Asset Prices with Wealth Dispersion *

Paul Ehling† Junjie Guo‡ Christian Heyerdahl-Larsen§
BI CUFE Indiana University

Current Draft: September 2021

*We thank participants at a brown bag seminar at Kelley for comments and suggestions. Paul Ehling thanks the Centre for Asset Pricing Research (CAPR) at BI for funding support. Junjie Guo acknowledges the financial support from the National Science Foundation of China (No. 72103216) and Program for Innovation Research in Central University of Finance and Economics.

†Department of Finance, BI Norwegian Business School, Nydalsveien 37, 0484 Oslo, paul.ehling@bi.no

‡School of Finance, Central University of Finance and Economics, Shahe Campus: Shahe Higher Education Park, Changping District, Beijing, 102206, junjguo@cufe.edu.cn

§Kelley School of Business, Indiana University, 1309 E. 10th Street, Bloomington, IN 47405, chheyer@indiana.edu

Abstract

With overlapping generations and heterogeneous risk aversion there is no unique relation between aggregate risk aversion and the real rate of interest, and this type of endogenous "noise" cannot arise in an economy where agents live forever. Our framework accommodates many agent types and the noise emerges precisely because all (but one) consumption shares drive the economy. Adding wealth dispersion to aggregate risk aversion sufficiently summarizes the rich dynamics of the model. Consistent with the model, we construct "level" and "slope" factors that do not require knowledge about agents' risk aversion to predict excess returns.

Keywords: Heterogeneous Risk Aversion, Overlapping Generations, Consumption Share Weighted Variance of the Risk-Tolerance

JEL Classification: E2, G10, G11, G12

1 Introduction

People are heterogeneous and risk premium or aggregate risk aversion vary over time. These two generally accepted observations raise the possibility that changes in the risk premium and changes in aggregate risk aversion are jointly determined through risk sharing by heterogeneous agents. Our point in revisiting this question is that an overlapping generations (OLG) model with a cross-section in risk aversion differs from the standard infinitely lived setting but also from an OLG model with two agents. In complete markets with infinitely lived agents only the level of aggregate consumption matters. In an OLG model with two agents, one has to only follow the consumption share of one of the two agents. In our framework with many types of agents in an OLG setting, one has to essentially follow all consumption shares, even though markets are complete and there is only one exogenous shock. Since aggregate risk aversion depends on the entire consumption share distribution it can take many values for a specific value of aggregate consumption. This property is inherited by all other asset pricing quantities such as the real rates of interest or the risk premium.

How does heterogeneity in risk aversion matter in such an OLG setting? We show that even though the entire consumption share distribution drive asset prices, most of the variation can be explained by aggregate risk aversion and the consumption share weighted variance of risk tolerance. This contrasts with the infinitely lived economy where only the aggregate risk aversion matters and has important implications for the equilibrium real rate of return and risk premium. Specifically, we show that the consumption share weighted variance of risk tolerance effectively amplifies the precautionary savings term, i.e., in times of high cross-sectional variance of risk aversion the real rate of return is low. In addition, a higher cross-sectional variance is associated with a higher stock market risk premium.

Since most of the variation in the asset pricing moments are explained by aggregate risk aversion and the consumption share weighted variance of risk tolerance it is important to understand the dynamics of the consumption shares. We show that the consumption shares of the agents with the highest or lowest risk aversion in the economy are mostly driven by the aggregate risk aversion in the economy, just like in an economy with infinitely lived agents or a two agent economy.[1] For agents in the middle of the risk aversion distribution the consumption share weighted variance of risk tolerance is more important than aggregate risk aversion. We exploit this heterogeneity of the loadings on risk aversion and variance of the risk tolerance to construct an empirical measure that does not require the knowledge about the risk aversion distribution. Specifically, we construct a "level" and "slope" factor that capture the dynamics of our model just as the aggregate risk aversion and the consumption share weighted variance of the risk-tolerance. With two agent types or infinitely lived agents the slope factor is spanned by the aggregate risk aversion and therefore it does not impact the economy. Consistent with the model, our proxy in the data for the level and slope factor are the log wealth share of Forbes 400 (the top 400 richest people in Forbes list) and the wealth share of Forbes 400 plus the wealth share of the top 10% group minus the wealth share of the top 0.1%, respectively. In bivariate predictive regression, we then regress the excess return on level and slope. As in the model, the level produces a negative sign whereas the slope produces a positive sign, where a higher wealth share of the top 0.1% indicates goods times and low aggregate risk aversion.

On a technical note, the usual approach to solve an economy with heterogeneous agents in complete markets is to aggregate them using a central planner with fixed Pareto weights.[2] In our continuous time OLG model, the Pareto weight of a type depends on the state of nature at birth and are determined as a part of the equilibrium. Since generally there is no closed-form solution for the Pareto weights, we solve the model by making the following assumption. Namely, that the consumption share at birth of each type is time invariant.[3] We construct such an equilibrium by assuming a redistributive shock to the endowment stream

[1] The cross-sectional variance also impact the consumption shares of the agents on the boundary of the risk aversion distribution, but relative to the variation driven by the aggregate risk aversion it is small.

[2] In the economics literature, it is standard to solve heterogeneous agent models with aggregate risk and idiosyncratic shocks by following agents' consumption or wealth share. See, for instance, Krusell and Smith (1998).

[3] See Ehling, Guo, Heyerdahl-Larsen, and Illeditsch (2021) for details.

