NBER WORKING PAPER SERIES

HETEROGENEITY AND ASSET PRICES:
A DIFFERENT APPROACH

Nicolae B. Gârleanu
Stavros Panageas

Working Paper 26607
http://www.nber.org/papers/w26607

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2020

We are grateful for comments and suggestions from Valentin Haddad, Jon Ingersoll, Ralph Koijen, Lawrence Schmidt, Pietro Veronesi, Motohiro Yogo, as well as participants at WFA, NBER Summer Institute on Asset Pricing, Princeton, Insead, NYU Stern, Berkeley Haas-Stanford GSB joint seminar, UCLA Macro-Finance lunch. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Heterogeneity and Asset Prices: A Different Approach
Nicolae B. Gârleanu and Stavros Panageas
NBER Working Paper No. 26607
January 2020
JEL No. E21,G12

ABSTRACT

We develop a tractable asset-pricing framework characterized by imperfect risk sharing among cohorts, who experience different levels of integrated life-time endowments. While all asset-pricing implications stem from the heterogeneity of consumption among investors, cross-sectional measures of inequality are non-volatile, only weakly related to asset prices, and far more persistent than the price-to-dividend ratio. We show how to identify a marginal agent's consumption growth in this framework by utilizing cross-sectional information. Our proposed notion of marginal-agent consumption growth exhibits different and more volatile low-frequency variation than the aggregate consumption growth per capita, which is normally used in representative agent models. These low frequency movements in our measure of marginal agent consumption growth can explain a large portion of the low frequency movements in real interest rates and, when combined with recursive preferences, can account quantitatively for the stylized asset-pricing facts (high market price of risk, equity premium, volatility, and return predictability).

Nicolae B. Gârleanu
Haas School of Business
F628
University of California, Berkeley
Berkeley, CA 94720
and NBER
garleanu@haas.berkeley.edu

Stavros Panageas
Anderson School of Management
University of California, Los Angeles
110 Westwood Plaza
Los Angeles, CA 90095-1481
and NBER
stavros.panageas@anderson.ucla.edu

1 Introduction

We construct equilibria of continuous-time overlapping-generations (OLG) economies in which different cohorts receive different endowments and experience different standards of living over the course of their lives. Imperfect risk sharing across cohorts implies volatile asset prices and high risk premiums despite constant aggregate consumption and dividend growths. While all asset-pricing implications stem from the heterogeneous consumption and income shocks experienced by investors, changes in the cross-sectional consumption and wealth distributions are smooth, exhibit weak correlation with asset-market fluctuations at high frequencies, and are substantially more persistent than these fluctuations, consistent with the data.

Second, we develop a methodology to infer the consumption growth of the marginal agent by utilizing cross-sectional data. It is important to realize that, due to imperfect risk sharing, the consumption growth of agents who are marginal for asset pricing differs from the aggregate consumption growth per capita; in fact, the OLG feature implies that the identity of marginal agents changes constantly. The methodology is flexible enough to account in a comprehensive manner for a multitude of factors, such as different cohort sizes, age-dependent life-cycle effects, shifts in the demographic pyramid, and different cohort productivities. We show that, empirically, marginal agents' consumption growth exhibits different and more volatile low frequency movements than the consumption growth per capita, which is the relevant quantity in representative agent models. These low-frequency movements can account for the secular variation in the real interest rate and, in conjunction with recursive preferences, also for the usual stylized asset-pricing facts.

The framework is a continuous-time, OLG economy. Agents arrive continuously, endowed with a claim to either a wage path ("workers") or a dividend stream ("entrepreneurs"). All shocks are exclusively redistributional; they drive the income shares obtained by firms and workers born at different times, while aggregate labor and dividend income grow at the same constant rate. Moreover, to differentiate our results from the literature, which has predominantly focused on models with lack of intra-cohort risk sharing, we assume that intra-cohort risk sharing is perfect, and the shares of labor or dividend income accruing to a given cohort of investors are locally deterministic processes,[1] albeit random over the long run.

In this setup we introduce a) imperfect inter-cohort risk sharing and b) recursive utility with a preference for early resolution of uncertainty. We utilize this framework to perform two exercises, one theoretical and one empirical.

Our theoretical exercise parallels the one in Constantinides and Duffie (1996). Specifically, we show the following "possibility" result: Share processes exist that support a broad class of given stationary processes for both the market price of risk and the price-dividend ratio. Our existence results are not abstract, but constructive; we use them to specify endowment-share processes that lead to closed-form expressions for asset-pricing quantities. The model is therefore highly tractable, despite the underlying heterogeneity.

One of the important differences between Constantinides and Duffie (1996) and our paper pertains to the time-series implications for inequality measures, such as the cross-sectional variance of consumption. Constantinides and Duffie (1996) relies on heterogeneous period-by-period changes in individual consumption-growth dispersion, which lead to period-by-period movements in inequality; by contrast, we rely on dispersion and uncertainty in the life-long, integrated consumption experienced by cohorts born at different times. In our approach inequality measures exhibit quite small volatility on a period-by-period basis and are essentially unrelated with the stock market at high frequencies, but they are quite persistent.

The key distinguishing feature of our model relative to a representative-agent economy is that the aggregate (per capita) consumption growth does not coincide with the consumption growth of the marginal agent. In our OLG economy, marginal-agent consumption growth over a given interval is the consumption growth of any cohort that entered before the beginning of the interval, and consequently does not include the portion of aggregate consumption accruing to newly arriving cohorts. Over short time intervals the discrepancy between marginal and aggregate consumption growths is small, but, due to its persistence, its cumulative effect at long horizons can be arbitrarily large. With recursive preferences

[1] Throughout the paper "locally deterministic" refers to a time-differentiable process. By definition, such a process has no diffusion component, but a possibly stochastic drift process.

the different long-run behavior of the two consumption processes has important asset pricing consequences.

We show how to use cross-sectional data to infer our notion of marginal agent's consumption growth. Essentially, we observe that the Euler equation implies a time, age, and cohort decomposition for log cohort consumption. This decomposition can be estimated from cross-sectional data, and the variation of the time effect corresponds to the consumption growth of a fixed cohort (i.e., marginal-agent consumption growth) inside our model. Since detailed cross-sectional data are available only since the mid-eighties, we also develop an indirect inference approach to combine information contained in estimated cohort and age effects with market clearing to extend the sample farther back in time.

One noteworthy feature of our methodology for inferring marginal agent consumption growth is that it requires relatively few assumptions. It accounts for several features of the data (time-varying population and cohort sizes, age profiles of consumption, etc.) to arrive at a measure of marginal-agent consumption growth that applies not only to our model, but to a wide range of OLG models featuring Euler equations. Moreover, similar to how aggregate consumption growth per capita encapsulates all relevant information for the stochastic discount factor in a representative agent economy (whether it is an endowment or production economy), our measure of marginal agent consumption growth encapsulates all relevant information for asset pricing in an OLG economy, irrespective of how one chooses to model production, the government, redistribution policies, demographics, aging effects, etc.[2] For this reason, our tractable OLG model, which features a rich enough specification of endowment dynamics to reproduce the dynamics of marginal agent consumption growth in the data, is sufficient for asset-pricing conclusions.

We examine the marginal agent consumption growth that results from our empirical exercise and compare it to the aggregate consumption growth per capita. We show that our measure of marginal agent consumption growth exhibits more persistence and predictability than aggregate consumption growth, and a stronger co-movement with the real expected interest rate over medium-run cycles. The reason is that economic forces that are typically

[2]Specifically, any two OLG economies calibrated to match the time series properties of our measure of marginal agent consumption growth have the same asset-pricing implications.

irrelevant for representative-agent asset pricing (e.g., the fact that the cohort of people born in 1940–1950 "did better" than the cohort born in the eighties) become priced sources of risk in our model.

To illustrate the quantitative implications of the model, we calibrate it to reproduce our inferred measure of marginal agent consumption growth. We show that the model produces realistic risk premiums, return predictability, interest rate levels, and volatility. We also show that the cross-sectional consumption variance has negligible volatility but follows a near-unit-root process, consistent with the data.

1.1 Relation to the literature

As mentioned above, the nature of our theoretical exercise is similar to Constantinides and Duffie (1996). We show a similar possibility result, but in a model with imperfect inter-, rather than intra-, cohort risk sharing, which breaks the tight link between cross-sectional variance fluctuations and asset-price movements. Some other differences between Constantinides and Duffie (1996) and our approach are: a) we don't have to take a stance on whether higher cross-sectional moments of the consumption distribution exist;[3] and b) our model is explicitly set in continuous time and can be time-integrated to any frequency, whereas the conclusions of Constantinides and Duffie (1996) are sensitive to the choice of time-interval.[4]

For our theoretical results, we employ a stochastic, endowment version of the Blanchard (1985) model. An advantage of this framework is that it allows us to isolate the notion of imperfect risk sharing across cohorts while sidestepping the technical complications of more conventional OLG models.[5] While we use a Blanchard (1985) model for our theoretical results, our empirical measure of marginal consumption growth and the results of our calibration do not depend on the simplifying assumptions of that framework, as we explain

[3] Toda and Walsh (2015) argues that higher moments of the cross sectional distribution of consumption may fail to exist, leading to erroneous conclusions in the Constantinides and Duffie (1996) model.

[4] Grossman and Shiller (1982) proves that idiosyncratic shocks don't matter for the market price of risk as the decision interval approaches zero in a Brownian-risk setting.

[5] Due to their tractability, in recent years perpetual-youth models have gained popularity in asset pricing. See, e.g., Campbell and Nosbusch (2007), Gârleanu and Panageas (2015), Ehling et al. (2018), Maurer (2017), Gomez (2017), Schneider (2017), or Farmer (2018) among others.

in Section 4.

Our paper also relates to the literature that studies the asset-pricing implications of OLG models,[6] to which we make two contributions. First, our paper highlights how to appropriately measure the consumption growth of a marginal agent — not just in this model but also in a broad family of OLG frameworks. Second, for many of the papers in this literature, the risk-price effects due to the lack of intergenerational risk sharing would vanish if trading was allowed at a high rate, as all demography-related shocks are locally predictable. Just as in Grossman and Shiller (1982), only the risk-free rate would be affected. Our framework helps both formalize this criticism (Section 3.1) and address it by employing recursive preferences (Section 3.3).

There are several papers in this literature that rely on perpetual-youth models with stochastic fluctuations in the labor income or profit share of each cohort, and to which ours relates more closely. Gârleanu et al. (2012) employs an i.i.d. framework, and thus cannot generate time variation in asset-pricing moments. Kogan et al. (2019) builds models of firm investment and heterogeneous rent allocation from technological progress. While present in these models, the lack of inter-generational risk sharing is not their centerpiece. In contrast, we consider a simpler endowment economy that allows a detailed theoretical analysis of the implications of the lack of inter-cohort risk sharing, and a methodology to evaluate the empirical connections between the lack of inter-cohort risk sharing and asset pricing.

Our paper also relates to the literature on long run risks, which was initiated by Bansal and Yaron (2004). The point we make in this paper is complementary to Bansal and Yaron (2004). We show (both theoretically and empirically) that inter-cohort risk sharing imperfections are an additional source of long run risk, largely independent from the long-run risks in aggregate per-capita consumption growth.

Further, by modeling dividend and labor income processes explicitly, we avoid the need for an intertemporal elasticity of substitution above one and we can model explicitly the

[6]Indicative examples of such papers include Constantinides et al. (2002), Gomes and Michaelides (2005), Storesletten et al. (2007), and Piazzesi and Schneider (2009). The literature on demographic shocks to asset prices, which we don't attempt to summarize here, is also (remotely) related to the present paper. Two indicative examples are Abel (2003), Geanakoplos et al. (2004).

source of non-cointegration between the dividends of existing stocks and marginal agent consumption in general equilibrium.

There is a voluminous empirical literature that employs time, age and cohort decompositions in repeated cross sections. We do not attempt to summarize this literature here. We simply mention that a large number of empirical papers (especially in labor economics) also find a significant role for cohort effects on people's incomes, consumption, health, etc.[7] Our contribution is to show how to utilize the information contained in a time, age, and cohort decomposition to reconstruct the consumption evolution of the "marginal agent."

The study of inequality is outside the scope of this paper, except for its covariance with asset prices. While the model could be easily extended to account for income inequality at birth, it cannot account for the increase of inequality over the life cycle, which is due to lack of intra-cohort risk sharing coupled with idiosyncratic income shocks. There is a debate in the literatutre on the asset-pricing implications of this type of inequality.[8] For reasons of conceptual differentiation, we abstract from it, without dismissing its importance or claiming that the two approaches to linking heterogeneity and asset prices are mutually exclusive.

The paper is organized as follows. Section 2 presents the model. Section 3 contains the possibility results. Section 4 develops the empirical implications of the model and uses them to measure the consumption and dividend share variations in the data. Section 5 calibrates the model to match the variation in consumption shares and derives the model's asset pricing applications. Section 6 concludes. All proofs are in Appendix A, while Appendix B provides several of the details of our empirical approach.

[7]This literature is too large to summarize here. An indicative list includes Oyer (2008), Oyer (2006), Kahn (2010), von Wachter and Schwandt (2018), and Oreopoulos et al. (2012) amongst many others.

[8]Indicatively, Krueger and Lustig (2010) points out that conventional modeling approaches to idiosyncratic endowment risk result in cross-sectional consumption dynamics that don't matter for risk pricing. As mentioned earlier, Grossman and Shiller (1982) also reaches the conclusion that incomplete risk sharing amongst existing agents does not matter for risk pricing in a continuous time, brownian setting. Schmidt (2015) and Constantinides and Ghosh (2017) enrich Constantinides and Duffie (1996) with recursive preferences and skewness to improve some of its empirical predictions, but don't consider lack of inter-cohort risk sharing as we do.

2 Model

We present the baseline model in two steps. In a first step we perform the analysis assuming that agents have expected utility, logarithmic preferences. In a second step we extend the analysis to recursive preferences.

2.1 Consumers

Time is continuous. Each agent faces a constant hazard rate of death $\lambda > 0$ throughout her life, so that a fraction λ of the population perishes at each instant. A new cohort of mass λ is born per unit of time, so that the total population remains at $\lambda \int_{-\infty}^{t} e^{-\lambda(t-s)} ds = 1$. Later (Proposition 3), we extend the model to accommodate a time-varying birth rate and a random population size.

Consumers maximize the utility they derive from their stream of consumption. In this section we illustrate our approach in the special case of logarithmic utility, i.e., consumers maximize

$$\mathrm{E}_s \left[\int_s^{\infty} e^{-\rho(t-s)} \log \left(c_{t,s} \right) dt \right], \tag{1}$$

where s is the time of their birth and t is calendar time. In Section 3.3, where we derive the main result, the preferences take the form of recursive utility with unitary intertemporal elasticity of substitution (IES). Consumers have no bequest (or gift) motives for simplicity.

2.2 Endowments

Following a long tradition in asset pricing, we consider an endowment economy. The total endowment of the economy is denoted by Y_t and evolves exogenously according to

$$\frac{\dot{Y}_t}{Y_t} \equiv g, \tag{2}$$

where $g > 0$. We intentionally model the aggregate endowment as a deterministic, constant-growth process in order to isolate the effect of redistribution shocks. Proposition 3 extends the model to allow for time-varying aggregate consumption growth.

This aggregate endowment accrues to the various agents populating the economy as follows. At birth, agents are of two types, to which we refer as "entrepreneurs" and "workers." They only differ with respect to their endowment.

Entrepreneurs join the economy at a rate of $\lambda\varepsilon$ per unit of time and constitute a fraction ε of the population. The entrepreneurs born at time s introduce a new cohort of firms into the market. The firms introduced at time s pay the following aggregate dividends at times $t \geq s$:

$$D_{t,s} = \alpha Y_t \eta_s^d e^{-\int_s^t \eta_u^d du}. \tag{3}$$

The term $\alpha \in (0,1)$ in equation (3) is a constant, while $\eta_t^d \geq 0$ is assumed to follow a non-negative diffusion

$$d\eta_t^d = \mu_t^d dt + \sigma_t^d dB_t, \tag{4}$$

for some processes μ_t^d and σ_t^d that we specify later. In equation (3) we can interpret α as the fraction of output that is paid out as dividends, and $\eta_s^d e^{-\int_s^t \eta_u^d du} \geq 0$ as the fraction of dividends accruing to firms of vintage s, since $\int_{-\infty}^t \eta_s^d e^{-\int_s^t \eta_u^d du} ds = 1$ for any path of η_t^d.[9] Accordingly, aggregating across firms of all vintages gives

$$D_t^A \equiv \int_{-\infty}^t D_{t,s} ds \ = \alpha Y_t \int_{-\infty}^t \eta_s^d e^{-\int_s^t \eta_u^d du} ds = \alpha Y_t. \tag{5}$$

Figure 2.2 illustrates the paths of dividends for firms of different vintages in the simple case in which $\eta_t^d = \eta^d$ is a constant. The figure shows that firms belonging to any given cohort s account for a smaller and smaller fraction of aggregate dividends as time t goes by. This is an empirically motivated feature of the model.

We next turn to workers. The specification of workers' endowments mirrors the one for

[9] Specifically, this statement holds true for paths of η_t^d satisfying $\int_{-\infty}^t \eta_s^d ds = \infty$. For the type of stochastic processes that we consider for η_t^d this property holds almost surely.

