{t-1}, \psi\right) = \frac{1}{\gamma r\left(1+\beta_0\right)^2 \sigma^2} \times \left[\alpha_0\left(1-r\right)\left(\frac{l\left(0,1\right)^2}{\left(1+\beta_0\right)^2} - \frac{l\left(0,1\right)}{\left(1+\beta_0\right)}\right) + \left(\left(\beta_1 - r\beta_0 - r\beta_1\right)\frac{l\left(0,1\right)^2}{\left(1+\beta_0\right)^2} - l(1,1)l\left(0,1\right)\right)d_{t-1} - l(1,1)l\left(0,1\right)\psi\right]$$

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 $tr_t^t\left(d_{t-1}, 0\right) \leq 0$. 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 $tr_t^t\left(d_{t-1}, \bar{\psi}\right) = 0$ and for all $\psi \geq \bar{\psi}$,$tr_t^t\left(d_{t-1}, \psi\right) \geq 0$. 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 $tr_t^t\left(d_{t-1}, 0\right) \geq 0$. By the same argument, there exists $\underline{\psi} \leq 0$ such that $tr_t^t\left(d_{t-1}, \bar{\psi}\right) = 0$ and thus for all $\psi \leq \underline{\psi}$,$tr_t^t\left(d_{t-1}, \psi\right) \leq 0$, 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.

□

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 Communismon People¡¯s Preferences. *American Economic Review 97*, 1507šC1528.

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.

Friedman, M. and A. J. Schwartz (1963). *A Monetary History of the United States.* Princeton, NJ: Princeton University Press.

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*, 534šC539.

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.*

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*, 1151šC1164.