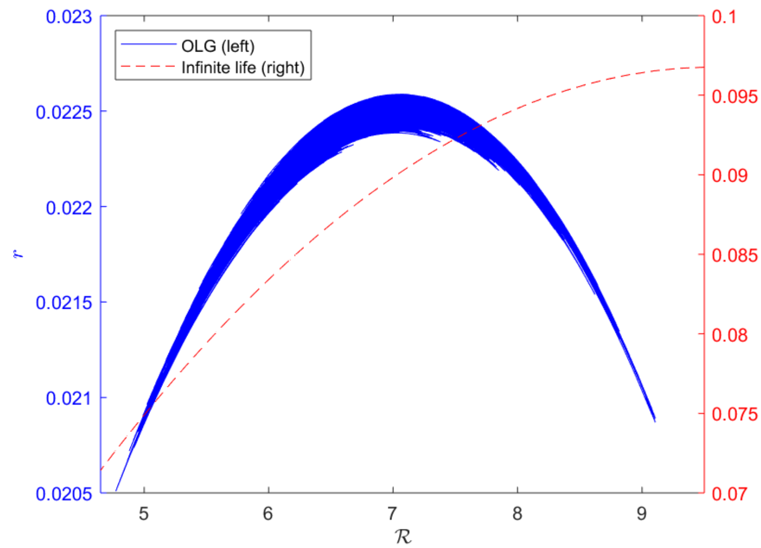
of all the agents in the economy. As this shock is the same shock as in aggregate output, there is no additional risk premium associated with it. A consequence of this is that consumption shares can be simulated forward. This then allows to use Monte-Carlo simulations, and in combination with Malliavin calculus,[4] we characterize the dynamics of stock prices and portfolios analytically. Thereby, we can circumvent the curse of dimensionality, and in principle have even hundreds of agent types, and still follow their consumption shares at a low computational cost. Without the time invariance, one has to solve for the optimal consumption-to-wealth ratio and the initial wealth, leading to a system of partial differential equations (PDEs). Even for a small number of types, this becomes quickly infeasible to solve.

In an OLG model such as ours, the per capita consumption of newborns relative to average consumption impacts the real rate of interest, which was first pointed out by Garleanu and Panageas (2015). Specifically, when the average consumption of the newborns are higher than the average consumption growth of the agents already in the economy this implies that their consumption growth has to be lower than the aggregate output growth. As the agents already present in the economy determine the interest rate it then declines due to their intertemporal smoothing motive. The displacement shock in our economy allows us to match the real rate of interest even for high levels of risk aversion. From Figure 1, we see that in our stationary economy the real short rate cannot move much beyond 2.25 percent while in an otherwise identical non-stationary infinite life economy the real short rate can reach levels above of 9.95 percent and this within the stationary range of the OLG model for aggregate risk aversion. In contrast to the infinitely lived economy, we can see from Figure 1 that there are many possible values for the real short rate (r) for a given value of aggregate risk aversion ($\mathcal{R}$). As we discussed above, this difference is mainly captured by the consumption share weighted variance of the risk tolerance in the economy.

Our paper builds on the broad literature studying asset prices and portfolio policies in

[4] See Nualart (1995) for an introduction to Malliavin calculus.

Figure 1: **Real Short Rate.** The figure plots the real short rate, r, in our OLG model (solid blue area) and in the standard model with infinitely lived agents (dashed red line) as a function of the aggregate risk aversion, $\mathcal{R}$. In both economies there are 10 agent types, where we set the lowest risk aversion to the 5th percentile (1.9) and the highest risk aversion to the 95 percentile (19.5) based on Kimball, Sahm, and Shapiro (2008). Further details are in Subsection 3.1.



infinite life economies populated by agents who differ in their level of risk aversion. We mention the classical papers by Dumas (1989), using two investors, Wang (1996), using two specific risk aversions that lead to closed form solutions, and Chan and Kogan (2002), using heterogeneous risk aversion with homogeneous habits.[5]

Our paper mainly builds on Garleanu and Panageas (2015), who also study a continuous-time overlapping generations economy.[6] Their focus is on the quantitative implications of heterogeneity with recursive preferences.[7] Because they have two classes of agent types, they

[5]Other more recent contributions to the literature on asset pricing with heterogeneous risk aversion with infinite life include Malamud (2008a), Malamud (2008b), Yan (2008), Zapatero and Xiouros (2010), Cvitanic and Malamud (2011), Cvitanic, Jouini, Malamud, and Napp (2012), Longstaff and Wang (2013), Bhamra and Uppal (2014), Ehling and Heyerdahl-Larsen (2017), Santos and Veronesi (2018) and Veronesi (2019), among many others.

[6]See Ehling, Graniero, and Heyerdahl-Larsen (2018), Garleanu and Panageas (2020a), Garleanu and Panageas (2020b) and Heyerdahl-Larsen and Illeditsch (2020) for related continuous-time overlapping generations economies that feature agents with logarithmic preferences.

[7]However, a working paper version (Garleanu and Panageas (2007)) of their model features heterogeneous risk aversion in a power utility framework similar to ours. It also produces comparable unconditional asset pricing moments. Another recent contribution to asset pricing theory using power preferences is Martin (2013).

can express the equilibrium problem as a system of ordinary differential equations (ODEs) but already with three levels for risk aversion one would have to solve a system of PDEs. In our examples, we use ten levels of risk aversion and stress that we could solve the model with hundreds of agent types.

Finally, we mention Gomez (2019), who studies the implications of heterogeneous risk aversion in a continuous-time overlapping generations economy for wealth inequality. He also points out that wealth inequality has predictive power for asset prices.

2 The Model

We consider a continuous-time exchange economy with overlapping generations in the style of Blanchard (1985). Agents consume a single perishable consumption good and, once born, they trade in a complete set of securities to share consumption and longevity risks. The agents are heterogeneous since they differ not only in when they are born and die but also in the curvature of their utility functions.

2.1 Populations, Earnings, and Output

To have a constant total population size, which we normalize to 1, agents die and are replaced at the same rate $\nu > 0$. Hence, at time t, the size of the cohort born at time $s < t$ is simply $\nu e^{-\nu(t-s)} ds$.

The agents receive an endowment of earnings $y^i_{s,t}$, where $y^i_{s,t} = \omega Y_t G_{s,t}$ for $\omega \in (0, 1)$ and i identifies the agent type in terms of preferences. The function G_s differs for every cohort and controls the life-cycle profile of earnings. We normalize the $G_{s,t}$ functions such that $\int_{-\infty}^{t} \nu e^{-\nu(t-s)} G_{s,t} ds = 1$. Here, the variation of $G_{s,t}$ over time can be thought of as a form of intergenerational displacement shock.