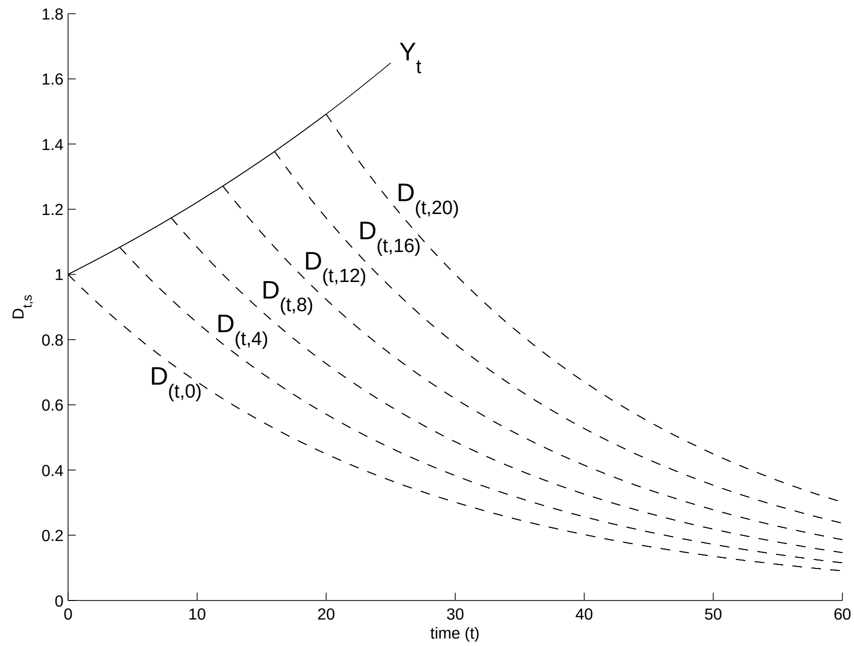

Figure 1: Illustration of $D_{t,s}$ for a constant η_t^d.

dividends and is a simple extension of the specification in Blanchard (1985). Specifically, per unit of time a mass $(1-\varepsilon)\lambda$ of workers is born. Accordingly, the time-t density of surviving workers who were born at time s is given by $l_{t,s} = \lambda\left(1-\varepsilon\right)e^{-\lambda(t-s)}$. The time-$t$ endowment $w_{t,s}$ of a worker who was born at time $s \leq t$ is given by

$$w_{t,s} \equiv \frac{\left(1-\alpha\right)Y_t \eta_s^l e^{-\int_s^t \eta_u^l du}}{l_{t,s}}, \tag{6}$$

where $\eta_t^l \geq 0$ is assumed to follow a non-negative diffusion

$$d\eta_t^l = \mu_t^l dt + \sigma_t^l dB_t,$$

for processes μ_t^l and σ_t^l that we specify later. As with dividend income, the term $\eta_s^l e^{-\int_s^t \eta_u^l du}$ can be interpreted as the share of aggregate earnings that accrues to the cohort of workers born at time s. Repeating the observations we made earlier, aggregate wage earnings are

$$\int_{-\infty}^{t} w_{t,s} l_{t,s} ds = (1-\alpha) Y_t.$$

In a nutshell, the model is a perpetual-youth endowment model with the additional feature that the shares of dividend and labor income accruing to new cohorts are stochastic.

2.3 Markets

Markets are dynamically complete. Investors can trade in instantaneously maturing riskless bonds in zero net supply, which pay an interest rate r_t. Consumers can also trade claims on all existing firms (normalized to unit supply). Following Blanchard (1985), investors can access a market for annuities through competitive insurance companies, allowing them to receive an income stream of $\lambda W_{t,s}$ per unit of time, where $W_{t,s}$ is the consumer's financial wealth. In exchange, the insurance company collects the agent's financial wealth when she dies. Entering such a contract is optimal for all agents, given the absence of bequest motives. A worker's dynamic budget constraint is given by

$$dW_{t,s} = (r_t + \lambda) W_{t,s} dt + (w_{t,s} - c_{t,s}) dt + \theta_{t,s} \left(dP_t + D_t^A dt - r_t P_t dt \right), \tag{7}$$

where P_t is the value of the market portfolio at time t and $\theta_{t,s}$ is the number of shares of the market portfolio. Specification (7) assumes that the consumer trades only in shares of the market portfolio, rather than individual firms. This is without loss of generality in our setup, since all existing firms have identical dividend growth rates, and therefore the same price-to-dividend ratios and consequently the same return to avoid arbitrage.

For a worker, $W_{t,t} = 0$. An entrepreneur's dynamic budget constraint is identical, except that the term $w_{t,s}$ is replaced by zero and the initial wealth $W_{t,t}$ is given by the value of the firm that the entrepreneur creates.

We note that, while agents can replicate any claim once they are alive (markets are "dynamically complete" in that sense), they cannot trade at times prior to their birth. This leads to a lack of inter-cohort risk sharing.

2.4 Equilibrium

The equilibrium definition is standard. We let $c^e_{t,s}$ (resp. $c^w_{t,s}$) denote the time-t consumption of an entrepreneur (resp. worker) born at time s and $\theta^e_{t,s}$ ($\theta^w_{t,s}$) her holding of stock. With $c_{t,s} = \varepsilon c^e_{t,s} + (1-\varepsilon)\, c^w_{t,s}$ the per-capita consumption of cohort s and, similarly, $\theta_{t,s} = \varepsilon \theta^e_{t,s} + (1-\varepsilon)\, \theta^w_{t,s}$, we look for consumption processes, asset allocations $\theta_{t,s}$, asset prices $P_{t,s}$, and an interest rate r_t such that a) consumers maximize objective (1) subject to constraint (7), b) the goods market clears, i.e., $\lambda \int_{-\infty}^{t} e^{-\lambda(t-s)} c_{t,s} = Y_t$, and c) assets markets clear, i.e., $\int_{-\infty}^{t} \lambda e^{-\lambda(t-s)} \theta_{t,s} ds = 1$ and $\int_{-\infty}^{t} \lambda e^{-\lambda(t-s)} (W_{t,s} - \theta_{t,s} P_t) ds = 0$.

3 Solution and Analysis

This section contains our theoretical results, which are in the spirit of the exercise in Constantinides and Duffie (1996). In particular, we establish the existence of share processes η^l_t and η^d_t that can support given processes for the asset pricing quantities as equilibrium outcomes.

The section is divided into four subsections. In Section 3.1 we derive, under the logarithmic-preference assumption, a key relation linking the processes η^d_t and η^l_t to the dynamics of the price-dividend ratio, which we denote by q_t. We use this relation in Section 3.2 to establish the existence of processes η^d_t and η^l_t that can support any given process for q_t as an equilibrium outcome. In Section 3.3 we enrich the setup to allow for recursive preferences and show how to obtain any (joint) dynamics for the price-dividend ratio and the market price of risk (Sharpe ratio) in equilibrium.

Besides providing a comprehensive mapping from assumptions on η^l_t and η^d_t to the equilibrium processes for the price-dividend ratio and the market price of risk, the propositions of this section have a practical implication: They can help determine the functional forms that one needs to assume for the diffusions η^d_t and η^l_t in order to ensure a given (closed-form) expression for the price-dividend ratio and the Sharpe ratio. We illustrate this statement with two examples.

Section 3.4 contains a discussion of the implications of the model for the joint dynamics of inequality and asset prices and highlights the differences of our framework from the literature. It also contains extensions to the baseline model to allow for random birth rates and time-varying per-capita consumption growth.

3.1 Logarithmic utility

We start by conjecturing that in this economy investors' consumption processes are locally deterministic. Given that agents have expected utility preferences, there are no risk premiums and the equilibrium stochastic discount factor m_t follows the dynamics

$$\frac{dm_t}{m_t} = -r_t dt, \tag{8}$$

for an interest rate process that will be determined in equilibrium. We employ the following definition.

Definition 1 *Let $q^d_{t,s}$ denote the ratio of the present value of the dividend stream $D_{u,s}$ to the current dividend:*

$$q^d_{t,s} \equiv \frac{E_t \int_t^\infty \frac{m_u}{m_t} D_{u,s} du}{D_{t,s}}. \tag{9}$$

Similarly, we define the respective valuation ratio for earnings, $q^l_{t,s}$:

$$q^l_{t,s} \equiv \frac{E_t \int_t^\infty e^{-\lambda(u-t)} \frac{m_u}{m_t} w_{u,s} du}{w_{t,s}}. \tag{10}$$

Remark 1 *Both $q^d_{t,s}$ and $q^l_{t,s}$ are independent of s, since $\frac{D_{u,s}}{D_{t,s}}$ and $\frac{w_{u,s} l_{u,s}}{w_{t,s} l_{t,s}}$ are not functions of s. Accordingly, we shall write q^d_t and q^l_t instead of $q^d_{t,s}$, respectively $q^l_{t,s}$.*

Due to unitary IES, the consumption to (total) wealth ratio is constant and equal to

$$\beta \equiv \rho + \lambda.$$

The next lemma uses this fact to derive a simple affine relationship between q_t^d and q_t^l.

Lemma 1 *In any (bubble-free) equilibrium,*

$$\alpha q_t^d + (1-\alpha)\, q_t^l = \frac{1}{\beta}. \tag{11}$$

Equation (11) is intuitive. It states that the sum of the present values of all dividend income accruing to existing firms, $q_t^d \alpha Y_t$, and all earnings accruing to existing agents, $q_t^l (1-\alpha)\, Y_t$, equals the present value of the aggregate consumption of existing agents ($\frac{C_t}{\beta}$). Since $C_t = Y_t$ in equilibrium, equation (11) follows.

By Lemma 1, q_t^l can be expressed as a simple (affine) function of q_t^d. Therefore, from now on we can concentrate our efforts on determining q_t^d, the price-dividend ratio, and we'll simplify notation by writing q_t instead of q_t^d.

In the remainder of this section we determine the equilibrium interest rate r_t and the drift of the price-dividend ratio q_t as functions of the input variables η_t^d and η_t^l. In the next section, we use these results to construct a mapping from the dynamics of q_t to those of η_t^d and η_t^l.

Applying Ito's Lemma to (9) yields the drift of the diffusion process q_t as[10]

$$\mu_{q,t} \equiv \left(r_t - g + \eta_t^d\right) q_t - 1. \tag{12}$$

Equation (12) is an indifference relation between investing in stocks and bonds. After some re-arranging, it states that the expected percentage capital gain on stocks, $\frac{\mu_{q,t}}{q_t}$, plus the dividend yield, $\frac{1}{q_t}$, minus the depreciation (or appreciation) rate $\eta_t^d - g$, should equal the interest rate r_t.

Equation (12) expresses the drift of the price-dividend ratio, $\mu_{q,t}$, in terms of the equilibrium interest rate r_t. To determine this equilibrium interest rate, we proceed in three steps. First, the Euler equation for agents with log preferences implies that the consumption

[10]To see this, note that (9) implies that $m_t q_{t,s}^d D_{t,s} + \int_{-\infty}^t m_u D_{u,s} du$ must be a martingale, and hence the drift of this expression must be zero.

dynamics of any agent are given by[11]

$$\frac{\dot{c}_{t,s}}{c_{t,s}} = -\left(\rho - r_t\right), \tag{13}$$

which also implies that $\frac{\dot{c}_{t,s}}{c_{t,s}}$ is independent of s.

Second, as we show in the appendix, the market clearing condition for aggregate consumption implies that the consumption growth of an existing agent equals

$$\frac{\dot{c}_{t,s}}{c_{t,s}} = g + \lambda - \lambda \frac{c_{t,t}}{Y_t}, \tag{14}$$

which is intuitive: the growth in consumption to an existing agent consists of the growth in aggregate consumption (g), plus the consumption share that perishing agents do not consume (λ), minus the consumption shares accruing to newly born agents ($\lambda c_{t,t}/Y_t$).

Finally, the intertemporal budget constraint at the time of a consumer's birth leads to the following result.

Lemma 2 *Let*

$$\varphi_t \equiv \eta_t^d - \eta_t^l, \tag{15}$$

$$\nu_t \equiv (1 - \alpha\beta q_t)\eta_t^l + \alpha\beta q_t \eta_t^d = \eta_t^l + \alpha\beta q_t \varphi_t. \tag{16}$$

The arriving agents' consumption is given by

$$\lambda \frac{c_{t,t}}{Y_t} = \nu_t. \tag{17}$$

Equation (17) states that the per-capita consumption of an arriving cohort of agents is given by ν_t, which is a weighted average of η_t^l and η_t^d. To derive equation (17), we use the fact that an arriving cohort's initial consumption is the product of the consumption-to-wealth ratio (β) with the sum of the value of the new firms, $\alpha\eta_t^d q_t Y_t$, and the cohort's present value

[11] For a derivation of the Euler equation in our perpetual youth model we refer to Gârleanu and Panageas (2015).

of labor income at birth, $(1-\alpha)\,\eta_t^l q_t^l Y_t$:

$$\lambda c_{t,t} = \beta\left(\alpha \eta_t^d q_t Y_t + (1-\alpha)\,\eta_t^l q_t^l Y_t\right).$$

Dividing both sides of the above equation by Y_t and using Lemma 1 and the definition of ν_t leads to (17).

Combining equations (14), (16), and (17) shows that the consumption drift of any given marginal agent is time varying, even though aggregate consumption growth is constant. The reason is that the existing cohorts, which are marginal in markets, and the arriving cohorts, which are endowed with a random fraction of the endowment at birth, cannot share that endowment risk.

Integrating equation (14) from s to t and using equation (17) implies that the time-t consumption share of the cohort of agents born at time s is given by

$$\frac{\lambda e^{-\lambda(t-s)} c_{t,s}}{Y_t} = \nu_s e^{-\int_s^t \nu_u du}, \tag{18}$$

an expression reminiscent of the specification we used for dividend and earnings shares in equations (3) and (6) respectively. Equation (18) shows that different cohorts of agents experience different integrated consumption growth rates over their lifetimes, a fact that we use in subsection 3.3.

Since any given agent's consumption drift is time-varying, so is the interest rate. Indeed, combining equations (13), (14), and (17) leads to the following result.

Lemma 3 *The equilibrium interest rate is given by*

$$r_t = \beta + g - \eta_t^l - \alpha\beta\varphi_t q_t. \tag{19}$$

Having solved for the equilibrium interest rate, we can now substitute (19) into (12) to obtain the following important result.

Lemma 4 *The drift of q_t is given by*

$$\mu_{q,t} = (\beta + \varphi_t)\, q_t - \beta\alpha\varphi_t q_t^2 - 1. \tag{20}$$

Equation (20) is central for our purposes, since it encapsulates all the equilibrium requirements that our model places on the drift of the price-dividend ratio.

3.2 Supporting a process for q_t as an equilibrium outcome

In this section we ask whether, taking two functions f and σ as given, one can specify a Markovian diffusion for $\varphi_t = \eta_t^d - \eta_t^l$ such that the equilibrium process for q_t is given by

$$dq_t = f(q_t)\, dt + \sigma(q_t)\, dB_t. \tag{21}$$

We leave some technical restrictions on f and σ to ensure that q_t is stationary and takes values in some bounded interval $[q^{\min}, q^{\max}]$ for the appendix, and present here the main argument, followed by an illustrative example and a general proposition.

Any process φ_t that supports (21) as an equilibrium price-dividend ratio must be such that $\mu_{q,t} = f(q_t)$. Using equation (20) allows us to explicitly solve for the process φ_t as a function of q_t:

$$\varphi_t = \varphi(q_t) = \frac{1 - \beta q_t + f(q_t)}{q_t(1 - \beta\alpha q_t)}. \tag{22}$$

We assume that φ thus defined is a strictly decreasing function of q_t, so that its inverse $\varphi^{-1}(\varphi_t)$ exists.[12] Combining equations (21) and (22), the dynamics of the process φ_t are

[12] We note that simple differentiation of (22) shows that φ decreases for $q \leq \frac{1}{2}\frac{1}{\alpha\beta}$ as long as f is decreasing. Hence, choosing the process q_t to have support in $[q^{\min}, q^{\max}]$ with $q^{\max} < \frac{1}{2}\frac{1}{\alpha\beta}$, or choosing a function f that has a sufficiently negative first derivative, is sufficient to ensure that φ is strictly decreasing.

easily obtained from Ito's Lemma as

$$d\varphi_t = \underbrace{\varphi'\left(\varphi^{-1}\left(\varphi_t\right)\right)\left(f\left(\varphi^{-1}\left(\varphi_t\right)\right) + \frac{1}{2}\sigma^2\left(\varphi^{-1}\left(\varphi_t\right)\right)\frac{\varphi''\left(\varphi^{-1}\left(\varphi_t\right)\right)}{\varphi'\left(\varphi^{-1}\left(\varphi_t\right)\right)}\right)}_{\equiv\mu_\varphi(\varphi_t)} dt$$
$$+ \underbrace{\varphi'\left(\varphi^{-1}\left(\varphi_t\right)\right)\sigma\left(\varphi^{-1}\left(\varphi_t\right)\right)}_{\equiv\sigma_\varphi(\varphi_t)} dB_t. \tag{23}$$

Equation (23) provides the answer to the question that we posed at the outset. Specifically, if we started out with the primitive assumption that φ_t follows the Markov diffusion

$$d\varphi_t = \mu_\varphi(\varphi_t)dt + \sigma_\varphi(\varphi_t)dB_t, \tag{24}$$

with $\mu_\varphi(\varphi_t)$ and $\sigma_\varphi(\varphi_t)$ defined in equation (23), then — by construction — the equilibrium dynamics of the price-dividend ratio are given by (21).

Before stating a formal general result, we illustrate the above ideas with a concrete example.

Example 1 *Suppose that we fix a process x_t obeying the following dynamics*

$$dx_t = \left(-v_1 x_t + v_2\left(1 - x_t\right)\right)dt - \sigma_x\sqrt{x_t\left(1 - x_t\right)}dB_t \tag{25}$$

where a_1, a_2, v_1, v_2, and σ_x are positive constants. It is established in the literature (see, e.g., Karlin and Taylor (1981), p. 221) that x_t has a stationary (Beta) distribution with support in $[0, 1]$ as long as $v_1 + v_2 > \frac{\sigma_x^2}{2}$.

Next suppose that we wish the equilibrium price-dividend ratio to be given by $q_t = a_1 + a_2 x_t$. Using (22), the (unique) dynamics of φ_t that support $q_t = a_1 + a_2 x_t$ as an equilibrium outcome are given as an explicit function of the Markov diffusion x_t:

$$\varphi_t = \frac{1 - \beta(a_1 + a_2 x_t) + a_2 v_2 - a_2\left(v_1 + v_2\right)x_t}{(a_1 + a_2 x_t)\left(1 - \beta\alpha(a_1 + a_2 x_t)\right)}. \tag{26}$$

Assuming that the right hand side of the above equation is declining in x_t (which is guaranteed *if $a_1 + a_2 < \frac{1}{2\alpha\beta}$ or if $v_1 + v_2$ is sufficiently large), then φ_t can be expressed as a Markovian diffusion, since it is a monotone function of the Markov process x_t.*[13]

The above example illustrates a practical benefit of our analysis, namely how to guide the choice of a functional form specification for the dynamics of φ_t that leads to a closed form solution for the dynamics of q_t and avoids the need for numerical techniques or approximations. For instance, one could postulate that φ_t follows the dynamics of equation (26), estimate the parameters a_1, a_2, v_1, v_2, and σ_x to match the moments of $\varphi_t = \eta_t^d - \eta_t^l$ in the data, and then examine whether the resulting stochastic process for the price-dividend ratio $q_t = a_1 + a_2 x_t$ is empirically plausible.

The following proposition provides the general result.

Proposition 1 *Suppose that technical Assumption 1 in the appendix is satisfied, and that the function $\varphi(\cdot)$ in equation (22) is decreasing. Then the equilibrium stochastic process for q_t is given by (21) if, and only if, φ_t follows the (Markovian) dynamics (23). Moreover, q_t is stationary and takes values in an interval $[q^{min}, q^{max}]$.*

Proposition 1 states that one can support a wide range of diffusions for q_t as an equilibrium outcome, even though aggregate consumption and dividends are both deterministic. A technical assumption is offered in the appendix to ensure a stationary distribution for q_t.

We conclude with two remarks. First, the process φ_t that supports a given equilibrium stochastic process for q_t is unique. Second, the process q_t only determines $\varphi_t = \eta_t^d - \eta_t^l$. The individual processes η_t^d and η_t^l can be chosen freely as long as their difference obeys the dynamics (23) and the processes are non-negative. (For instance, one choice is to set $\eta_t^l = \eta^l = \varphi(q^{\max})$ and $\eta_t^d = \eta^l + \varphi_t$, which ensures that both processes are non-negative.)

[13] The price-dividend ratio (i.e., the inverse function $q_t = \varphi^{-1}(\varphi_t)$) can be computed explicitly as

$$\varphi^{-1}(\varphi_t) = \frac{1}{2\beta\alpha}\left(\frac{\varphi_t + \beta + v_1 + v_2}{\varphi_t} - \sqrt{\left(\frac{\varphi_t + \beta + v_1 + v_2}{\varphi_t}\right)^2 - \frac{4\beta\alpha}{\varphi_t}\left(1 + a_2 v_2 + \alpha_1 (v_1 + v_2)\right)}\right).$$

Using this expression for $\varphi^{-1}(\varphi_t)$ inside (23) allows one to derive a stochastic differential equation for φ_t.

3.3 Recursive preferences and risk premiums

With expected-utility preferences the model faces an important limitation: Any agent's consumption is locally deterministic and so is their marginal utility. Therefore the market price of risk in this economy is zero.

To introduce a non-zero market price of risk, in this section we allow for recursive preferences and show how to support any given dynamics for the price-dividend ratio and the market price of risk jointly. The construction of the appropriate processes η_t^d and η_t^l is conceptually similar to the construction in the previous section. Hence, in order to avoid repetition, we simply state the main results and refer the reader to the appendix for the derivations.

Specifically, we continue to assume that investors have unit intertemporal elasticity of substitution, but allow for a risk aversion higher than one. In mathematical terms, the consumer's instantaneous utility flow is given by the aggregator

$$f\left(c_{t,s}, V_{t,s}\right) = \beta\gamma V_{t,s}\left(\log\left(c_{t,s}\right) - \gamma^{-1}\log\left(\gamma V_{t,s}\right)\right), \tag{27}$$

in that

$$V_{t,s} = \mathrm{E}_t\left[\int_t^\infty f\left(c_{u,s}, V_{u,s}\right) du\right]. \tag{28}$$

Here, $V_{t,s}$ is a consumer's value function and $\gamma < 0$ is a parameter that controls the risk aversion of the investor. Utilities of this form are introduced and discussed extensively in Duffie and Epstein (1992). They correspond to the continuous-time limit of Epstein-Zin-Weil utilities with unit elasticity of substitution.

Since preferences are homothetic, the hazard rate of death is constant, and the investment opportunities are the same for all existing agents, it follows that $\frac{\dot{c}_{t,s}}{c_{t,s}}$ continues to be independent of the cohort s to which the consumer belongs. Accordingly, equation (14) continues to hold and so do Lemmas 1 and 2, which follow from agents' budget constraints. Since $\frac{\dot{c}_{t,s}}{c_{t,s}}$ is independent of s, we shall henceforth write $\frac{\dot{c}_t}{c_t}$.

The only object that changes when agents have recursive preferences is the stochastic discount factor, described by the following result.

Lemma 5 *Define the process Z_t as the solution to the backward stochastic differential equation*

$$Z_t \equiv E_t \int_t^{\infty} e^{-\beta(u-t)} \left(\gamma \left(\frac{\dot{c}_u}{c_u} \right) + \frac{1}{2} \sigma_{Z,u}^2 \right) du, \tag{29}$$

where $\sigma_{Z,t}$ is the volatility of Z_t. Then the stochastic discount factor evolves according to

$$\frac{dm_t}{m_t} = -r_t dt - \kappa_t dB_t, \tag{30}$$

where r_t, the interest rate in this economy, continues to be given by equation (19)*, while κ_t is the market price of risk in this economy and is given by $\kappa_t = -\sigma_{Z,t}$.*

Recursive preferences imply a risk premium. Intuitively, a risk premium arises because investors worry not only about the immediate impact of the fluctuations associated with the share processes, but also about the "long run" impact of these risks on their consumption. As we discussed in the previous subsection, while the immediate impact is locally predictable, the long run impact is uncertain. This long run impact is captured by the definition of the process Z_t, and the magnitude of the market price of risk (or Sharpe ratio) κ_t reflects the volatility of Z_t.

We next ask a question similar to the one we asked in the previous subsection. Is it possible to choose diffusion processes for η_t^l and η_t^d to support given stock-market dynamics (q_t) and given dynamics of the Sharpe ratio (κ_t)?

To provide an answer to this question, we proceed as in the previous section. Specifically, we fix functions f_Z, σ_Z, f_q, and σ_q and intervals $[Z^{\min}, Z^{\max}]$ and $[q^{\min}, q^{\max}]$ and try to determine a (vector) Markov process (η_t^l, η_t^d) such that the equilibrium process Z_t — to target a particular Sharpe ratio κ_t, all we need is that the process Z_t have volatility $\sigma_Z(Z_t) = -\kappa_t$ — has support in $[Z^{\min}, Z^{\max}]$ and follows the dynamics

$$dZ_t = f_Z(Z_t) + \sigma_Z(Z_t) dB_t, \tag{31}$$

while the process for q_t has support in $[q^{\min}, q^{\max}]$ and follows the dynamics

$$dq_t = f_q(q_t) + \sigma_q(q_t)\, dB_t. \tag{32}$$

As we show in the appendix, the equilibrium dynamics of Z_t and q_t obey equations (31) and (32) when and only when the functions f_Z and f_q satisfy the relations

$$f_Z(Z_t) = \beta Z_t + \gamma\nu_t - \frac{1}{2}\sigma_Z^2(Z_t) - \gamma(\lambda + g) \tag{33}$$

$$f_q(q_t) = (\beta + \varphi_t)\, q_t - \beta\alpha\varphi_t q_t^2 - 1 - \sigma_Z(Z_t)\sigma_q(q_t). \tag{34}$$

Comparing the right-hand sides of (34) and (20) shows that the two expressions are identical, except for the last term in equation (34), which captures the presence of an equity premium.

Solving for ν_t from equation (33) and for φ_t from equation (34) leads to

$$\nu(Z_t) = \frac{1}{\gamma}\left(f_Z(Z_t) + \frac{1}{2}\sigma_Z^2(Z_t) - \beta Z_t\right) + \lambda + g \tag{35}$$

$$\varphi(q_t, Z_t) = \frac{1 - \beta q_t + f_q(q_t) + \sigma_Z(Z_t)\,\sigma_q(q_t)}{q_t(1 - \beta\alpha q_t)}. \tag{36}$$

Once again, we wish to be able to invert this mapping, which we can do under the conditions given in the following lemma.

Lemma 6 *Suppose that* $\frac{d\nu}{dZ} > 0$ *for all* $Z \in \left[Z^{\min}, Z^{\max}\right]$ *and also* $\frac{\partial\varphi}{\partial q} < 0$ *for any* $Z \in \left[Z^{\min}, Z^{\max}\right]$ *and* $q \in \left[q^{\min}, q^{\max}\right]$*. Then the mapping* (35)–(36) *is invertible.*

Given invertibility, we obtain, from Ito's Lemma, two (jointly Markov) diffusion processes for ν_t and φ_t that support (33) and (34) as equilibrium outcomes. The values of η_t^d and η_t^l follow easily as solutions to the linear two-by-two system constituted by $\varphi_t = \eta_t^d - \eta_t^l$ and equation (16):[14]

$$\eta_t^d = \nu_t + (1 - \alpha\beta q_t)\,\varphi_t \tag{37}$$

$$\eta_t^l = \nu_t - \alpha\beta q_t \varphi_t. \tag{38}$$

We record the formal result and provide an example immediately thereafter.

Proposition 2 *Consider intervals* $[q^{min}, q^{max}] \subset \left(0, \frac{1}{\alpha\beta}\right)$ *and* $\left[Z^{\min}, Z^{\max}\right]$ *and continuous functions* f_Z*,* f_q*,* σ_Z*, and* σ_q *such that the assumptions of Lemma 6 hold. Then there exists a unique pair of Markov diffusions* (ν_t, φ_t) *such that the equilibrium stochastic process for* Z_t *and* q_t *are given by the diffusions (31) with support* $\left[Z^{\min}, Z^{\max}\right]$ *and (32) with support* $[q^{min}, q^{max}]$*.*

The main goal of Proposition 2 is to provide an explicit mapping between assumptions on the share processes η_t^d and η_t^l and the resulting diffusion processes for the Sharpe ratio and the price-to-dividend ratio. The restrictions placed on these latter two processes by the assumptions of Lemma 6 are rather mild and in practical applications amount to simple parametric restrictions, as the next example illustrates.

Example 2 *Suppose that* x_t *follows the process* (25) *and that we wish to obtain* $Z_t = b_1 + b_2 x_t$ *and* $q_t = a_1 + a_2 x_t$ *as equilibrium outcomes with* $b_1 = \frac{\gamma}{\beta}(\lambda + g)$ *and some constants* $a_1 > 0$*,* $a_2 > 0$*, and* $b_2 < 0$*.*

In that case equation (35) implies that ν_t *must be given by*

$$\nu_t = -\frac{b_2}{\gamma}\left((v_1 + v_2)\,x_t + \beta x_t - \frac{b_2\sigma_x^2}{2} x_t(1 - x_t) - v_2\right). \tag{39}$$

We require

$$v_1 + v_2 + \beta + \frac{b_2\sigma_x^2}{2} > 0$$

[14]We note that adding a constant to both η^d and η^l shifts Z_t by a constant, but leaves its dynamics (as well as the process q_t) the same. As a consequence, one can always construct positive processes η^d and η^l by starting from an arbitrary pair $(\hat{\eta}^d, \hat{\eta}^l)$ and letting $\eta^i = \hat{\eta}^i + k$, $i \in \{d, l\}$, for k large enough — in particular, $k \geq -\min_{Z \in [Z^{min}, Z^{max}], q \in [q^{min}, q^{max}]} \hat{\eta}^i(q, Z)$, where the function $\hat{\eta}^i(q, Z)$ defined by plugging (35)–(36) inside (37)–(38) is continuous.

so that the right hand side of (39) increases in x_t for all $x_t \in [0,1]$ and ν_t can be expressed as a (Markovian) stochastic differential equation.[15] *With this specification, the Sharpe ratio is given by $|b_2|\sqrt{x_t(1-x_t)}$.*

The dynamics of φ_t that support $q_t = a_1 + a_2 x_t$ follow from equation (36)*, namely*

$$\varphi_t = \frac{1-\beta(a_1+a_2x_t)+a_2v_2-a_2\left(v_1+v_2\right)x_t+a_2b_2\sigma_x^2x_t(1-x_t)}{\left(a_1+a_2x_t\right)\left(1-\beta\alpha(a_1+a_2x_t)\right)}. \tag{40}$$

We assume that parameters are such that the right-hand side of (40) *is decreasing in x_t, so that φ_t can be expressed as a Markovian diffusionm and the relation between φ_t and q_t is invertible. (A sufficient condition is $v_1+v_2+b_2\sigma_x^2+\beta>0$ and $a_1+a_2<\frac{1}{2\alpha\beta}$.) We note here that $a_2b_2<0$ means that the Sharpe ratio is negatively related to the price-dividend ratio.*

As a final remark, we note that we have assumed throughout that Z_t and q_t (and by implication η_t^d and η_t^l) are driven by the same Brownian motion. It is straightforward to extend the analysis to allow Z_t and q_t to be driven by separate Brownian motions with some correlation coefficient ρ.[16]

3.4 Discussion and extensions

Proposition 2 is a "possibility" result, similar to the one provided in Constantinides and Duffie (1996), but predicated on qualitatively different specifications of inequality dynamics and market incompleteness.[17] It shows that the model is able to produce a wide range

[15] The function $Z(\nu_t)$ can be written explicitly as

$$Z_t = \nu^{-1}\left(\nu_t\right) = b_1 + \frac{\frac{\sigma_x^2}{2}b_2 - v_1 - v_2 - \beta + \sqrt{\left(v_1+v_2+\beta-\frac{\sigma_x^2}{2}b_2\right)^2 + 2\sigma_x^2\left(b_2v_2 - \frac{\gamma}{\beta}\nu_t\right)}}{\sigma_x^2}.$$

Using the same steps as the ones we used to arrive at (23), one can derive the stochastic differential equation for $d\nu_t$.

[16] The only modification required in that case is the replacement of the term $\sigma_Z\sigma_q$ in equations (34) and (36) with $\rho\sigma_Z\sigma_q$. As a result, equation (40) in Example 2 would feature Z on the right-hand side, as well, so that q_t follows as a function of both φ and Z, and therefore φ and ν. (To ensure market completeness, one would also need to introduce a zero net supply asset to "span" the second Brownian shock.)

[17] The market incompleteness in Constantinides and Duffie (1996) doesn't allow an existing cohort of agents to share risk, while our incompleteness stems from the fact that unborn generations cannot trade before they are born.

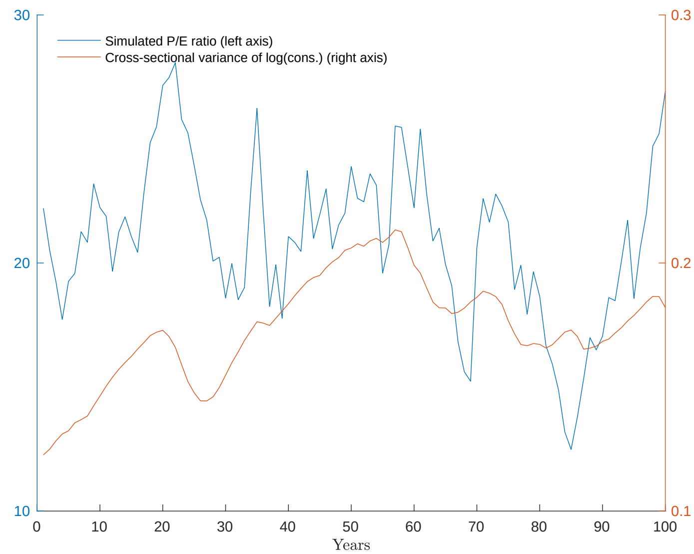

Figure 2: An indicative, model-implied path of the price-earnings ratio (left scale) and the cross-sectional standard deviation of log consumption (right scale).

of dynamics for the price-dividend ratio and the Sharpe ratio despite constant aggregate consumption and dividend growth.

It is an empirical matter to estimate the share processes and establish whether they are quantitatively consistent with the observed asset-pricing moments. We address this question in Section 4. Here we discuss (i) how this model differs from Constantinides and Duffie (1996) and (ii) how the key insights are robust to various model extensions.