Besides that, there also exists a representative firm paying out $D_t = (1-\omega)Y_t$ in dividends. Thus, aggregate output is given by $\int_{-\infty}^{t} \nu e^{-\nu(t-s)} \omega Y_t G_{s,t} ds + D_t = Y_t$ and it evolves over time as follows

$$dY_t/Y_t = \mu_Y dt + \sigma_Y dz_t, \tag{1}$$

where z_t is a standard Brownian motion.

2.2 Security Markets and Prices

Agents trade an instantaneously risk-free asset, shares in the representative firm, and annuities that allow hedging mortality risk in frictionless markets. The instantaneously risk-free asset is in zero net supply and it evolves over time with dynamics given by

$$dB_t/B_t = r_t dt, \tag{2}$$

where r_t denotes the real short rate to be determined in equilibrium.

We normalize the supply of shares in the representative firm to one. It trades at a price of P_t, where the return process, R_t, evolves according to

$$dR_t = (dP_t + D_t dt)/P_t = \mu_{R,t} dt + \sigma_{R,t} dz_t, \tag{3}$$

where $\mu_{R,t}$ and $\sigma_{R,t}$ are to be determined in equilibrium.

As in Yaari (1965), an annuity contract generates an income stream of $\nu W^i_{s,t}$ per unit of time, where $W^i_{s,t}$ is the financial wealth at time t of an agent with risk aversion i born at time s. The annuities are supplied by a competitive insurance industry and all agents purchase an annuity to insure against longevity risk. To break-even, the insurance industry receives the financial wealth of the deceased agents.

It is convenient to summarize the price system in terms of a stochastic discount factor. Its dynamics follow

$$d\xi_t/\xi_t = -r_t dt - \theta_t dz_t, \tag{4}$$

where θ_t is the price of risk.

2.3 Preferences and the Dynamics of Wealth

We assume that each generation born at time s consists of a continuum of agents differing from each other with respect to the curvature of their utility functions, which is given by

$$U^i\left(c^i\right) = \mathbb{E}_s\left[\int_s^{\zeta} e^{-\rho(t-s)}\frac{\left(c^i_{s,t}\right)^{1-\gamma_i}}{1-\gamma_i}dt\right], \tag{5}$$

where $\mathbb{E}$ denotes the expectation operator, ζ is the stochastic time of death, ρ is the subjective time preference rate, $c_{s,t}$ is the instantaneous consumption of generation s at time t, γ_i is the coefficient of relative risk aversion of type i with finite support over $i \in \left[\underline{\gamma}, \overline{\gamma}\right]$ with $0 < \underline{\gamma} <= \overline{\gamma} < \infty$.

The dynamics of financial wealth, $W^i_{s,t}$, of an agent of type i born at time s entitled to the endowment stream $y^i_{s,t}$, is

$$dW^i_{s,t} = \left(r_t W^i_{s,t} + \pi^i_{s,t}\left(\mu_{R,t} - r_t\right) + \nu W^i_{s,t} + y^i_{s,t} - c^i_{s,t}\right)dt + \pi^i_{s,t}\sigma_{R,t}dz_t, \tag{6}$$

where $\pi^i_{s,t}$ denotes the dollar amount invested in the risky security. As agents are born without financial wealth, we have that $W^i_{s,s} = 0$.

2.4 Heterogeneity within a Generation

Agents are heterogeneous in their risk aversion and in each generation a fraction α^i_t exhibits a coefficient of relative risk aversion of γ_i. Assuming that there is a continuum of agents of each type, we have

$$\int_i \alpha^i_t di = 1. \tag{7}$$

Specifically, we allow there to be a continuum of types as well as a discrete number of types. In case we consider a discrete number of types, the integral in Equation (7) is replaced by a sum. Further, we allow α^i_t to be stochastic and, therefore, the distribution of agent types can change from period to period. When α^i_t is stochastic, we assume that it is adapted to

z_t.

2.5 Infinitely Lived Agents

It is useful to compare our economy with finitely lived agents with the more standard infinitely lived agents case. To this end, we consider the limit case with $\nu = 0$. Specifically, we know from the welfare theorems that the marginal utilities of the agents are equated with constant Pareto weights, i.e., $\lambda^i \left(c_t^i\right)^{-\gamma_i} = \lambda^j \left(c_t^j\right)^{-\gamma_j}$. Imposing market clearing, $\int_i c_t^i di = Y_t$, implies the following standard equilibrium results.

Proposition 1. *Consider an economy with infinitely lived agents, then the optimal consumption of an agent of type i is only a function of aggregate output at that point in time:*

$$c_t^i = c^i \left(Y_t\right). \tag{8}$$

Further, aggregate risk aversion and prudence are only a function of aggregate output at that point in time:

$$\mathcal{R}_t = \mathcal{R}\left(Y_t\right), \tag{9}$$

$$\mathcal{P}_t = \mathcal{P}\left(Y_t\right), \tag{10}$$

where $\mathcal{R}$ *is monotonic with following limits:* $\lim_{Y\to\infty}\left[\mathcal{R}_t\right] = \underline{\gamma}$ *and* $\lim_{Y\to 0}\left[\mathcal{R}_t\right] = \overline{\gamma}$. *Importantly, aggregate risk aversion can be inverted to serve as state variable, that is,*

$$Y_t = \left(\mathcal{R}_t\right)^{-1}, \tag{11}$$

and aggregate prudence can be expressed as a function of aggregate risk aversion

$$\mathcal{P}_t = \mathcal{P}\left(\mathcal{R}_t\right). \tag{12}$$

In applied work, see for instance the seminal work of Dumas (1989), typically the Pareto weights are chosen directly and this implies the initial wealth distribution. However, even for a given initial wealth distribution one can solve a fixed point problem to find a set of corresponding Pareto weights. A consequence of the above is that the consumption share distribution does not serve as a state variable since aggregate output is sufficient. As we show below, in the OLG economy the aggregate output is not sufficient to derive the consumption share distribution anymore.

2.6 Equilibrium

Since after birth, the market is dynamically complete for each single agent, we solve the individual optimization problems by martingale methods as in Cox and Huang (1989). Because death is independent of output and exponentially distributed, we integrate it out to write an agent's static lifetime maximization problem as

$$\max \mathbb{E}_s \left[\int_s^\infty e^{-(\rho+\nu)(t-s)} \frac{\left(c_{s,t}^i\right)^{1-\gamma_i}}{1-\gamma_i} dt \right], \tag{13}$$

s.t.

$$\mathbb{E}_s \left[\int_s^\infty e^{-\nu(t-s)} \frac{\xi_t}{\xi_s} c_{s,t}^i dt \right] = \mathbb{E}_s \left[\int_s^\infty e^{-\nu(t-s)} \frac{\xi_t}{\xi_s} y_{s,t}^i dt \right] = H_{s,s}^i, \tag{14}$$

where $H_{s,s}^i$ is the endogenous present value of lifetime earnings for an agent born at time s with risk aversion γ_i.

At time s, the first order condition (FOC) from the martingale solution for all agents born at time s implies

$$c_{s,s}^i = \left[\kappa_{i,s}\right]^{-1/\gamma_i}, \tag{15}$$

whereas at time t, the FOC for all agents born at time s implies

$$c_{s,t}^i = \left[e^{\rho(t-s)} \frac{\xi_t}{\xi_s} \kappa_{i,s} \right]^{-1/\gamma_i}, \tag{16}$$

where $\kappa_{i,s}$ is the Lagrange multiplier of the optimization problem of an agent i born at time s. The Lagrange multiplier relates to the initial consumption through $c^i_{s,s} = \kappa_{i,s}^{-1/\gamma_i}$.

In equilibrium, markets must clear and, thus, after integrating over the times of birth and over agent types, we have that

$$\int_{-\infty}^{t} \left(\int_i c^i_{s,t} \alpha^i_s di \right) \nu e^{-\nu(t-s)} ds = Y_t, \tag{17}$$

formally stating that the sum of all individuals' consumption equals aggregate output.

To help manipulate the market clearing condition, we define the aggregate consumption of cohort s, $C_{s,t}$, and the aggregate consumption of agent type i, C^i_t, as follows

$$C_{s,t} = \int_i c^i_{s,t} \alpha^i_s di, \tag{18}$$

$$C^i_t = \int_{-\infty}^{t} c^i_{s,t} \alpha^i_s \nu e^{-\nu(t-s)} ds, \tag{19}$$

$$f^i_t = \frac{C^i_t}{Y_t}, \tag{20}$$

where f^i_t denotes the share of consumption of agent type i in aggregate consumption at time t.

Definition 1. *Let $\tau_i = \frac{1}{\gamma_i}$ denote the risk tolerance of agents of type i. We define the consumption share weighted moments of order k as*

$$\mathcal{E}_t \left(\tau^k \right) = \int_i f^i_t \tau^k_i di. \tag{21}$$

The relative risk aversion is given by

$$\mathcal{R}_t = \mathcal{E}_t \left(\tau \right)^{-1}, \tag{22}$$

and the relative prudence $\mathcal{P}_t$ is given by

$$\mathcal{P}_t = \left(\frac{\mathcal{E}_t\left(\tau^2\right)}{\mathcal{E}_t\left(\tau\right)^2} + \mathcal{R}_t \right). \tag{23}$$

Aggregate risk aversion in heterogeneous risk aversion models is bounded between the risk aversion of the individual agents. This, however, is not the case for aggregate prudence as discussed for instance in Wang (1996) and Bhamra and Uppal (2014).

The next corollary provides a novel characterization of prudence.

Corollary 1. *Aggregate prudence in Equation (23) can be rewritten as*

$$\mathcal{P}_t = (1 + \mathcal{R}_t) + \frac{\mathcal{V}_t\left(\tau\right)}{\mathcal{E}_t\left(\tau\right)^2}, \tag{24}$$

where $\mathcal{V}_t\left(\tau\right) = \mathcal{E}_t\left(\tau^2\right) - \mathcal{E}_t\left(\tau\right)^2$ is the consumption share weighted cross-sectional variance of the risk tolerance.

Hence, the prudence is higher in the heterogeneous risk aversion economy as long as the consumption share distribution is not concentrated on one agent type. The characterization in Corollary 1 is general but we emphasize that only in an OLG framework with more than two types $\mathcal{R}$ and $\mathcal{P}$ can move independently.

Next, consider the individual consumption, $c^i_{s,t}$, of an agent of type i born at time s. Given the optimal consumption in Equation (16) and the dynamics of the stochastic discount factor in Equation (4), we have the following.

Proposition 2. *Consider the consumption of an agent of type i born at time s. The dynamics of the consumption follows*

$$dc^i_{s,t} = \mu_{c^i,t} c^i_{s,t} dt + \sigma_{c^i,t} c^i_{s,t} dz_t, \tag{25}$$

where

$$\mu_{c^i,t} = \frac{\mathcal{R}_t}{\gamma^i}\left(\mu_Y + \nu\left(1-\beta_t\right)\right) + \frac{1}{2}\frac{\mathcal{R}_t}{\gamma^i}\left(\frac{\mathcal{R}_t}{\gamma^i} - \frac{\mathcal{E}_t\left(\tau^2\right)}{\mathcal{E}_t\left(\tau\right)^2}\right)\sigma_Y^2, \tag{26}$$

and

$$\sigma_{c_s^i,t} = \frac{\mathcal{R}_t}{\gamma^i}\sigma_Y. \tag{27}$$

According to the proposition, the drift and diffusion coefficients of the individual consumption growth do not depend on the date of birth. Hence, all agents of same type have the same consumption dynamics. Moreover, from the consumption dynamics in Proposition 2, we can see that the volatility of the consumption is decreasing in the risk aversion. In addition, the expected consumption growth depends on two components. The first part, $\frac{\mathcal{R}_t}{\gamma^i}\left(\mu_Y + \nu\left(1-\beta_t\right)\right)$, shows that an agent with high risk aversion relative to the consumption share weighted harmonic mean, $\mathcal{R}_t$, has a lower expected consumption growth. Of course, in this case, it is actually not the risk aversion per se, but instead the elasticity of intertemporal substitution that drives this result. We stress that the expected consumption growth is that of the agents currently alive and, therefore, it will be different from the aggregate endowment growth unless newborns consume all their endowment.