One obvious difference to Constantinides and Duffie (1996) is that we do not require independent innovations to the stochastic discount factor; instead we can accommodate a Markovian structure. However, the more important difference between the two models (and indeed relative to many other models of heterogeneous agents) pertains to the dynamic behavior of inequality. To see this, it is useful to define the cross-sectional variance of log consumption as

$$\mathcal{V}_t = \lambda \int_{-\infty}^{t} e^{-\lambda(t-s)} \left(\log\left(c_{t,s}\right) - \lambda \int_{-\infty}^{t} e^{-\lambda(t-u)} \log\left(c_{t,u}\right) du \right)^2 ds.$$

Time-differentiating $\mathcal{V}_t$ we obtain the following dynamics

$$d\mathcal{V}_t = -\lambda \mathcal{V}_t dt + \lambda \left(\log\left(c_{t,t}\right) - \lambda \int_{-\infty}^{t} e^{-\lambda(t-u)} \log\left(c_{t,u}\right) du \right)^2 dt \tag{41}$$

An immediate implication of the above equation is that $\mathcal{V}_t$ is a locally deterministic process, i.e., it has no diffusion component. Accordingly, when integrated over short periods of time (say, a quarter or a year), the innovations to this process will have negligible volatility and exhibit an essentially zero covariance with asset returns. Inspection of the first term on the right hand side of (41) shows that the process $\mathcal{V}_t$ is quite persistent, since it mean-reverts at the rate λ, the rate of population renewal. By contrast, the mean reversion of the price-to-dividend ratio may be considerably higher than λ; that is, the price-dividend ratio may exhibit faster mean reversion than the cross-sectional variance of log consumption. Figure 2 provides an illustration of these properties by plotting an indicative path (of length similar to that of the post-war sample) of the price-dividend ratio and the cross-sectional variance of log consumption in the calibrated version of the model that we describe in greater detail in Section 5.

As can be seen from the figure, the model implies a positive, but weak, association between inequality and the log price-dividend ratio, with the cross-sectional variance of log consumption exhibiting a smooth path compared to the log price-dividend ratio. Moreover, the log price-dividend ratio and the cross-sectional variance of log consumption exhibit different persistence (in the calibrated version of the model the persistence of the log price dividend ratio is 0.89, while the one for the cross-sectional variance of consumption is approximately 0.97). These patterns are consistent with the data, where consumption inequality only changes by a few basis points on a yearly basis, but exhibits an autocorrelation very close to one.

We also note that this weak association between inequality and asset price movements differentiates this model from other heterogeneous-agent models in which asset price movements are driven by agents having different preferences, beliefs, etc. In these models there is strong contemporaneous correlation between asset returns and cross-sectional inequality, which is absent here.

We now turn to a couple of realistic extensions extensions of the model. Specifically, suppose that both the population size N_t and the per-capita consumption growth $\frac{C_t}{N_t}$ are stochastic and given by

$$dN_t = (-\lambda + b_t)N_t dt \tag{42}$$

$$d\left(\frac{C_t}{N_t}\right) = g_t \frac{C_t}{N_t} dt, \tag{43}$$

for some birth process b_t and some per-capita consumption growth process g_t. We have the following result.

Proposition 3 *Suppose that the share processes η_t^d and η_t^l support a given set of dynamics for the price-dividend ratio and the Sharpe ratio in the baseline economy, i.e., with $g_t = g$ and $b_t = \lambda$. Then, the modified share processes $\hat{\eta}_t^d = \eta_t^d - g + g_t + b_t - \lambda$ and $\hat{\eta}_t^l = \eta_t^l - g + g_t + b_t - \lambda$ support the same dynamics for the price-divided ratio, the interest rate, and the Sharpe ratio in an economy described by* (43).

The intuition behind Proposition 3 is straightforward. The adjustments to the share processes ensure that the entire additional growth accrues to the new cohort, leaving marginal agents' consumption growth, $\frac{\dot{c}_{t,s}}{c_{t,s}}$, the same as in the baseline model. Accordingly, the model's asset pricing implications remain unaffected.

Proposition 3 illustrates a more general point. The asset-pricing implications of the model depend only on the properties of the marginal-agent consumption growth. Modifications or extensions of the model (e.g., production, government, redistribution policies, etc.) that imply the same stochastic process for marginal agent consumption growth as our simple fluctuating endowment-share economy will have identical asset pricing implications.

With this observation in mind, we next propose a methodology to infer marginal agent consumption growth in the data, and then calibrate our model so as to reproduce its properties.

4 Empirical Implications

We focused so far on the theoretical possibility of supporting given dynamics for the price-to-dividend ratio and the market price of risk as equilibrium outcomes. In Section 4.1 we develop a methodology to infer the consumption growth rate of a cohort $\frac{\dot{c}_{t,s}}{c_{t,s}}$ and we compare its properties to the notion of aggregate consumption growth per capita. In Section 4.2 we measure η^d. In Section 5 we calibrate the model to reproduce the time-series properties of these time series.

4.1 Measuring cohort consumption growth

4.1.1. Data and preliminary observations

To motivate the results of this section, we present two figures that illustrate the differences between the consumption allocation across different cohorts. We use the tables readily available on the website of the Consumer Expenditure Survey, which report household consumption by 10-year age groups[18] (25–34, 35–44, etc.) for the year 1972 and then annually from 1984 to 2016. We isolate the years 1972, 1984, 1994, 2004, and 2014, so that there is a ten year gap between cross sections (with the exception of the first cross section where the gap is 12 years). By creating a ten year gap between cross sections we can follow the same cohort across time, since the people who were, say, born between 1950 and 1960 and were in the 25–34 age group in the 1984 cross section will be in the 35–44 age group in the 1994 cross section. For each cohort and year of observation we divide by the average (across all age groups) household consumption in that year from the same CEX tables.

Figure 3 shows that every cohort exhibits a hump-shaped consumption over the life cycle, which has a similar shape across cohorts and across time. The right plot highlights that these hump-shaped consumption profiles are roughly parallel to each other indicating the presence of a permanent, cohort-specific effect on consumption. For example, compare the cohort born between 1940–1950 with the one born between 1960–1970. The first cohort consumed 1.1 of average consumption when aged 25–34 as opposed to 0.96 for the second cohort when

[18] Age refers to the age of the head of the household.

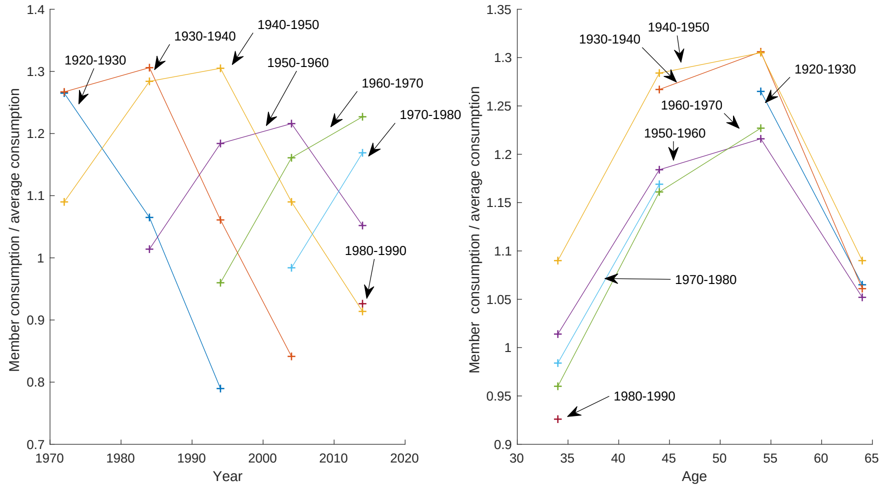

Figure 3: Average per-household consumption of cohorts born in different decades divided by average household consumption at the time of observation. The left plot depicts the information as a function of time, the right plot as a function of age.

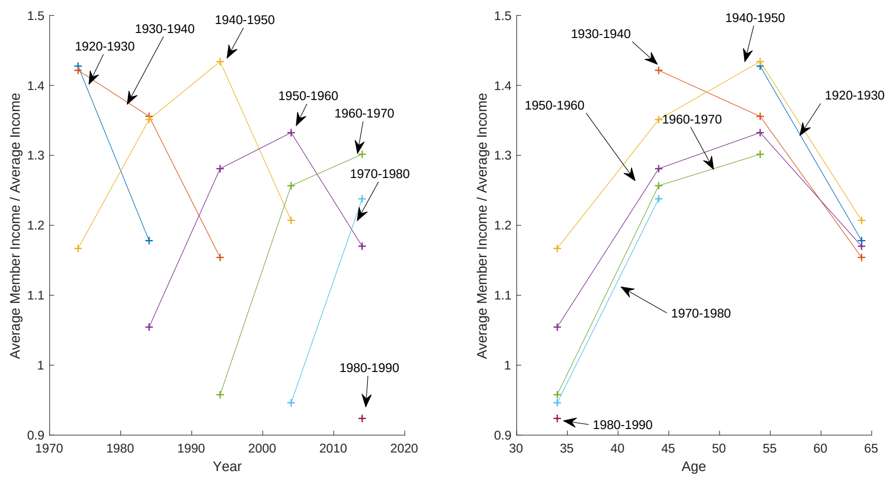

Figure 4: Average per-person income of cohorts born in different decades divided by average income at the time of observation. The left plot depicts the information as a function of time, the right plot as a function of age. For cohorts born during the decades 1920–1930, 1930–1940, and 1940–1950, we cannot provide a direct counterpart fo the last observation in Figure 3, due to reporting differences in the CPS and CEX tables.

aged 25–34. This gap persists across the life cycle with the first cohort consuming 1.29 of average consumption when aged 35–44, as opposed to 1.16 for the second cohort (when aged 35–44) and similarly for the next phase of the life cycle. In short, the figure shows that some cohorts obtain a “head start” and maintain it over their entire life as opposed to other, less lucky cohorts.

To address concerns that these patterns are special to the CEX sample, or the result of measurement biases, Figure 4 shows that exactly the same patterns characterize the income data in tables P9 and P10 on the website of the Current Population Survey (CPS). For both Figures 3 and 4 we used directly the summary tables on the websites of the CEX and the CPS (rather than the publicly available micro-data). The tables have the advantage that they contain the raw (non-top-coded) data available to these agencies. We note that the CPS provides data at the person rather than the household level, and whatever measurement errors are unlikely to be correlated across these two surveys. In results that we don't report here to save space, we have confirmed that the patterns in Figure 4 remain the same when we examine median rather than mean income, and when we produce the graph by males and females separately.

The fact that — throughout their lives — some cohorts earn and consume more or less compared to others when they were at a similar point in their life cycle is a central feature of our model. Indeed, with $s' < s$, equation (18) gives $\frac{c_{t,s'}}{c_{t,s}} = \frac{\nu_{s'}}{\nu_s} e^{-\int_{s'}^{s}(\nu_u - \lambda)du}$, which is independent of t. The same property characterizes income differences as well, which motivates the functional forms we used for the endowment specifications.

4.1.2. Estimation

Here we propose and implement a methodology to measure the gap between marginal agent- $(\frac{\dot{c}_{t,s}}{c_{t,s}})$ and aggregate consumption growth rate per capita.

In performing this exercise, we pay particular attention to the facts that i) in the data cohorts have different sizes, ii) there are hump-shaped consumption profiles over the life cycle, and iii) the population mass at any given point in time may be an arbitrary function of t and s. We take these issues into account when identifying the empirical counterpart of ν_t that we use for our calibration.

We start by remarking that, inside the model,[19] log-consumption of a given agent can be

[19]This follows by integrating the Euler equation, which gives $A_s = \log c_{s,s} - \int_{t_0}^{s} r_u du$, $L_t = \int_{t_0}^{t} r_u du$, and $G_{t-s} = -\rho(t-s)$.

decomposed as the sum of a cohort effect, a time effect, and an age effect:

$$\log c_{t,s} \equiv A_s + L_t + G_{t-s}. \tag{44}$$

To allow for salient empirical features, we do not impose the linear age effects implied by our baseline model, but rather estimate a general function G_{t-s}.

Differentiating (44) with respect to t, we obtain the consumption growth of the marginal agent as

$$d\log c_{t,s} = dL_t + \dot{G}_{t-s}dt. \tag{45}$$

There are two components to the fluctuations of $d\log c_{t,s}$: The first component captures the (stochastic) changes in the time effect, dL_t, which are common across all cohorts. The second component captures the deterministic changes arising from aging effects, $\dot{G}_{t-s}dt$. For any asset pricing model that implies an Euler equation at the level of a cohort, only the first component is of interest: Since all cohorts face the same investment opportunity set, all fluctuations in the investment opportunity set (i.e., movements in the real interest rate and the market price of risk) must be reflected in dL_t.[20] Besides, the aging effect is deterministic, so it cannot reflect stochastic fluctuations in the investment opportunity set. (A practical implication consequence of this observation is that models may differ in their implications for age effects, but will still have the same *asset pricing implications* as long as they imply the same dL_t.) Therefore, our objective is to measure L_t, and reproduce its properties in our calibration.

In principle, L_t, together with A_s and G_{t-s}, can be obtained by regressing $\log c_{t,s}$ on time, age, and cohort dummies. Such an approach is, however, limited by the lack of availability of long time series of cross-sectional consumption data. (Annual CEX cross sections start

[20] To see this more formally, suppose that we introduce age-effects into our model by introducing an age-dependent discount factor ρ_{t-s}. Then the Euler equation becomes $\frac{\dot{c}_{t,s}}{c_{t,s}} = r_t - \rho_{t-s}$, which can be written in the regression form (44) with $L_t = \int^t r_u du$, $A_s = \log c_{s,s} - \int^s r_u du$, $G_{t-s} = -\int_s^t \rho_u du$. This implies that the relation $dL_t = r_t dt$ is unaffected by the presence of arbitrary age effects in consumption. The same goes for cohorts of different sizes, etc., which do not affect the validity of the Euler equation for each member of a cohort.

in the mid-eighties). Quite remarkably, however, equation (44) allows an indirect inference approach facilitating the identification of L_t even for times t for which cross-sectional data are not available.

Specifically, in a first step, the cohort (A_s) and age (G_{t-s}) effects in equation (44) can be estimated from the available cross sections of consumption data. (A limitation of this regression is that A_s and G_{t-s} can only be identified up to an affine term, but, as we explain shortly, this does not matter for our purposes).

Second, with $\Lambda_{t,s}$ the population of cohort s at time t, aggregating equation (44) yields

$$C_t = \int_{-\infty}^{t} \Lambda_{t,s} e^{A_s + L_t + G_{t-s}} ds. \tag{46}$$

Defining

$$F_t \equiv \log \left(\int_{-\infty}^{t} \Lambda_{t,s} e^{A_s + G_{t-s}} ds \right) \tag{47}$$

and taking logarithms in equation (46) implies

$$\log(C_t) = L_t + F_t, \tag{48}$$

and therefore

$$dL_t = d\log(C_t) - dF_t. \tag{49}$$

Equation (49) presents an indirect way to infer the variation in dL_t. Computing F_t does not require a long time series of cross-sectional consumption data, since the only estimated quantities that enter the equation are the cohort effects (A_s) and the age effects (G_{t-s}) — thus, not the time effects (L_t). In principle, just two cross sections $T-1$ and T suffice to compute a long path of cohort and age effects, with more cross sections reducing the estimation error.

A common concern with regression (49), which forms the foundation for our indirect inference, is that time, age, and cohort effects can only be identified up to an affine term. This means that the data cannot distinguish the model of equation (49) from the alternative model

$$\log c_{t,s} = \underbrace{A_s + \chi s}_{\text{modified cohort effect}} + \underbrace{L_t - \chi t}_{\text{modified time effect}} + \underbrace{G_{t-s} + \chi\,(t-s)}_{\text{modified age effect}},$$

for some arbitrary constant χ. The implication of this non-identifiability is that dF_t (and hence dL_t) can only be identified up to an additive constant.[21] However, the variation of dF_t (and hence dL_t) around its mean is uniquely identified, and this variation is the relevant quantity for asset-pricing purposes.

To contrast with the classical, representative-agent based, approach to asset pricing, we focus on the portion of L_t that is in excess of the per-capita aggregate consumption growth, $d\log(C_t/N_t)$:

$$dL_t - d\log(C_t/N_t) = -dF_t + d\log(N_t). \tag{50}$$

In the model, the quantity on the right-hand side equals $-\nu_t$ (up to a constant), and it is therefore the variations in this quantity that our calibration of ν_t targets.

To implement this approach, we need data on annual aggregate consumption growth $d\log C_t^A$ (available from the BEA since 1929), the population $\Lambda_{t,s}$ of people alive at time t that were born at time s (available from the Census at annual frequency since 1910), and cross-sectional consumption data to estimate the cohort effects A_s and the age effects G_{t-s} (available from the Consumer Expenditure Survey since 1996 and from the NBER between 1984-2003).[22]

Since in the data we can only obtain estimates of cohort effects for people born from

[21] This statement can be proven by using the modified age and cohort effects inside (47) to obtain

$$dF_t^{(\chi)} \equiv d\log\left(\int_{-\infty}^{t} \Lambda_{t,s} e^{A_s+\chi s+G_{t-s}+\chi(t-s)} ds\right) = d\log\left(e^{\chi t}\int_{-\infty}^{t} \Lambda_{t,s} e^{A_s+G_{t-s}} ds\right) = \chi dt + dF_t.$$

[22] The online data on the CEX start in 1996. The National Bureau of Economic Research (NBER) contains CEX extracts that cover the period 1984-2003.

1890 onwards, we extrapolate linearly cohort effects prior to 1890. In Appendix B we report results from alternative extrapolation or truncation methods. We show that, as long as we focus attention on post world-war-II data (and especially post 1970 data), all methods give the same results, since the population weight of pre-1890 cohorts becomes relatively small. That appendix also contains a detailed exposition of many further choices that we need in order to obtain a measurement of F_t, along with a discussion of potential issues related to measurement error, discrepancies between NIPA and CEX data, etc.. We also perform a validation exercise, whereby we compare the results of our indirect inference approach to estimating L_t with a direct estimation of L_t since the seventies.

We conclude this section with a comment. Throughout our empirical exercise we chose to focus on consumption data rather than income data, since it is consumption (rather than income) that enters investors' Euler equation. An alternative approach to calibration would involve measuring η^l utilizing a time, age, and cohort decomposition for income. Inspection of Figures 3 and 4 shows that the cohort and age effects for income and consumption data are quite similar. This implies that the age and cohort-effect estimates that enter our measure of F_t are unlikely to be driven by measurement errors that are specific to consumption data.

To focus on results, we relegate all measurement-related details to the appendix and continue with a presentation of the results.

4.1.3. Comparing aggregate and marginal-agent consumption growth

Figure 5 plots the estimated (age-independent component of the) consumption growth of marginal agents, $\Delta L_t = \Delta \log C_t - \Delta \log F_t$, where Δ is the first difference operator at annual frequency. To ensure that our results are not driven by extrapolations of pre-1890 cohorts etc., we focus on the post-1950 sample. For comparison purposes, the figure also plots per capita consumption growth $\Delta \log C_t - \Delta \log N_t$, where N_t is the US population at time t. It is useful to recall that the the average level of marginal consumption growth cannot be identified, and therefore only the variations (rather than the level) of the series are meaningful.

Clearly, the two series look quite similar at annual frequencies, since in the short run

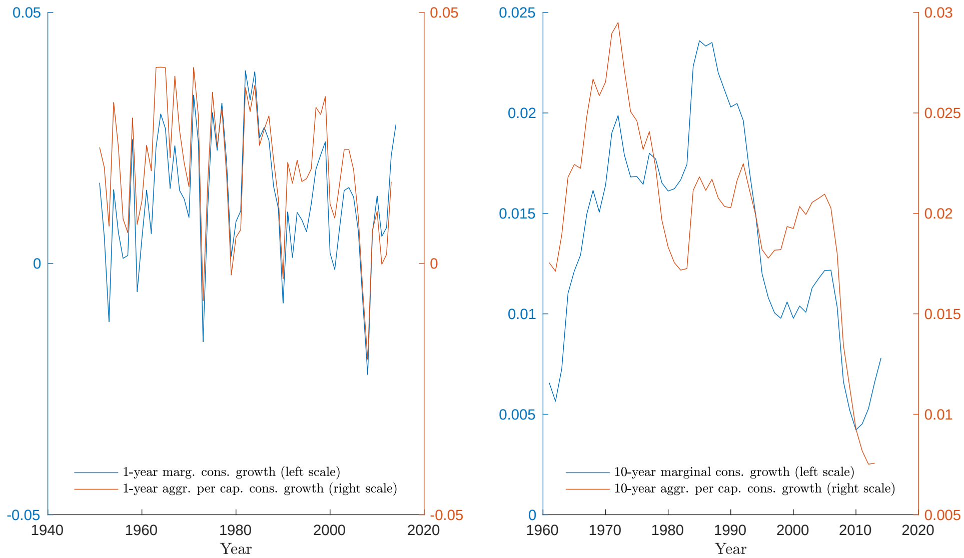

Figure 5: Left plot: Yearly per capita consumption growth of the marginal agent $\Delta L_t = \Delta \log C_t - \Delta \log F_t$ (left scale) and yearly (aggregate) per capita consumption growth $\Delta \log C_t - \Delta \log N_t$, (right scale). Right plot: Same plot, but for 10-year moving averages of both series.

the movements in both series are dominated by their common component, namely aggregate consumption growth $\Delta \log C_t$. However, the two time series look noticeably different when we time aggregate them over 10 years, suggesting that they have different low-frequency components.

Figure 6 illustrates the source of the low-frequency difference by plotting the difference between marginal and aggregate consumption growth rate per-capita, $\Delta \log N_t - \Delta \log F_t$. The figure shows that the difference between marginal and per-capita consumption growth rate has small year-over-year volatility, but is quite persistent, consistent with our model.

The noticeable low-frequency cycles exhibited in Figure 6 are driven by two basic forces in the data, namely the baby boom and the relative economic weakness of cohorts that are born after the baby boomers. The figure shows that the difference between marginal consumption growth and aggregate per capita consumption growth peaks in 1980, hits a trough in 2000 from which point onward it starts increasing again. The slowdown starting in the 1980s is driven by the fact that the fraction of aggregate consumption accruing to middle-aged and older age-groups (say ages 45 and above) has been fairly stable over time (Figure 13

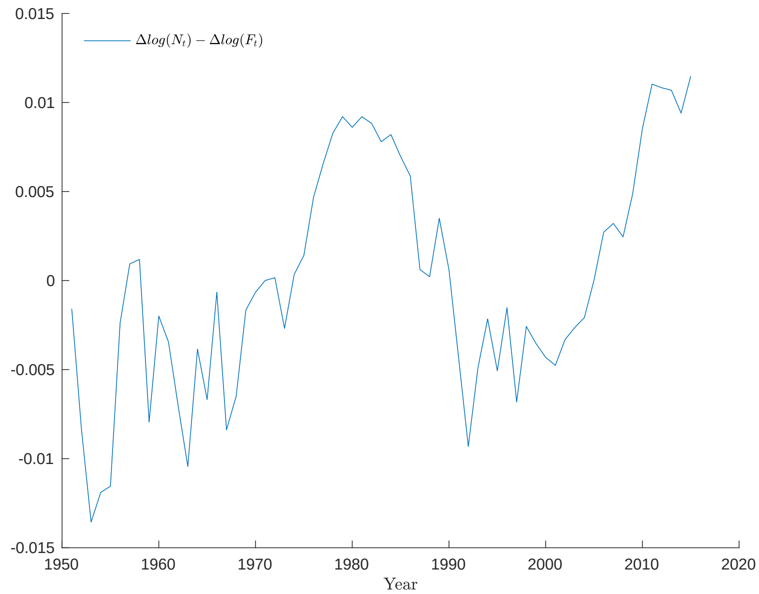

Figure 6: Difference between marginal agent consumption growth and per capita consumption growth. The difference is normalized to have mean zero.

in Appendix B). As the populous baby boomers start becoming members of the middle-aged group in the eighties and early nineties, the mathematical implication is that the per-household consumption growth of the cohorts that are middle-aged in the mid-eighties must exhibit a slowdown when compared to aggregate per capita consumption growth. This effect reverses in early-to-mid 2000, as the cohorts that start joining the middle-aged population are both smaller and less economically successful.

We note, however, that aggregate consumption growth per capita has itself slowed down since mid 2000, which explains why the rolling, 10-year moving average of marginal-agent consumption growth in Figure 5 has been steadily declining since the late eighties, unlike aggregate consumption growth per capita, which has remained fairly stable from the mid-seventies to the mid 2000s. Indeed, as the right plot of Figure 5 illustrates, the ten-year moving average of marginal-agent consumption growth has been on a steady decline since the mid-eighties, unlike aggregate consumption growth rate per capita, which starts declining in the mid-2000s.

Table 1 provides a formal econometric framework to model the joint time series properties

	$\Delta \log C_t - \Delta \log N_t$	$\Delta \log N_t - \Delta \log F_t$
Lag $\Delta \log C_t^A - \Delta \log N_t$	0.4546	-0.1004
	(0.1334)	(0.3128)
Lag $\Delta \log N_t - \Delta \log F_t$	-0.0134	0.9203
	(0.0301)	(0.0707)
R^2	0.23	0.81
σ_ϵ	0.0111	0.0025
N(obs)	87	87

Table 1: Bivariate Vector Autoregression of i)per capita consumption growth ($\Delta \log C_t - \Delta \log N_t$) and ii) the difference between marginal and per capita consumption growth ($\Delta \log N_t - \Delta \log F_t$).

of i) aggrregate consumption growth per capita and ii) the difference between marginal and per capita consumption growth rate as a bivariate first order vector autoregression

$$\begin{bmatrix} \Delta \log C_t - \Delta \log N_t \\ \Delta \log N_t - \Delta \log F_t \end{bmatrix} = \begin{bmatrix} \Delta \log C_{t-1} - \Delta \log N_{t-1} \\ \Delta \log N_{t-1} - \Delta \log F_{t-1} \end{bmatrix} B + \begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix}.$$

Using the estimates for B in the above equation and the covariance matrix of the residuals Σ thus obtained, we compute the long-run covariance matrix of the two time series,

$$\Omega \equiv \left[I + B + B^2 + ...\right] \Sigma \left[I + B + B^2 + ...\right]' = \frac{1}{100} \times \begin{bmatrix} 0.0481 & -0.0280 \\ -0.0280 & 0.1061 \end{bmatrix}. \tag{51}$$

The two time series exhibit low correlation at frequency zero (the correlation implied by Ω is around -0.39). Hence, the re-distributional risk that arises from imperfect risk sharing is a separate source of long-run consumption uncertainty, and fluctuations in aggregate consumption growth per capita don't seem to offset them. (This is part of the reason why we chose to abstract from fluctuations in aggregate consumption growth in our model.)

An additional implication of this low correlation is that the the marginal agent consumption growth, which is the sum of $\Delta \log C_t - \Delta \log N_t$ and $\Delta \log N_t - \Delta \log F_t$), has almost twice as high a long-run variance (the sum of all elements of Ω) as the long-run variance of aggregate consumption growth rate per capita (the top left element of Ω). This higher volatility of the low-frequency components of marginal consumption growth results in a higher risk

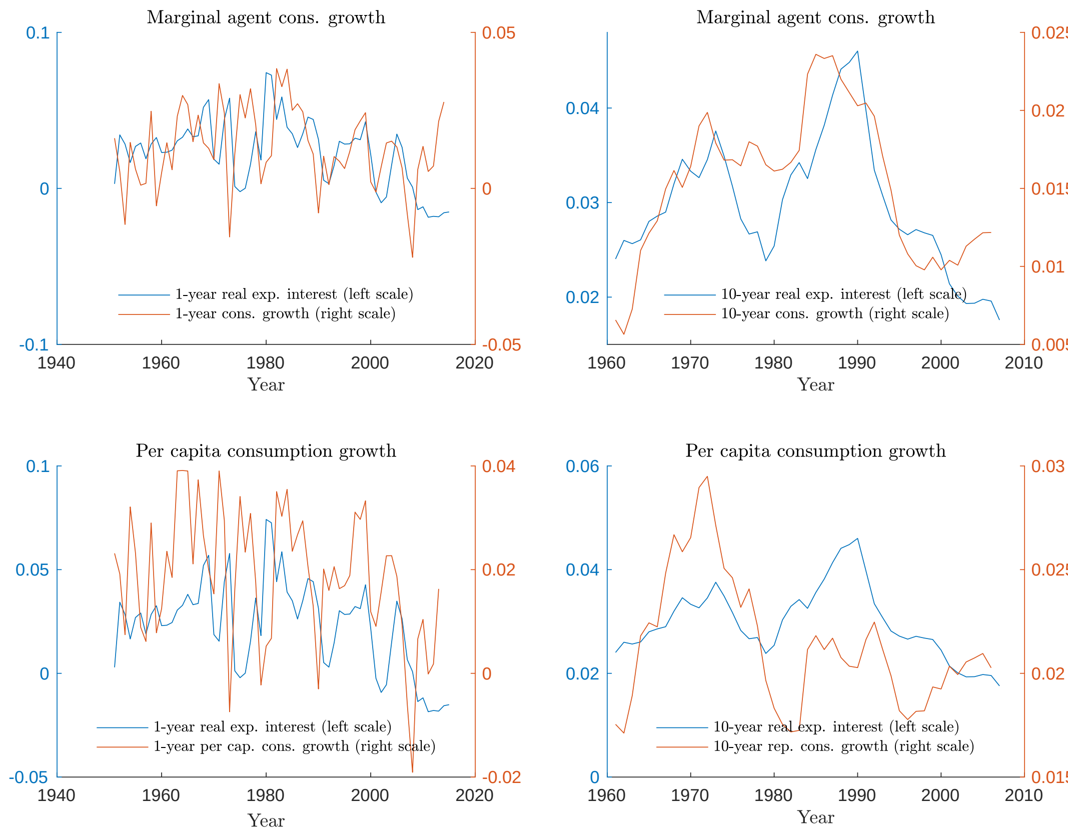

Figure 7: Top left plot: Expected real interest rate at the beginning of each year and marginal agent consumption growth over the year. Top right plot: 10-year moving averages yearly marginal agent consumption growth and 10-year moving average of expected real interest rate. Bottom left and right plots: Identical to the top plots, except that marginal agent consumption growth is replaced with aggregate consumption growth per capita.

premium with recursive preferences.

4.1.4. Relating marginal-agent consumption growth to the interest rate

We conclude this section by investigating the model-implied link between our measure of marginal-agent consumption growth and the expected real interest rate.

To measure the expected real interest rate, we use the short term nominal interest rate from Robert Shiller's online data set minus the (ex ante) expected inflation rate as formed in December of the preceding year (Source: Philadelphia Fed inflation expectations survey).

The top left plot of Figure 7 plots the expected real interest rate and marginal agent consumption growth for the respective year. Clearly the two series differ, because in reality there are shocks to aggregate consumption from which our model abstracts. However, if the real interest rate reflects the expected (rather than the realized) marginal agent consumption growth, then we should find that the co-movement between the two series rises as we time aggregate the two series over longer horizons. Indeed, this is what the top right plot of Figure 7 shows. Ten-year moving averages of expected consumption growth and real expected interest rates exhibit very similar fluctuations (the correlation coefficient is approximately 0.8). The bottom two plots show that, by comparison, this co-movement is weaker for aggregate per capita consumption growth. This strong correlation between the low frequency movements of the real interest rate and the marginal-agent consumption growth is driven by the secular decline in marginal-agent consumption growth that starts in the mid-eighties, which is the same time when the real interest rate starts declining.

In the appendix (Figure 15), we show that the high R^2 obtained when regressing 10-year moving averages of marginal agent consumption growth on similar moving averages of expected real interest rates is unlikely to be the result of randomness by performing a bootstrap exercise enforcing the null hypothesis that the two series are uncorrelated.

4.2 The measurement of η_t^d

The model implies the following result.

Lemma 7 *Let $P_t^A = \int_{-\infty}^t P_{t,s} ds$ denote aggregate market capitalization and let $\pi_s = \frac{P_{t,s}}{P_t^A}$ denote the market-capitalization weight of firms of vintage s. Then*

$$\eta_t^d = \underbrace{\overbrace{\frac{P_{t,t}}{P_t^A}}^{\substack{\textit{Market capitalization} \\ \textit{of firms added in year } t}}}_{\textit{Total market cap}} = \underbrace{\frac{dP_t^A}{P_t^A}}_{\substack{\textit{Aggregate} \\ \textit{market capitalization growth}}} - \underbrace{\int_{-\infty}^t \pi_s \frac{dP_{t,s}}{P_{t\,s}} ds}_{\substack{\textit{Market capitalization growth of} \\ \textit{firms already in the market portfolio}}} . \tag{52}$$

The first equality in (52) provides a straightforward empirical proxy for η_t^d: the ratio of the market value of additions to the market index to the total market value of the index.

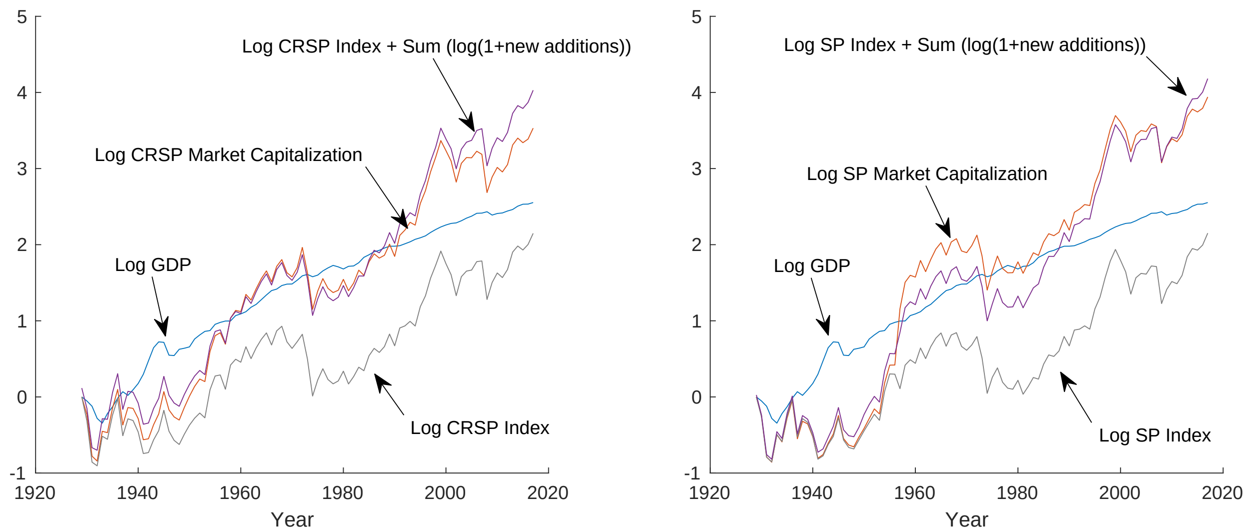

Figure 8: Left plot: Logarithm of GDP, logarithm of the CRSP market index (obtained by cumulating ex-dividend CRSP gross returns), log market capitalization, and log market capitalization plus the cumulative sum of the logarithm of 1+addition rate to the market. The addition rate is defined as the market value of additions to the index (valued at the end of each year) divided by the total value of the index at the end of each year. All series are deflated by subtracting the logarithm of the CPI. Right plot: Same as left plot, except that the market index is the S&P 500.

We use this measure in our calibration. According to the model, this ratio also equals the discrepancy between aggregate market capitalization growth and the market capitalization growth of firms already in the market portfolio, i.e., the (ex-dividend) return on the index.

Figure 8 illustrates equation (52) in the data. The solid line in the figure depicts the (log) level of the market index. The figure also depicts the aggregate gross domestic product (GDP) series and the total stock market capitalization. The figure shows that the log-level of the index (which is identical to the cumulative sum of the log gross rates of ex-dividend returns) follows a markedly slower growth than the aggregate stock market capitalization. Interestingly, the discrepancy between these two comes down almost entirely to the value of additions to the index: adding the cumulative sum of $\log\left(1+\frac{P_{t,t}}{P_t^A}\right)$ to the log index results in a series that tracks the market capitalization growth very closely.