As all agents of the same type have the same dynamics for their consumption, the total consumption of all agents of a specific type only differ from the individual dynamics by the birth and death of agents of that type. This is illustrated in the next corollary.

Corollary 2. *The dynamics of the total consumption of agents of type i is*

$$dC_t^i = \left(\nu\left(\alpha_t^i c_{t,t}^i - C_t^i\right) + \mu_{c^i,t}C_t^i\right)dt + \sigma_{c_s^i,t}C_t^i dz_t. \tag{28}$$

Computing the dynamics of the stochastic discount factor requires the consumption shares of each agent type, $f_t^i = \frac{C_t^i}{Y_t}$. Thus, the next proposition characterizes the dynamics of the consumption shares.

Proposition 3. *The dynamics of consumption share i is*

$$df_t^i = \mu_{f^i,t}dt + \sigma_{f^i,t}dz_t, \tag{29}$$

where

$$\begin{aligned}\mu_{f^i,t} &= \nu\left(\alpha_t^i\beta_t^i - f_t^i\right) \\ &+ f_t^i\left(\left(\frac{\mathcal{R}_t}{\gamma^i} - 1\right)\mu_Y + \nu\frac{\mathcal{R}_t}{\gamma^i}\left(1-\beta_t\right) + \left(1 - \frac{1}{2}\frac{\mathcal{R}_t}{\gamma^i}\mathcal{P}_t + \frac{1}{2}\left(1+\frac{1}{\gamma^i}\right)\frac{\mathcal{R}_t^2}{\gamma^i} - \frac{\mathcal{R}_t}{\gamma^i}\right)\sigma_Y^2\right),\end{aligned} \tag{30}$$

and

$$\sigma_{f^i,t} = f_t^i\left(\frac{\mathcal{R}_t}{\gamma^i} - 1\right)\sigma_Y, \tag{31}$$

and where $\beta_t^i = \frac{c_{t,t}^i}{Y_t}$.

Inserting the FOC at time t into the market clearing at t yields

$$\int_{-\infty}^{t}\left(\int_i\left[e^{\rho(t-s)}\frac{\xi_t}{\xi_s}\kappa_{i,s}\right]^{-1/\gamma_i}\alpha_{s,t}^i di\right)\nu e^{-\nu(t-s)}ds = Y_t, \tag{32}$$

which we rewrite as

$$\int_{-\infty}^{t}\int_i \nu e^{-\nu(t-s)}c_s^i e^{-\rho(t-s)/\gamma_i}\left[\frac{\xi_t}{\xi_s}\right]^{-1/\gamma_i} dids = Y_t, \tag{33}$$

where $c_s^i = \left[\kappa_{i,s}\right]^{-1/\gamma_i}\alpha_s^i$ is the initial consumption of agents' of type i at time s, i.e., at the time of entering the economy. An application of Ito's lemma to both sides of Equation (33) and equating the drift and diffusion terms yields the dynamics of the stochastic discount factor.

Proposition 4. *The real short rate is*

$$r_t = \rho + \mathcal{R}_t\mu_Y - \frac{1}{2}\mathcal{R}_t\mathcal{P}_t\sigma_Y^2 + \mathcal{R}_t\nu\left(1-\beta_t\right), \tag{34}$$

where

$$\beta_t = \frac{C_{t,t}}{Y_t}. \tag{35}$$

The market price of risk is

$$\theta_{Y,t} = \mathcal{R}_t \sigma_Y. \tag{36}$$

From Proposition 4, we see that the market price of risk takes the standard form in economies with heterogeneous risk aversion as discussed in Ehling and Heyerdahl-Larsen (2017). The risk-free rate also has a similar form as in an economy with heterogeneous preferences and infinite life except for the last term, $\mathcal{R}_t \nu (1 - \beta_t)$. This term is related to the birth and death of agents in the economy. From Equation (35), we see that β_t is the ratio of the consumption of the newborn agents to the consumption of the total population. Hence, when $\beta_t > 1$ newborns consume on average more than the rest of the population. Whenever the average consumption of the newborns differs from aggregate consumption, then the interest rate deviates from the interest rate in an infinitely lived economy. Specifically, when $\beta_t > 1$, the average consumption of the newborns are higher than the total, implying that the average consumption growth of the agents already in the economy has to be lower than the aggregate output growth, μ_Y. As the interest rate is determined by the agents already present in the economy and not the newborns, the consumption growth is lower and, therefore, the interest is also lower due to the intertemporal smoothing motive as discussed in Garleanu and Panageas (2015).

Using the consumption share weighted cross-sectional variance of the risk tolerance, the next lemma relates the risk-free rate in the homogeneous preferences case to the heterogeneous preference case.

Lemma 1. *Consider a homogeneous agent economy with a coefficient of relative risk aversion of* γ*. Let the real risk-free rate in this economy be* $r_t^H = \rho + \gamma\mu_Y - \frac{1}{2}\gamma(1+\gamma)\sigma_Y^2 + \gamma\nu\left(1 - \beta_t^H\right)$*, where* β^H *is the average consumption of the newborns relative to the total consumption in this economy. Consider a heterogeneous agent economy at time* t *with* $\mathcal{R}_t = \gamma$ *and real risk* *free rate of r_t, then the difference between the real rates in the two economies is*

$$r_t - r_t^H = \gamma \left(\beta_t - \beta_t^H \right) - \frac{1}{2} \gamma \frac{\mathcal{V}_t (\tau)}{\mathcal{E}_t (\tau)^2} \sigma_Y^2. \tag{37}$$

Moreover, the market prices of risk are the same in the two economies.

From Lemma 1, we see that the difference in the real rates is due to the overlapping generations structure and the prudence. Importantly, the real rate in the heterogeneous agent economy is as if there is a representative agent that has desire to increase the precautionary savings when the consumption share weighted cross-sectional variance goes up. Hence, for a fixed level of relative risk aversion, the interest rate is decreasing in the cross-sectional variance of the risk tolerance.