Figure 9 compares the path of the discrepancy between marginal and per capita consumption growth (ν_t) that we obtained in the previous subsection with the path of η_t^d. The graph starts in 1963 so that AMEX is already incorporated in CRSP and we treat 1972 (the

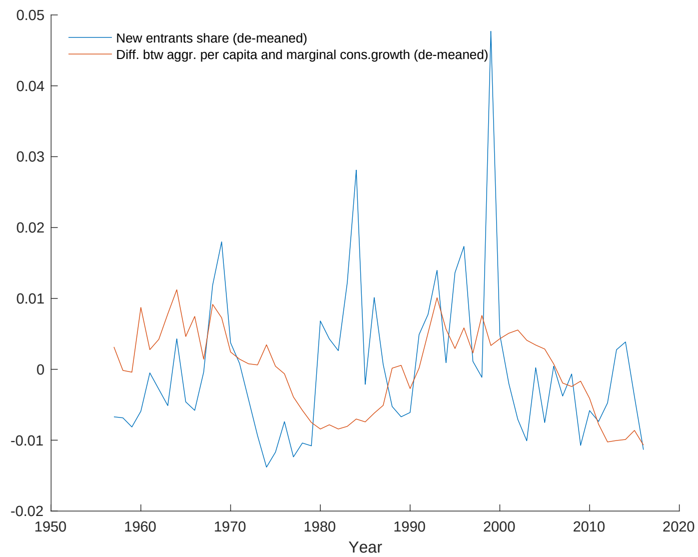

Figure 9: Market share of new entrants (η_t^d) and difference between aggregate per capita and marginal agent consumption growth (ν_t).

year when NASDAQ is added to CRSP) as an outlier (we just replace the value in 1972 by the average of the values in 1971 and 1973). An additional advantage of focusing on the post-1963 sample is that the weight of cohorts born prior to 1890 does not impact ν_t. The figure shows that while the two series differ on a year-to-year basis, they appear to share the same low-frequency cycle.[23]

5 Calibration

The exercise we perform is straightforward. First, we choose functional forms for the dynamics of η_t^l and η_t^d. We choose these parametric forms judiciously so as to support closed-form solutions for the dynamics of the price-dividend ratio q_t and the Sharpe ratio κ_t. Second, we choose the parameters governing the dynamics of η_t^l and η_t^d to match the empirical moments

[23]To formalize this idea we computed the cumulative sums of $\log(1+\nu_t)$ and $\log(1+\eta_t^d)$ and performed a Johansen rank statistic for co-integration with restricted trend, since the mean of ν_t is not identified. That test fails to reject the hypothesis of a co-integrating relation between the two cumulative sums at the 5% significance level. This means that a shock that widens the distance between the market capitalization of existing firms and total market capitalization in a permanent fashion and a shock that widens the distance between marginal agent and per-capita consumption growth in a permanent fashion are perfectly correlated.

v_1	0.012	β	0.03	γ	8
v_2	0.078	α	0.3	g	0.025
σ_x	0.12	a_1	1.1	a_2	2.25
$\lvert b_2 \rvert$	7				

Table 2: Parameters used in the model calibration.

of η_t^d and ν_t in the data. (Note that by equations (38) and (37) there is a one-to-one correspondence between the pairs (η_t^l, η_t^d) and (η_t^d, ν_t).) Then we examine the resulting moments of asset-price dynamics and compare them to the data.

Specifically, we employ a functional form specification similar to Example 2. Using the definition of x_t in equation (25), we specify ν_t as in equation (39). For φ_t we choose the functional specification

$$\varphi_t = \frac{e^{-a_1 - a_2 x_t} - \beta + a_2\left(-v_1 x_t + v_2(1 - x_t)\right) + \left(\frac{a_2^2}{2} - |b_2|\right)\sigma_x^2 x_t(1 - x_t)}{1 - \alpha\beta e^{a_1 + a_2 x_t}}. \tag{53}$$

With these specifications, the price-to-dividend ratio is log-linear in x_t, $\log(q_t) = a_1 + a_2 x_t$, while the Sharpe ratio is given by $|b_2|\sqrt{x_t(1 - x_t)}$.

We fix preference parameters to $\beta = 0.03$ (sum of discount and death rates) and $\gamma = -8$, which maps into a risk aversion coefficient of $1 - \gamma = 9$. We set α to a level that reflects the share of capital income in output (0.3). The aggregate growth rate is set to 0.025, in line with historical data. We choose the six parameters (v_1, v_2, σ_x, b_2, a_1, and a_2) that govern the dynamics of ν_t and φ_t in equations (39), respectively (53), to (approximately) match six moments, namely the stationary mean, stationary standard deviation, and autocorrelation of the inferred values of ν_t and η_t^d in the data. To confirm that the functional forms that we chose (motivated by the desire to obtain simple, closed-form solutions for the price-dividend ratio and the Sharpe ratio) are consistent with the data, we perform a Kolmogorov-Smirnov test, which cannot reject that the model-implied stationary distributions of ν_t and η_t^d is different from the respective stationary distributions in the data. We note that since the mean value of ν_t cannot be identified with the time-, age-, and cohort- decomposition method of the previous section, we target instead $\frac{c_{t,t}}{Y_t} \approx 1$ motivated by the evidence in Figure 3.

	Data	Model
Median arrival rate of new firms	2.2%	2.30% (1.19%)
Standard deviation of the arrival rate of new firms	1.6%	1.94% (0.98%)
Autocorrelation of the arrival rate of new firms	0.77	0.88 (0.0617)
Median value of ν_t		2.97% (1.19%)
Standard deviation of of ν_t	0.55%	0.55% (0.23%)
Autocorrelation of imputed ν_t	0.89	0.89 (0.0613)

Table 3: Targeted moments, model and data. We simulate 1000 independent paths of similar length as the data, and compute each of the six moments for every path. We then report the mean and standard deviation (across the 1000) paths for each moment. The term "arrival rate of new firms" refers to the ratio of the value of the market value of additions to the market portfolio to the total value of the market portfolio.

This implies a mean value of ν_t close to the sum of the birth and (net) immigration rate, which is about 2.5% in our sample.[24]

Table 3 shows that these parameter choices allow us to plausibly reproduce the targeted empirical moments within our model. To account for estimation error, we do not only report average values of the targeted moments within our model, but also the standard deviation for the model-implied values, when we simulate our model over similar sample lengths to the data. The table shows that the moments in the data are within two standard deviations of their simulated means inside the model. Figure 10 provides an alternative, graphical illustration of Table 3 by comparing the empirical and the simulated distributions of ν_t and φ_t.

Having determined the parameters to match the moments of the share processes, we next examine what these parameter choices imply for asset pricing moments. Table 4 provides a comparison between the model-implied unconditional moments and the respective moments in the data. In reporting the results we follow the approach of Barro (2006) to relate the

[24]We compute the sum of birth and immigration rates as the sum of the population growth rate (around 1.5% for most of our sample) plus the death rate (around 1% for most of the sample of interest).

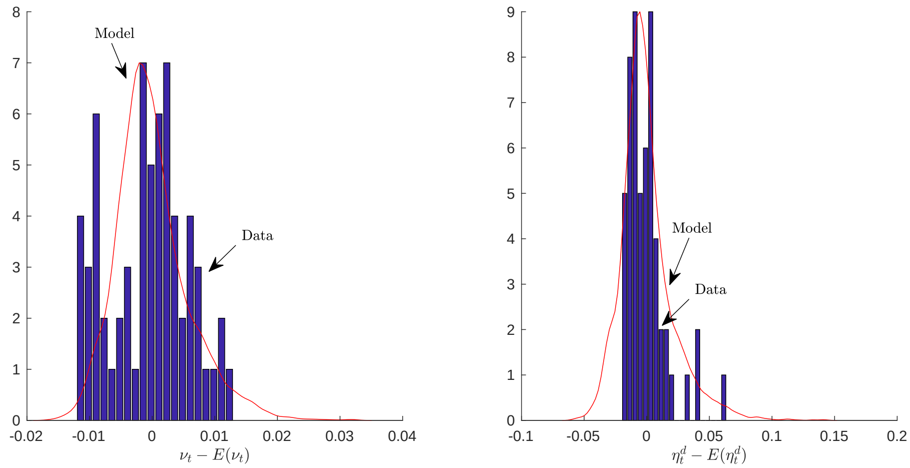

Figure 10: Left plot: Histogram of (de-meaned) inferred values of ν_t and kernel density of the respective quantity inside the model. Right plot: Histogram of the market capitalization of new index additions over the existing market capitalization of the index (de-meaned) and kernel density of the respective quantity inside the model. To obtain the model-implied quantities, we simulate 1000 paths of length identical to the length of the data sample, and de-mean the simulated data separately on each sample path, to account for sampling error in the computation of the means. We then compare the distribution of the de-meaned simulated data to the de-meaned empirical data.

results of our model (which produces implications for an all-equity financed firm) to the data (where equity is levered). Specifically, we use the well known Modigliani-Miller formula, according to which the levered equity return is equal to the un-levered equity return times 1.7 (the leverage ratio in the data (see e.g., Barro (2006)).

Inspection of Table 4 shows that the model accounts for a sizable fraction of all asset pricing moments. To put these numbers in the proper relation to the literature, it is worth highlighting that aggregate consumption and dividend growth are constant in this model. The numbers should therefore be interpreted as the asset-pricing moments that would obtain in an economy where one abstracts from all aggregate sources of uncertainty and examines the impact of the share processes in isolation.

Table 4 only pertains to unconditional moments. To evaluate the model's ability to account for variations in conditional moments we turn to Figure 11 and Table 5. Figure 11 plots the equity premium, market price of risk, interest rate, and stock-return volatility as a function of the log-price-earnings ratio. The Sharpe ratio and the equity premium are both

	Data	Model
Aggregate consumption growth rate	2.3%	2.3%
Standard deviation of consumption growth rate	3.3%	0%
Sharpe ratio	0.29	0.26
Stock market volatility	18.2%	14.32%
Equity premium	5.2%	4.13%
Average interest rate	2.8%	2.37%
Standard deviation of real interest rate	0.92%	0.72%
Average (log) PD ratio	2.9	3.05
Standard deviation of (log) PD ratio	0.27	0.21
Autocorrelation of (log) PD ratio	0.89	0.91

Table 4: Unconditional moments for the data and the model. The data for the average equity premium, the volatility of returns, and the level of the interest rate are from the long historical sample available from the website of R. Shiller (http://www.econ.yale.edu/?shiller/data/chapt26.xls). The volatility of the real rate is inferred from the yields of 5-year constant maturity TIPS as in Gârleanu and Panageas (2015).

declining functions of the log-price-earnings ratio. This counter-cyclicality is responsible for the model's ability to reproduce the predictability relations documented in Table 5. This table reports results of simulated predictability regressions inside the model and compares the results with the data. The main takeaway of the table is that the model-implied predictability is close to the respective time-variation in the data.

Finally, for a discussion of the model's implications for inequality, we refer the reader to Section 3.4 and specifically the discussion of Figure 2.

6 Conclusion

In this paper we propose a simple mechanism to relate low frequency movements in inequality with volatile asset-price movements. We exploit the structure of an overlapping generations economy, which allows different cohorts of agents to experience different (and random) consumption growth paths over their lifetimes, even though aggregate consumption evolves deterministically. Combining this observation with recursive preferences, we prove that an appropriate specification of endowment shares can account for a host of styl-

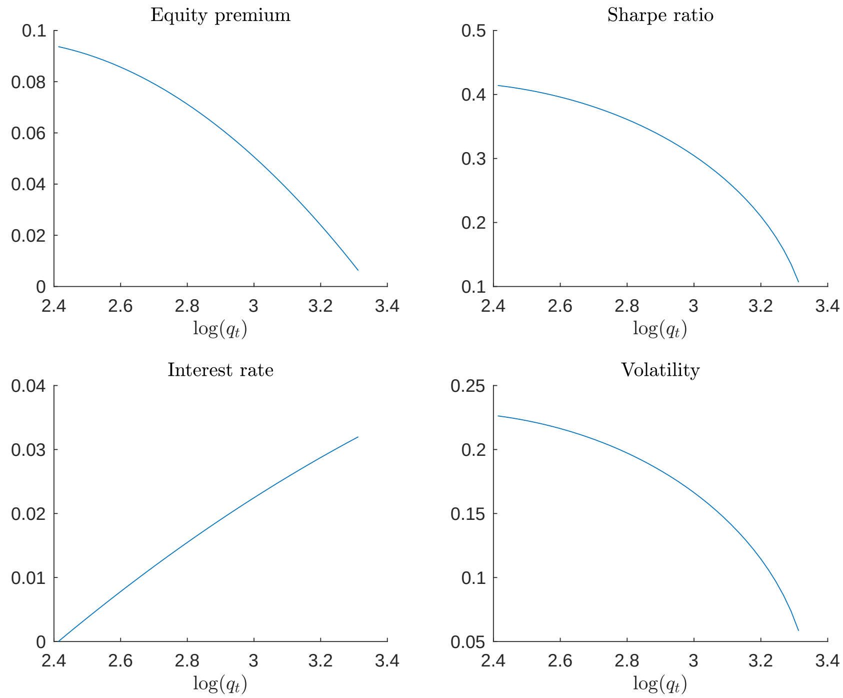

Figure 11: Calibration results. Equity premium, market price of risk (Sharpe ratio), interest rate, and stock return volatility for the baseline parametrization. We plot each variable against the (log) price-to-earnings ratio $\log(q_t)$. The range of values of $\log(q_t)$ correspond to the interval between the bottom 1% and the top 99% percentiles of the stationary distribution of $\log(q_t)$.

ized asset-pricing facts. Unlike Constantinides and Duffie (1996), our framework implies an inequality process that is non-volatile, weakly related to asset-price fluctuations at high frequencies, and substantially more persistent than the price-to-dividend ratio.

The main difference, for the purposes of asset pricing, between OLG and representative-agent economies is that the Euler equation applies at the level of a given cohort, but not at the aggregate level. This feature is quite general and applies to a broad class of OLG models. Motivated by this fact, we develop an empirical strategy to infer the discrepancy between the consumption growth at the aggregate and cohort levels by utilizing a time, age, and cohort decomposition of cross-sectional consumption data and imposing market clearing. Since it does not require time-series information, this technique can be implemented using existing cross-sectional data sources.

Year	β (Data)	β (Model)	R^2(Data)	R^2(Model)
1	-0.130	-0.086	0.040	0.032
		[-0.200 -0.007]		[0.000 0.091]
3	-0.350	-0.232	0.090	0.086
		[-0.516 -0.012]		[0.000 0.220]
5	-0.600	-0.343	0.180	0.128
		[-0.744 0.005]		[0.000 0.321]
7	-0.750	-0.429	0.230	0.163
		[-0.893 0.023]		[0.000 0.435]

Table 5: Long-horizon regressions of excess returns on the log P/D ratio. The simulated data are based on 1000 independent simulations of 100-year long samples. For each of these 100-year long simulated samples, we run predictive regressions of the form $\log R^e_{t\to t+h} = \alpha + \beta \log(P_t/D_t)$, where $\log R^e_{t\to t+h}$ denotes the time-t gross excess return over the next h years. We report the mean values for the coefficient β and the R^2 of these regressions, along with the respective [0.025, 0.975] percentiles.

We evaluate the model quantitatively and show that the discrepancy between aggregate and cohort-level consumption growth exhibits strong persistent components, which are largely independent from fluctuations in aggregate consumption growth. With recursive preferences, these persistent components lead to sizable and time-varying risk premiums, even in a world of deterministic aggregate consumption growth. Cohort-level consumption growth also exhibits a stronger medium-run comovement with expected real interest rates than does per capita consumption growth, consistent with our theory.

References

Abel, A. B. (2003). The effects of a baby boom on stock prices and capital accumulation in the presence of social security. *Econometrica 71*(2), 551–578.

Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance 59*(4), 1481–1509.

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics 121*(3), 823–66.

Blanchard, O. J. (1985). Debt, deficits, and finite horizons. *Journal of Political Economy 93*(2), 223–247.

Campbell, J. Y. and Y. Nosbusch (2007). Intergenerational risksharing and equilibrium asset prices. *Journal of Monetary Economics 54*(8), 2251–2268.

Constantinides, G., J. B. Donaldson, and R. Mehra (2002). Junior can't borrow: A new perspective on the equity premium puzzle. *The Quarterly Journal of Economics 117*(1), 269–296.

Constantinides, G. and D. Duffie (1996). Asset pricing with heterogeneous consumers. *Journal of Political Economy 104*, 219–240.

Constantinides, G. M. and A. Ghosh (2017). Asset pricing with countercyclical household consumption risk. *The Journal of Finance 72*(1), 415–460.

Duffie, D. and L. G. Epstein (1992). Stochastic differential utility. *Econometrica 60*(2), 353–394.

Ehling, P., A. Graniero, and C. Heyerdahl-Larsen (2018). Asset prices and portfolio choice with learning from experience. *The Review of Economic Studies 85*(3), 1752–1780.

Farmer, R. E. (2018). Pricing assets in a perpetual youth model. *Review of Economic Dynamics 30*, 106–124.

Gârleanu, N., L. Kogan, and S. Panageas (2012). The demographics of innovation and asset returns. *Journal of Financial Economics 105*, 491–510.

Gârleanu, N. and S. Panageas (2015). Young, old, conservative, and bold: The implications of heterogeneity and finite lives for asset pricing. *Journal of Political Economy 123*, 670–685.

Geanakoplos, J., M. Magill, and M. Quinzii (2004). Demography and the long-run predictability of the stock market. *Brookings Papers on Economic Activity 2004*(1), 241–307.

Gomes, F. and A. Michaelides (2005). Optimal life-cycle asset allocation: Understanding the empirical evidence. *Journal of Finance 60*(2), 869–904.

Gomez, M. (2017). Asset prices and wealth inequality. Working Paper, Columbia University.

Grossman, S. J. and R. J. Shiller (1982). Consumption correlatedness and risk measurement in economies with non-traded assets and heterogeneous information. *Journal of Financial Economics 10*(2), 195 – 210.

Kahn, L. B. (2010). The long-term labor market consequences of graduating from college in a bad economy. *Labour Economics 17*(2), 303 – 316.

Karlin, S. and H. Taylor (1981). *A Second Course in Stochastic Processes*. Academic Press.

Kogan, L., D. Papanikolaou, and N. Stoffman (2019). Left behind: Creative destruction, inequality, and the stock market. Journal of Political Economy, forthcoming.

Krueger, D. and H. Lustig (2010). When is market incompleteness irrelevant for the price of aggregate risk (and when is it not)? *J. Economic Theory 145*(1), 1–41.

Maurer, T. (2017). Asset pricing implications of demographic change. Working Paper, Olin School of Business.

Oreopoulos, P., T. von Wachter, and A. Heisz (2012). The short- and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics 4*(1), 1–29.

Oyer, P. (2006). Initial labor market conditions and long-term outcomes for economists. *Journal of Economic Perspectives 20*(3), 143–160.

Oyer, P. (2008). The making of an investment banker: Stock market shocks, career choice, and lifetime income. *The Journal of Finance 63*(6), 2601–2628.

Piazzesi, M. and M. Schneider (2009). Inflation and the price of real assets. Federal Reserve Bank of Minneapolis, Staff Report: 423.

Schmidt, L. D. W. (2015). Climbing and falling off the ladder: Asset pricing implications of labor market event risk. Unpublished working paper, MIT Sloan.

Schneider, A. (2017). Risk sharing and the term structure of interest rates. Unpublished working paper, Federal Reserve Board of Governors.

Storesletten, K., C. I. Telmer, and A. Yaron (2007). Asset pricing with idiosyncratic risk and overlapping generations. *Review of Economic Dynamics 10*(4), 519–548.

Toda, A. A. and K. Walsh (2015). The Double Power Law in Consumption and Implications for Testing Euler Equations. *Journal of Political Economy 123*(5), 1177–1200.

von Wachter, T. and H. Schwandt (2018). Unlucky cohorts: Earnings, income, and mortality effects from entering the labor market in a recession. forthcoming, Journal of Labor Economics.

A Proofs

Proof of Lemma 1. The absence of bubbles together with the assumption of a unit elasticity of substitution implies that aggregate consumption is given by $C_t = \beta\left(\bar{W}_t + \bar{H}_t\right)$, where

$$\bar{W}_t = \int_{-\infty}^{t} q_{t,s}^{d} D_{t,s} ds = \alpha q_t^d Y_t \tag{54}$$

is the present value of all dividends to be paid by existing firms. Similarly, the total value of all human capital of existing agents is

$$\bar{H}_t = \int_{-\infty}^{t} q_{t,s}^{l} l_{t,s} w_{t,s} ds = (1-\alpha)\, q_t^l Y_t. \tag{55}$$

Combining goods market clearing ($C_t = Y_t$) with (54) and (55) and re-arranging leads to (11).

■

Proof of Lemma 2. The present value of all newly-born workers' wages is given by $(1-\alpha)\,\eta_t^l q_t^l Y_t$, while the present value of all newly created firms is $\alpha \eta_t^d q_t^d Y_t$. The sum of these quantities gives the total wealth of newly born agents. Given that the consumption-to-wealth ratio for investors is β, the per-capita consumption of the newly born, as a proportion of total consumption, is given by (17). ■

Proof of Lemma 3. The only step of the proof not made completely explicit in the proof is the one yielding equation (14). To show this relation, time-differentiate aggregate consumption $C_t = \lambda \int_{-\infty}^{t} e^{-\lambda(t-s)} c_{t,s} ds$ to get

$$\dot{C}_t = -\lambda C_t + \lambda \int_{-\infty}^{t} e^{-\lambda(t-s)} \dot{c}_{t,s} ds + \lambda c_{t,t} = -\lambda C_t + \frac{\dot{c}_{t,s}}{c_{t,s}} \left(\lambda \int_{-\infty}^{t} e^{-\lambda(t-s)} c_{t,s} ds\right) + \lambda c_{t,t}, \tag{56}$$

where we used Leibniz's rule and the fact that $\frac{\dot{c}_{t,s}}{c_{t,s}}$ is independent of s. Dividing both sides of (56) by C_t, using $C_t = \lambda \int_{-\infty}^{t} e^{-\lambda(t-s)} c_{t,s} ds = Y_t$ and $\frac{\dot{C}_t}{C_t} = g$ leads to (14). ■

Proof of Lemma 4. Contained in the text. ■

The following technical restrictions on f and σ ensure existence of stationary q.

Assumption 1 *The functions f and σ are Lipschitz continuous on the bounded interval $[q^{min}, q^{max}] \subset (0, \frac{1}{\alpha\beta})$. Moreover, f is twice differentiable, monotonically decreasing, and satisfies $f\left(q^{min}\right) > 0$ and $f\left(q^{max}\right) < 0$. Finally, $\sigma(q) \geq 0$, $\sigma\left(q^{min}\right) = \sigma\left(q^{max}\right) = 0$, and*

$$\lim_{q \to q^{max}} \frac{\sigma^2(q)}{q^{max} - q} < 2\left|f\left(q^{max}\right)\right|, \quad \lim_{q \to q^{min}} \frac{\sigma^2(q)}{q - q^{min}} < 2\left|f\left(q^{min}\right)\right|. \tag{57}$$

Proof of Proposition 1. Let $\hat{q}_t$ be a solution to the stochastic differential equation (21) with support in $[q^{\min}, q^{\max}]$. By construction, the process $\varphi(\hat{q}_t)$ solves the SDE (23). Since $\hat{q}_t$ is bounded, so is φ_t, and we can construct two positive processes η_t^l and η_t^d such that $\eta_t^d - \eta_t^l = \varphi_t$.

Taking these two processes as given, posit that the equilibrium price dividend ratio is $q_t^d = q_t = \hat{q}_t$ and q_t^l is given by Lemma 1. Further, conjecture that the interest rate r_t is given by equation (19). We next confirm that these postulates for q_t^d, q_t^l and r_t constitute an equilibrium.

Given the dynamics of q_t, the definition of $\varphi(\cdot)$, and the definition of r_t, pricing equation (12) is satisfied. Further, using also the definition of q_t^l, which implies $\alpha dq_t^d + (1-\alpha)dq_t^l = 0$, we obtain the analogous pricing equation for q_t^l:

$$\mathrm{E}[dq_t^l] = \left(r_t - g + \eta_t^l\right) q_t^l - 1. \tag{58}$$

Agents' consumption optimality require $c_{t,t} = \beta W_{t,t}$, yielding Lemma 2, as well as the Euler equation (13). Starting with equation (13), then applying Lemma 2 and equation (19) in succession, we obtain

$$\begin{aligned} C_t &= \int_{-\infty}^{t} \lambda e^{-\lambda(t-s)} e^{\int_s^t (r_u - \rho)du} c_{s,s} ds && (59)\\ &= \int_{-\infty}^{t} e^{\int_s^t (r_u - \beta)du} \left(\eta_s^l + \alpha\beta q_s \varphi_s\right) Y_s ds && (60)\\ &= \int_{-\infty}^{t} e^{\int_s^t (r_u - \beta - g)du} \left(\beta + g - r_s\right) Y_t ds && (61)\\ &= Y_t, && (62) \end{aligned}$$

given that r_t is bounded above away from $\beta + g$. Proposed consumption processes are therefore optimal and clear the consumption market, given the interest rate.

Finally, with q_t^d and q_t^l the valuation ratios, the total wealth in the economy is

$$\frac{1}{\beta} C_t = \frac{1}{\beta} Y_t = \alpha q_t^d Y_t + (1-\alpha) q_t^l Y_t. \tag{63}$$

To see that asset markets clear, note that integrating forward the budget constraint (7) of all agents born at time s and alive at time t and taking expectations gives

$$\lambda e^{-\lambda(t-s)} W_{t,s} = \lambda e^{-\lambda(t-s)} E_t \left[\int_t^{\infty} \frac{m_u}{m_t} e^{-\lambda(u-t)} (c_{u,s} - (1-\epsilon) w_{u,s}) du \right],$$

where m_t is given by (8). Aggregating across all cohorts, and using the same arguments as in Lemma 1 shows that

$$\lambda \int_{-\infty}^{t} e^{-\lambda(t-s)} W_{t,s} = \int_{-\infty}^{t} P_{t,s} ds,$$

i.e., the stock market is clearing. Consumption market clearing and stock market clearing implies bond market clearing by Walras' law. Uniqueness of the process φ_t is a direct consequence of the analysis in the text, in particular equation (20).

We end the proof with a technical detail — a sketch of an argument that shows that q_t is stationary. We make use of results in Karlin and Taylor (1981). Specifically, we start by defining

$$s(q) \equiv \exp\left\{ - \int^{q} \frac{2 f(\xi)}{\sigma^2(\xi)} d\xi \right\},$$

noting that by assumption (57) there exists $\bar{v} > 1$ such that, for ε small enough and $q \in (q^{\max} - \varepsilon, q^{\max})$ we have

$$\frac{s(q)}{s(q^{\max} - \varepsilon)} = \exp\left\{-\int_{q^{\max}-\varepsilon}^{q} \frac{2f(\xi)}{\sigma^2(\xi)} d\xi\right\} < \exp\left\{-\int_{q^{\max}-\varepsilon}^{q} \frac{\bar{v}}{q^{\max} - \xi} d\xi\right\} = \left(\frac{q}{q^{\max} - \varepsilon}\right)^{-\bar{v}}.$$

Hence, for q "close" to $q^{\max}$ the function $s(q)$ (and accordingly the speed measure $S(q) = \int^q s(\eta)\, d\eta$) behaves as in Example 5 on page 221 in Karlin and Taylor (1981). (A similar argument applies to the boundary $q = q^{\min}$.) It then follows that the boundaries $q^{\min}$ and $q^{\max}$ are entrance boundaries whenever condition (57) holds and a stationary distribution exists. ■

Proof of Lemma 5. The fact that m_t is a (spanned) stochastic discount factor (SDF) means

$$d\log(m_t) = -r_t dt - \frac{\kappa_t^2}{2} dt - \kappa_t dB_t, \tag{64}$$

where κ_t is the market price of risk (the maximal Sharpe ratio). In the special case when preferences are specified by (27), and given the existence of annuities, optimality implies that the process $\log(m_t)$ satisfies[25]

$$d\log(m_t) = \beta\left(\gamma \log(c_t) - \log(\gamma V_t)\right) dt - \rho dt + d\log(\gamma V_t) - d\log(c_t). \tag{65}$$

An agent's value function V is homogeneous of degree γ in the her total wealth $\hat{W}_t$, which is the sum of her financial wealth and the present value of her future earnings. We consequently write

$$V_t(\hat{W}_t) = \frac{\hat{W}_t^{\gamma}}{\gamma} e^{\tilde{Z}_t} \tag{66}$$

for an appropriate process $\tilde{Z}_t$. Furthermore, from the envelope condition we have

$$\frac{\gamma}{\hat{W}_t} V_t = \frac{\partial V_t}{\partial \hat{W}_t} = f_c = \frac{\beta \gamma V_t}{c_t},$$

giving $c_t = \beta \hat{W}_t$.

For any $s < t$, the definition of V_t implies

$$V_t + \int_s^t \beta\gamma V_u \left(\log(c_u) - \frac{1}{\gamma}\log(\gamma V_u)\right) du = E_t \int_s^{\infty} \beta\gamma V_u \left(\log(c_u) - \frac{1}{\gamma}\log(\gamma V_u)\right) du.$$

Since the right-hand side is a martingale, the drift of the left-hand side equals zero, implying that $dV_t + \left(\beta\gamma V_t \left(\log(c_t) - \gamma^{-1}\log(\gamma V_t)\right)\right) dt$ is a martingale increment and therefore, upon applying Ito's Lemma we obtain

$$d\log(\gamma V_t) = -\beta\gamma\left(\log(c_t) - \gamma^{-1}\log(\gamma V_t)\right) dt - \frac{1}{2}\frac{\sigma_{V,t}^2}{V_t^2} dt + \frac{\sigma_{V,t}}{V_t} dB_t, \tag{67}$$

[25]See Duffie and Epstein (1992) for details.

where $\sigma_{V,t}$ denotes the instantaneous volatility of V_t at time t. Combining (67) and (65), we obtain

$$d\log(m_t) = -\beta dt - \frac{1}{2}\frac{\sigma_{V,t}^2}{V_t^2}dt - d\log(c_t) + \frac{\sigma_{V,t}}{V_t}dB_t.$$

Comparison with (64), along with the fact that aggregate consumption growth is deterministic and as a result individual consumption growth needs to be locally deterministic implies

$$\frac{\sigma_{V,t}}{V_t} = -\kappa_t \tag{68}$$

$$\dot{c}_t = (r_t - \rho)c_t. \tag{69}$$

Consumption c_t is therefore locally deterministic, and so is $\hat{W}_t = \beta^{-1}c_t$, which, upon using equation (66), leads to

$$\frac{\sigma_{V,t}}{V_t} = \sigma_{\tilde{Z}_t} = -\kappa_t.$$

Combining $c_t = \beta\hat{W}_t$ and (66) leads to

$$d\log(\gamma V_t) = d\tilde{Z}_t + \gamma d\log c_t = -\beta\left(\gamma\log\beta - \tilde{Z}_t\right)dt - \frac{1}{2}\frac{\sigma_{V,t}^2}{V_t^2}dt + \frac{\sigma_{V,t}}{V_t}dB_t, \tag{70}$$

where the second equality follows from (67). Letting $Z_t \equiv \tilde{Z}_t - \gamma\log(\beta)$ and noting that $\frac{\sigma_{V,t}}{V_t} = \sigma_{\tilde{Z},t} = \sigma_{Z,t}$ leads to

$$dZ_t = -\gamma d\log c_t + \beta Z_t dt - \frac{1}{2}\sigma_{Z,t}^2 dt + \sigma_{Z,t}dB_t. \tag{71}$$

Integrating (71), and noting that σ_Z is bounded, gives equation (29).

■

Proof of Lemma 6. Since the right hand side of (35) depends only on Z_t, it is immediate that strict monotonicity is equivalent to invertibility. Fixing Z_t and therefore ν_t, $\frac{\partial\varphi(q_t,Z_t)}{\partial q_t} < 0$ implies that there is a unique $q_t = \varphi^{-1}(\varphi_t, \nu_t)$. ■

Proof of Proposition 2. The proof of the proposition follows the same logic as that of Proposition 1. In the interest of completeness, we start by invoking Ito's Lemma to write down the SDE for ν:

$$\begin{aligned} d\nu_t = {} & \nu'\left(\nu^{-1}(\nu_t)\right)\left(f_Z\left(\nu^{-1}(\nu_t)\right) + \frac{\sigma_Z^2\left(\nu^{-1}(\nu_t)\right)}{2}\frac{\nu''\left(\nu^{-1}(\nu_t)\right)}{\nu'\left(\nu^{-1}(\nu_t)\right)}\right)dt \\ & + \nu'\left(\nu^{-1}(\nu_t)\right)\sigma_Z\left(\nu^{-1}(\nu_t)\right)dB_t \end{aligned} \tag{72}$$

Similiarly, one can write the dynamics of

$$\varphi_t = \varphi(q_t, Z_t) \tag{73}$$

based on the dynamics of q_t and Z_t, and then plug in $q_t = \varphi^{-1}(\varphi_t, \nu_t)$ and $Z_t = \nu^{-1}(\nu_t)$.

The existence of the inverse functions ν^{-1} and φ^{-1} is ensured by Lemma 6. To avoid repetition, we only justify two key statements in the text, namely equations (33) and (34).

As before, the definition of q_t implies that

$$m_t q_t D_{t,s} + \int_s^t m_t D_{t,s} = E_t \int_s^\infty m_u D_{u,s} du. \tag{74}$$

is a martingale. Using Ito's Lemma and $\kappa_t = -\sigma_{Z,t}$ yields equation (34).

From equations (71), (14), and the definition of ν_t, the drift of Z_t equals

$$\beta Z_t - \gamma \frac{\dot{c}_t}{c_t} - \frac{1}{2}\sigma_Z^2(Z_t) = \beta Z_t - \gamma(\lambda + \rho - \nu_t) - \frac{1}{2}\sigma_Z^2(Z_t), \tag{75}$$

which is equated to $f_Z(Z_t)$ to yield equation (33). ∎

Proof of Proposition 3. The modification of the share processes $\eta_t^d - g + g_t, \eta_t^l - g + g_t$ imply that $\frac{\dot{c}_{t,s}}{c_{t,s}}$ and $\frac{\dot{D}_{t,s}}{D_{t,s}}$ remain unaffected. Hence the real interest rate, the Sharpe ratio and the price-dividend ratio are identical in the two economies. ∎

Proof of Lemma 7. The first equality follows from $\eta_t^d = \frac{D_{t,t}}{D_t^A} = \frac{q_t D_{t,t}}{q_t D_t^A} = \frac{P_{t,t}}{P_t^A}$. The second equality follows upon time-differentiating $P_t^A = \int_{-\infty}^t P_{t,s} ds$ to obtain $dP_t^A = \int_{-\infty}^t d_t P_{t,s} ds + P_{t,t}$ and then dividing by P_t^A. ∎

B Details on the computation of marginal agent consumption growth

In this appendix, we provide more details on the construction of our empirical measure for marginal agent consumption growth dL_t and the discrepancy between marginal agent consumption growth and aggregate per capita consumption growth $d\log N_t - d\log(F_t)$. We also discuss issues related to measurement error and discrepancies between CEX and NIPA data, and provide a direct validation exercise for the part of the sample where we can estimate time effects in both a direct and indirect fashion.