2.7 Curse of Dimensionality

In Proposition 3, the dynamics depends on β_t^i which is determined through the initial consumption of each agent when born. However, to solve for the initial consumption one has to solve for the optimal consumption-to-wealth ratio and the initial wealth. To this end, consider the consumption-to-wealth ratio of an agent born at time s of type i

$$\phi_{s,t}^i = \frac{c_{s,t}^i}{\tilde{W}_{s,t}^i}, \tag{38}$$

where $\tilde{W}_{s,t}^i = H_{s,t} + W_{s,t}^i$ denotes the total wealth of the agent. One can show that the consumption-to-wealth ratio is independent of the time of birth, that is $\phi_{s,t}^i = \phi_t^i$ for all $s \leq t$. Hence, the initial consumption of an agent born at time s of type i can be written as

$$c_{t,t}^i = \phi_t^i H_{t,t}. \tag{39}$$

One could write down the pricing equation and express the problem as a system of partial differential equations (PDEs). However, this depends on the entire consumption share distribution, so an economy with N different agent types $i = (1, \ldots, N)$ leads to a system of PDEs with $N-1$ independent variables. Even for small N, this becomes quickly infeasible to solve. This is specific to the overlapping generations (OLG) setting as there is a subtle form of incompleteness in the OLG economy. Agents not born yet cannot hedge against being born in a bad time. Therefore, the constant Pareto weight representation from an economy with infinitely lived agents is not feasible in general. The Pareto weights in the OLG economy are related to the Lagrange multipliers κ_s^i, which themselves depend on the state of the economy at time s. As discussed, in the infinitely lived agent economy this state dependency is not present. Another way to see this is to consider the dynamics of the consumption share in Proposition 3 when ν approaches 0, which corresponds to the infinitely lived economy. In this case, the dependence on β^i disappears and the system is only a forward system. Thus, for a given initial consumption share distribution it can be easily simulated forward. When $\nu \neq 0$ this is no longer possible as β_t^i depends on the valuation ratios at time t, which again depend on the entire consumption share distribution in the future.

2.8 A Solution to the Curse of Dimensionality through a Local Pareto Weight Representation

Once we have $\nu \neq 0$, then the fraction of each type born every period, α_t^i, becomes an additional state variables. Consider the total consumption of all newborns at time t

$$C_{t,t} = \int_i \alpha_t^i c_{t,t}^i di, \tag{40}$$

and define the within cohort consumption share among the newborns as

$$\lambda_t^i = \frac{\alpha_t^i c_{t,t}^i}{C_{t,t}}. \tag{41}$$

Therefore, $\int_i \lambda_t^i di = 1$ for all t. We make two assumptions.

Assumption 1. *We posit that the dynamics of α^i and $G_{s,t}$ for $s \leq t$, such that the following two assumptions hold*

1. $\lambda_t^i = \lambda^i$ *(time invariant within cohort scaled Pareto weight);*

2. $C_{t,t} = \beta Y_t$ *(time invariant consumption share of newborns).*

Assumption 1 greatly simplifies the solution of the model. The key here is that we use the α processes to control the intragenerational importance of the different agents and the G_s function to control the intra-generational allocations. For instance, consider an economy with N agent types. In this case, we need $N-1$ degrees of freedom to match the λ^i weights and one degree to match the β. Specifically, it implies that $\alpha_t^i \beta_t^i = \alpha_t^i \frac{c_{t,t}^i}{Y_t} = \alpha_t^i \frac{c_{t,t}^i}{C_{t,t}} \frac{C_{t,t}}{Y_t} = \lambda^i \beta$.[8]

When Assumption 1 is satisfied, we have that the consumption shares in Proposition 3 only depend on $\mathcal{R}_t$ and $\mathcal{P}_t$, both known functions of f_t, and the consumption shares themself. A consequence of this is that the dynamics of the consumption shares is only a forward system and, therefore, one can easily simulate it. Further, the stochastic discount factor is also easy to simulate forward as it only depends on the consumption shares since $\beta_t = \beta$. This makes it feasible to use Monte-Carlo simulations to solve for valuations such as individual wealth and stock prices, and numerical analysis will not suffer as much from the curse of dimensionality. Moreover, as we will show in the next section, in combination with Malliavin calculus one can explicitly characterize the dynamics of stock prices and portfolios with low computational cost.

2.9 Stock Price and Return Dynamics

Since the stock market is the claim to the aggregate dividends, D_t, we have that $P_t = \mathbb{E}_t\left[\int_t^{\infty} \frac{\xi_s}{\xi_t} D_s ds\right]$. As $\xi_t P_t + \int_{-\infty}^{t} \frac{\xi_s}{\xi_t} D_s ds = \mathbb{E}_t\left[\int_{-\infty}^{\infty} \frac{\xi_s}{\xi_t} D_s ds\right]$ is a local martingale, we can apply the Clark-Ocone formula to find return dynamics.

[8]Further details including proofs and extensive numerical analysis are available in a companion paper (Ehling, Guo, Heyerdahl-Larsen, and Illeditsch (2021)).

Proposition 5. *The stock market return process, R_t, evolves as*

$$dR_t = \mu_{R,t}dt + \sigma_{R,t}dz_t, \tag{42}$$

where

$$\mu_{R,t} = r_t + \theta_t \sigma_{R,t}, \tag{43}$$

with

$$\sigma_{R,t} = \sigma_Y - \frac{\mathbb{E}_t\left[\int_t^\infty \xi_s D_s \int_t^u \left(\mathbb{D}_t r_u + \theta_u \mathbb{D}_t \theta_u\right) duds\right]}{\mathbb{E}_t\left[\int_t^\infty \xi_s D_s ds\right]}, \tag{44}$$

and where $\mathbb{D}_t X_u$ denotes the Malliavin derivative at time t of X_u.