For the construction of our measure of marginal consumption growth we need to choose: a) the measure of aggregate consumption C_t^A, b) how to estimate cohort and age effects in the data, c) the demographic table $\Lambda(t,s)$, d) how to extrapolate cohort effects that we can't measure back in time, and finally e) where to set the age cutoff at which agents enter the population and become marginal.

For choice a) For aggregate consumption growth $d\log C_t^A$ we use NIPA aggregate consumption expenditure of goods and services deflated by the respective PCE deflator (available from the Bureau of economic Analysis since 1929). To construct aggregate per capita consumption growth we subtract total population growth $d\log N_t$ (available from the Census) from aggregate consumption growth. Whether we use NIPA aggregate consumption of goods and services, total NIPA consumption expenditure or even the aggregate expenditure as measured by the CEX makes no difference to the calculation of the discrepancy between marginal agent and aggregate per capita consumption growth, which is the measure we use for our calibration. The reason is that aggregate consumption growth cancels out when we compute the difference $d\log(N_t) - d\log(F_t)$. (We revisit the choice of the appropriate notion of aggregate consumption when we discuss the issue of the choice of a cutoff age.) As a robustness check, we also re-computed our measure of the discrepancy between marginal agent and aggregate per capita consumption growth using the population growth of the adult rather than the general population (ages greater or equal to 20). We report the results in column A of Table 6. We note that using either the adult population growth in the US or the growth in the number of households in the US (available since 1947) gives similar results (correlation of 0.8 between the two measures of $d\log(N_t)-d\log(F_t)$) , but the measure of households in the data is volatile and contains a few abrupt year-over-year changes in some years, presumably due to data revisions. Besides, the demographic breakdown of households by age is sparse (households are binned in 10-year age-groups), and we need a more granular information on the demographic pyramid (see point c below).

For choice b), there are two sets of issues to address. First, whether to aggregate at the household level or the cohort level, and second, how to adjust for different cohort sizes. Concerning the first issue, we chose to use the public use micro data (PUMD) and the computer programs provided by the CEX to aggregate consumption expenditure for an entire cohort s in year t and then divide by the number of households in year t whose head of household was born in year s, using the weighting variable provided by the CEX. This gives us a set of 1,354 observation of $c_{t,s}$, which can then be regressed on time, age, and cohort dummies. The advantage of this approach is that the CEX code isolates expenditure within a given calendar year and performs a "months in scope" adjustment. Moreover, it becomes possible to compare results with the published CEX tables, which are not subject to top-coding and correspond to a larger sample than the Public Use Micro-Data.

An alternative approach to the estimation of time, age, and cohort effects is to aggregate expenditure at the level of a household (across all four interviews) and then regress that quantity on the appropriate dummies. The main disadvantage of this method is that, since households start and end their interviews at different times, part of a household's consumption will span two different years. As a robustness check, we used this alternative approach to the estimation of time, age, and cohort effects using the convention that year t corresponds to the calendar year of the earliest interview for each household. Using this convention, we re-estimated the cohort and age effects, and recomputed our measure of the difference between marginal and aggregate per-capita consumption growth. In performing this exercise, we combined the PUMD data with the CEX extracts on the NBER website (provided by John Sabelhaus), which reach back to the early eighties (we report the results in Table 6 column B). The correlation between our baseline measure of the discrepancy and this alternative measure of the discrepancy was quite high (0.92).

Our results also don't depend on how we adjust for household size. Robustness check C recomputes our baseline measure by estimating time, age ,and cohort effects but using only households with at least 2 members (to make sure that trends in single-member households don't affect our results). Robustness check D recomputes our baseline measure by aggregating consumption at the household level (as outlined in the above paragraph) and then including a control for log family size along with the dummies. The correlations with the baseline measure are above 0.9 in both cases.

For point c), we need a measure of the population of people at time t born in year s. These demographic tables are available from the Census going back to 1910 at annual frequency. Older

	Baseline	A	B	C	D	E	F	G	H
Correlation (1930-2016)	1.00	0.82	0.92	0.95	0.91	0.96	0.85	0.99	0.99
Correlation (1971-2016)	1.00	0.86	0.92	0.95	0.89	1.00	0.92	0.98	0.99
Standard Deviation (1930-2016)	0.74	0.97	0.70	0.77	0.64	0.61	0.68	0.79	0.73
Standard Deviation (1971-2016)	0.56	0.82	0.54	0.67	0.53	0.57	0.54	0.64	0.51

Table 6: The first two rows display the correlations between the measures of the discrepancy between marginal agent growth and per-capita consumption growth $d\log(N_t) - d\log(F_t)$ computed as in the baseline specification, respectively according to alternative specifications. The last two rows display the standard deviations of these measures, expressed in percent per annum.

ages are binned together past a cutoff for some census tables (typically past age 90). In these cases, we used the age-appropriate survival rates of the census closest to the calendar date (survival rates are available decennially) to split up the binned population, so as to make sure that for all years we have the population going to age 100. This procedure played no role in the end, because even if we simply truncate the sum F_t at age 90, we get (essentially) identical results.

Ideally, it would be best to use the age distribution of heads of households. Unfortunately, this time series is available only for 10-year binned age groups (25–34, 35–44, etc.) and only post 1960. We were able to confirm, however, that, the age distribution of heads-of-household largely mirrors the age distribution of adult persons when we bin the population of adult persons by 10-year age-groups and compare to the respective age distribution of heads of households (for the years where this information is available).

For choice d), in our baseline construction we used information on cohorts going back to 1907 only. (The earliest PUMD cross section is 1996 and age information is top-coded at 90, so 1907 is the earliest cohort we had information on). For cohorts prior to that we extrapolated the cohort estimates A_s by using linear extrapolation.[26] We performed three robustness checks. In column E of table 6, we used constant instead of linear extrapolation (i.e., we set $A_s = A_{1907}$ for $s < 1907$). In column F, we combined the PUMD data sets and the NBER CEX extracts (as outlined above), which allowed us to increase the number of cross sections back to 1980 and estimate cohort dummies back to 1891, with linear extrapolation prior to that. (We note here that cohort dummies prior to 1907 have quite large confidence bands, since these cohorts are not very well populated.) We also re-computed our measure, introducing an upper cutoff age of 75 when computing the moving sum in F_t (column G). Any of these modifications made little difference to our baseline measure for the entire sample. Especially in the post 1970 subsample, the three measures produce practically identical results, because the measure of people born prior to 1900 has become quite small by that time.

Before explaining how we make choice e), we note that this cut-off age is immaterial as long as the same cut-off is used in the computation of the integrals in equations (46) and (47), to compute C_t, respectively F_t. Figure 12 illustrates this point in the data: Starting with 1984, the CEX tables allow us to compute aggregate consumption for age groups above given cutoffs. The different lines in the figure depict dL_t by using equation (49) with different minimal age cutoffs (i.e., summing across age groups above 25, 35, etc., in the computation of both C_t and F_t). Since we are interested

[26]Specifically, for $s < 1907$, we set $A_s = A_{1907} + \chi(s - 1907)$ and used χ to correspond to the average value of ΔA_s for $s \in [1907, 1937]$.

Figure 12: Difference between the consumption growth rate of the marginal agent and aggregate consumption growth using different minimal age cutoffs. All series have been de-meaned. We depict 5-year moving averages to illustrate the low frequency movements of the series, which is the focus of our analysis.

in low frequency co-movements between the series, we illustrate the co-movements of 5-year moving averages, which are essentially the same no matter which cutoff is used.

For our baseline results we choose the minimal age cutoff to be 45. There are three reasons for this: First, for our purposes "birth" refers to the age at which an agent joins the financial market and her Euler equation starts holding. Accordingly, we want to ensure that agents have reached an age where borrowing constraints, which may invalidate equation (44) for younger cohorts, are likely to be irrelevant. We therefore aimed for a cut-off where the consumption-age profile starts reaching its peak. Second, immigration (which is prevalent in younger cohorts and acts similar to birth in our framework) is unlikely to be important past age 45. Third and most importantly, the aggregate consumption C_t of people in the included age groups should be a relatively stable fraction of aggregate NIPA consumption. Indeed, Figure 13 shows that a cut-off of 45 meets this goal. (The ratio for the 1972 cross section is essentially the same as for the 1984 cross section and all the subsequent cross sections). We note that such stationarity does not characterize the ratio of aggregate consumption in the CEX data to NIPA aggregate consumption, which is about 0.8 in the 1972 cross section and in the 1984 cross section, then exhibits an almost linear decline to around 0.6 until 2003 and fluctuates around that level thereafter.

The fact that the ratio of aggregate consumption in the CEX to aggregate consumption in NIPA is below one is not a particular concern for our purposes.[27] If the two series were in a (constant) proportion to each other, this would not impact the computation of $d \log C_t$. What is more disconcerting is the period between 1984 and 2003, when ratio of aggregate consumption according to CEX divided by aggregate consumption according to NIPA exhibits a downward trend.

[27]This discrepancy is most likely due to different definitions, weighting schemes, and under-reporting of certain consumption categories in the CEX

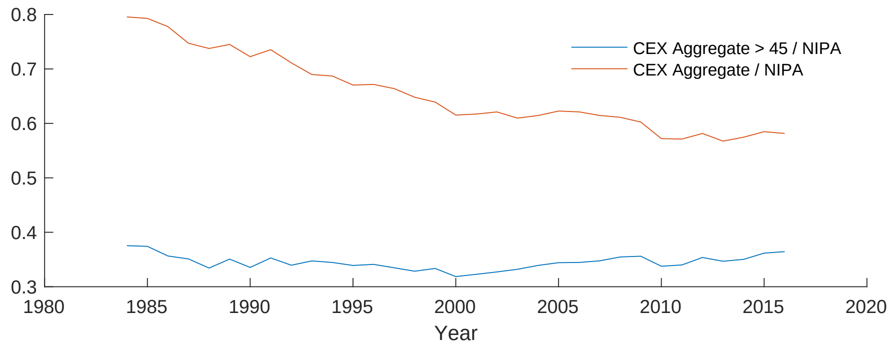

Figure 13: Ratio of aggregate consumption accruing to households 45 years or older according to the CEX divided by NIPA aggregate consumption expenditure. We also report the ratio of aggregate consumption according to the CEX divided by NIPA aggregate consumption expenditure.

While the consumption aggregate above our minimal age cutoff has remained a roughly constant fraction of aggregate NIPA consumption, as a robustness check in column H of Table 6, we reran our results using only the CEX cross sections after 2003 for the estimation of age and cohort effects that we use in the computation of F_t. For these cross sections, the ratio between CEX and NIPA aggregates has remained relatively stable. As column H of Table 6 shows, our results are unchanged.

In the context of measurement error, we would also like to note that the left plots of Figures 3 and 4 imply that the same patterns describe both the consumption and income of different cohorts. Income data are based on substantially more observations and are not subject to the same type of measurement errors as consumption surveys. Hence, whatever the source of measurement error in the CEX, it is unlikely to be impacting the measurement of age- and cohort- effects, which enters the measurement of F_t.

We conclude this appendix with a simple validation exercise that we can perform for a subsample. We compare our indirect inference approach for the estimation of the time-effects with a direct approach that can be performed by using data from the published CEX tables. These tables provide information on the average per-household consumption of 10-year binned age groups (25–34, 35–44, 45–54, 55–64, and 65–74). We isolate the cross sections 1972 (the earliest cross section), then the first cross section that is available thereafter (1984), and then cross sections every ten years (1994, 2004, 2014) so that there is approximately (for the first) and exactly (for all subsequent) a 10-year gap between the cross sections. Letting $t \in \{1972, 1984, 1994, 2004, 2014\}$ and i index the age-groups 25–34, 35–44, 45–54,55–64,65–74, we compute the expression

$$\log c_{t+10}^{i+1} - \log c_t^i - (\log C_{t+10} - \log C_t) - \frac{1}{4}\sum_t \left[\log c_{t+10}^{i+1} - \log c_t^i - (\log C_{t+10} - \log C_t)\right]. \quad (76)$$

If equation (44) holds, then the expression (76) should be equal to the de-meaned, decennial time effect $L_{t+10} - L_t$ minus the de-meaned, decennial aggregate consumption growth rate $\log C_{t+10} - \log C_t$.[28] Furthermore, according to equation (49), it also equals — up to measurement error —

[28]Intuitively, the term $\log c_{t+10}^{i+1} - \log c_t^i$ is the consumption growth rate for a fixed, 10-year binned cohort; by subtracting the term $\frac{1}{4}\sum_t \left[\log c_{t+10}^{i+1} - \log c_t^i\right]$ we remove the part of that consumption growth that can be attributed to the age-effect, so that the remainder is the de-meaned time effect.

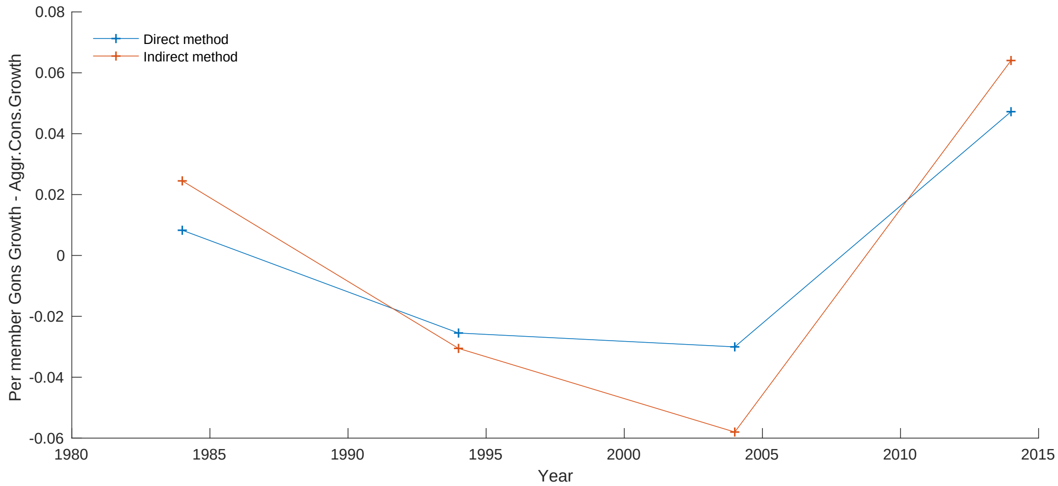

Figure 14: Comparison of direct method (i.e., by directly inferring the decennial time effects $L_{t+10} - L_t$ minus decennial aggregate consumption growth $\log C_t - \log C_t$) and indirect method of inferring the same concept with our measure of $\log(F_t) - \log F_{t+10}$. Each point in the graph corresponds to the difference in the two growth rates over the course of the preceding decade, with the exception of the first observation, which spans 12 years (1972-1984).

$\log(F_t) - \log(F_{t+10})$, consistent with Figure 14.

C Additional results

Figure 15 shows that the high correlation between real expected interest rates and marginal agent consumption growth is unlikely to be the result of sampling error. The figure shows the R^2 of the regression of the 1-, 2-, ...,12-year moving average of marginal consumption growth on the 1-, 2-, ...,12-year moving average of the real expected interest rate. The solid line refers to the data, the dotted line refers to the 95% confidence bands obtained by constructing 10,000 artificial time series by drawing from the empirical distribution of marginal agent consumption growth, time-aggregating both the real interest rate and the marginal agent consumption growth for each sample to obtain 1-, 2-, ...,12-year moving averages of both series, computing the R^2 for each of those 10,000 samples, and reporting the top 95-th percentile of R^2. The R^2 obtained in the data for years larger than 2 is higher than the bootstrapped 95-th percentile of R^2, implying that the high correlation between the two series in the data is unlikely to be random.

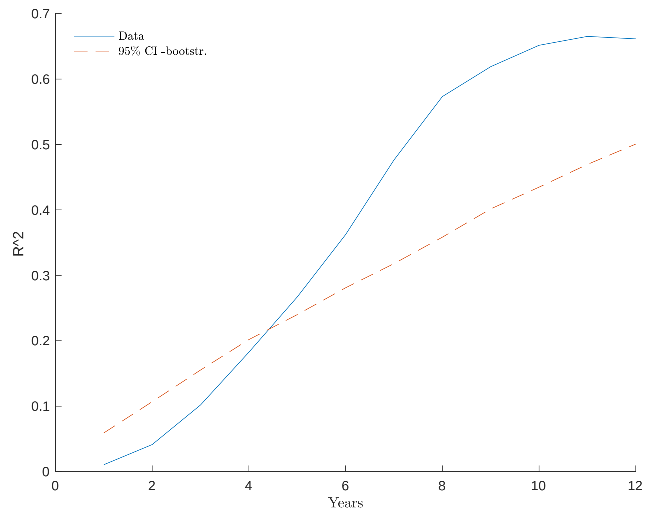

Figure 15: Top left plot: R^2 of Regression of 1-, 2-, ...,12-year moving average of marginal consumption growth on 1-, 2-, ...,12-year moving average of the real expected interest rate. The solid line refers to the data, the dotted line refers to the 95% confidence bands obtained by drawing 10,000 random time-series of marginal agent consumption growth with replacement from the data, time-aggregating both the real interest rate and the marginal agent consumption growth for each sample, computing the R^2 for each of those 10,000 samples, and reporting the top 95-th percentile of R^2.