From Proposition 5, we see that the loading on the consumption shock, $\sigma_{R,t}$, depends as expected on the output volatility σ_Y. In addition, the second term takes into account how a consumption shock changes the discount rate through the market price of risk and the real short rate in the future. This is captured by the Malliavin derivatives.[9] Intuitively, if a consumption shock increases the future real interest rate or the market price of risk, keeping everything else constant, this will tend to lower the returns. However, as one can see from Equation (44), it is the product between the discounted cash flows $\xi_t D_t$ and the Malliavin derivative that matters and, therefore, if the response of a shock today on the future real short rate or market price of risk (captured through the Malliavin deriavatives) covaries with the discounted cash flows this will impact the loading of the returns. Keeping everything else fixed, a higher covariance between the discounted cash flows, $\xi_t D_t$, and the response in the discount rate, the lower $\sigma_{R,t}$, will be. To evaluate the expectation in Equation (44) the Malliavin derivatives of the interest rate and the market price of risk have to be calculated. To this end, note that both the risk free rate and the market price of risk are functions of

[9] Malliavin derivatives, Nualart (1995), were introduced to economics and finance in Detemple and Zapatero (1991).

the consumption shares:

$$\mathbb{D}_t r_u = \int_i \frac{\partial r_u}{\partial f^i} \mathbb{D}_t f_u^i di, \tag{45}$$

$$\mathbb{D}_t \theta_u = \int_i \frac{\partial \theta_u}{\partial f^i} \mathbb{D}_t f_u^i di. \tag{46}$$

Hence, as long as we know the dynamics of the Malliavin derivatives of the consumption shares, the expectations can be easily evaluated by Monte-Carlo simulations.

Proposition 6. *Let* $L_{t,u}^k = \mathbb{D}_t f_u^k$*, we have that*

$$dL_{t,u}^k = \left(\int_i \frac{\partial \mu_{f^i,u}}{\partial f^i} \mathbb{D}_t L_{t,u}^i di \right) dt + \left(\int_i \frac{\partial \sigma_{f^i,u}}{\partial f^i} \mathbb{D}_t L_{t,u}^i di \right) dz_t, \tag{47}$$

where $L_{t,u}^k = \sigma_{f^i,t}$ *and the expressions for* $\mu_{f^i,t}$ *and* $\sigma_{f^i,t}$ *are as in Proposition 3.*

From Proposition 6, we can see that the Malliavin derivatives of the consumption shares form an autonomous system of stochastic differential equations. For a finite number of agent types, the dynamics of the Malliavin derivatives can then be easily simulated.

3 Numerical Analysis

3.1 Parameters

We present numerical examples with 10 agent types, where we set the lowest risk aversion to the 5th percentile (1.9) and the highest risk aversion to the 95 percentile (19.5) from the survey imputed risk aversion in Kimball, Sahm, and Shapiro (2008) and assign equal weight to each agent type. The percentiles of risk aversion and corresponding weights are shown in Table 1.

Besides the values for relative risk aversion and the initial Pareto weights, which we also use for the benchmark case with infinite life, our model has another six parameters $(\rho, \nu, \mu_Y, \sigma_Y, \varphi, \beta)$. As in Garleanu and Panageas (2015), we use a time discount factor, ρ, of 0.001, and a hazard rate, ν, of 0.02. The drift, μ_Y, and volatility, σ_Y, of aggregate output are 2% and 3%, respectively. We set the leverage of the consumption claim, φ, to 3 where the aggregate stock market is the claim to Y_t^{φ}. The ratio of the consumption of the newborn agents to the consumption of the total population, β, is 1.65. We employ it to match the real short rate.

With these parameter values, summarized in Table 2, the equity premium on the leveraged claim is 1.4% with a standard deviation of 10.2%. Thus, the model does not match the equity premium because the average risk aversion in our economy is 4.6. To increase the equity premium one would have to increase the weight on the more risk averse agents in the economy. The average real interest rate is matched at 2.1%.

3.2 Consumption Shares

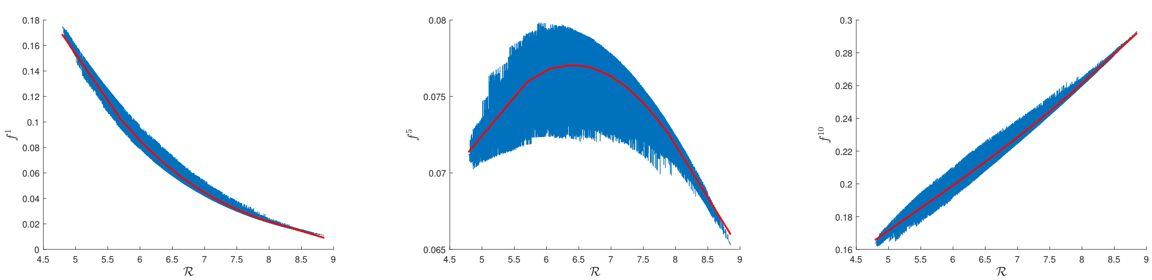


Figure 2: **Consumption Shares.** The left, middle and right panel in the figure plot the consumption shares for the first agent (lowest risk aversion), the fifth agent, and the tenth agent (highest risk aversion) in our OLG model as a function of the aggregate risk aversion, $\mathcal{R}$. The red line is a nonlinear regression of the consumption shares on the aggregate risk aversion to capture the non-monotonic relation for agents in the middle of the risk aversion distribution.

Figure 2 plots the consumption shares of three of the agents as a function of aggregate risk aversion. The left (right) plot shows the agent with the lowest (highest) risk aversion. As expected, the consumption share is decreasing (increasing) as a function of the aggregate risk aversion just as in an economy without overlapping generations. The middle plot shows the consumption share of agent five. Here we see that the relation between the aggregate

Table 1: **Risk Aversion with Initial Pareto Weights.** Percentiles of survey imputed risk aversion in Kimball, Sahm, and Shapiro (2008) with our uniform initial baseline weight (λ).

Percentile	Risk aversion	Initial baseline weight (λ)
5	1.9112	0.10
15	2.9716	0.10
25	3.8638	0.10
35	4.7656	0.10
45	5.7534	0.10
55	6.9041	0.10
65	8.3352	0.10
75	10.2806	0.10
85	13.3675	0.10
95	19.4695	0.10

Table 2: **Parameters.**

Parameter	Value	Description
ρ	0.001	Time discount factor
ν	0.02	Hazard rate
μ_Y	2%	Drift term of aggregate output
σ_Y	3%	Diffusion term of aggregate output
φ	3	Leverage of consumption claim
β	1.65	Ratio of consumption of the newborn agents to total consumption

risk aversion and the consumption share is much weaker. There are two reasons for this. First, the relation is non-monotonic as the consumption volatility is lower than the aggregate consumption volatility when the aggregate risk aversion is low and higher than aggregate consumption volatility when the aggregate risk aversion is high. This is also true for an economy without overlapping generations. However, we see that the consumption share can take many values for a fixed value of the aggregate risk version. This is unique to the overlapping generations economy, where there are $N-1$ state variables in an economy with N types of risk aversion.

To illustrate the importance of the additional state variables, Table 3 shows the R^2 from two different regressions. The first is based on a non-linear regression of the consumption shares on the aggregate risk version. This captures the red line in Figure 2, but for all the 10 agents types. As we can see, the risk aversion captures the lion's share for the agents with either high or low risk aversion. However, for the agents in the middle this is not the case. Yet, once we also include the consumption share weighted variance of the risk-tolerance, $\mathcal{V}(\tau)$, even for agent five, the R^2 reaches 97.92%. Hence, although there are $N-1$ state variables, the aggregate risk aversion, $\mathcal{R}$, and the cross-sectional variance of the risk tolerance, $\mathcal{V}$, capture most of the variation.

Figure 3 shows the loadings of the consumption shares on the aggregate risk aversion, b_R and the cross-sectional consumption share weighted variance of the risk tolerance, $b_{\mathcal{V}}$. In the regression, we orthogonalize the variance to capture the independent variation in the cross-sectional variance of the risk tolerance. From the figure, we can see a positive relation between the risk aversion of the agents and the aggregate risk aversion. The loading on the cross-sectional variance is hump-shaped. Independent variations in $V(\tau)$ increases the consumption share of the high and low risk aversion agents while it decreases the consumption share of agents with an intermediate risk aversion.

The fact that the cross-sectional variance $\mathcal{V}(\tau)$ is important for the consumption shares has asset pricing implications. Consider the diffusion coefficient of the stock market, $\sigma_{R,t}$.

Table 3: R^2 **for Consumption Share Regressions.** The table shows the R^2 from two regressions. The first is a non-linear regression of the consumption shares on the aggregate risk aversion, $\mathcal{R}$ and the second is from regressing the consumption shares on aggregate risk aversion and the consumption share weighted variance of the risk tolerance.

Agent	Non-linear $\mathcal{R}$	$\mathcal{R}$ and $\mathcal{V}(\tau)$
1	99.96	99.91
2	97.14	99.45
3	92.57	99.35
4	84.76	99.07
5	63.34	97.92
6	78.25	98.40
7	98.93	99.77
8	99.86	99.93
9	99.02	99.93
10	98.00	99.89

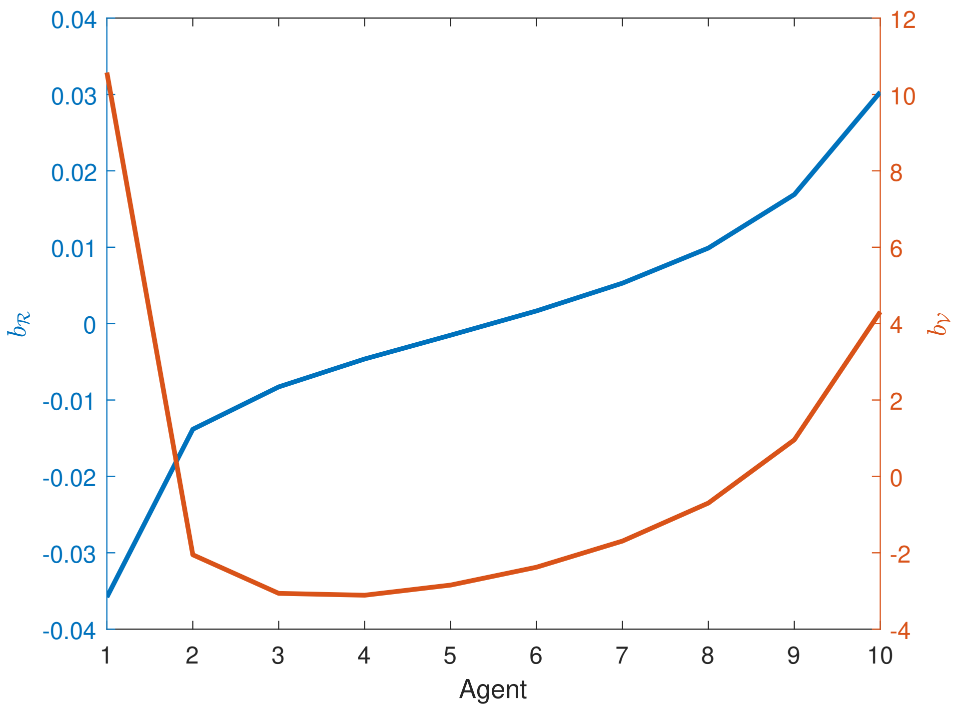


Figure 3: **Consumption Share Regression Coefficients.** The figure shows the relation between the consumption shares and the aggregate risk aversion $\mathcal{R}$ on the left axis and the (orthogonalized) consumption share weighted cross-sectional variance of the risk tolerance $V(\tau)$ on the right axis.

Regressing this volatility on aggregate risk aversion using a non-linear regression to capture non-monotonicity leads to an R^2 of 82.79%. This would be 100% in an economy with infinitely lived agents. Including the variance $\mathcal{V}(\tau)$ increases the R^2 to 98.25%. The loading on the independent variation in the variance $\mathcal{V}(\tau)$ is positive. Hence, as the consumption shares of the agents with high *and* low risk aversion increases at the expense of the agents with intermediate risk version the stock market volatility goes up. As the price of risk is only driven by the aggregate risk aversion, this implies that expected excess returns are higher in states with high consumption share weighted variance in the risk tolerance (for fixed aggregate risk aversion).

3.3 Predictive Regression

The previous subsection shows that the independent variation in the consumption share weighted cross-sectional variance of the risk tolerance, $\mathcal{V}(\tau)$, is important for the variation in the consumption shares. Moreover, given the impact on the expected excess return, an implication of our model that differs from previous models with heterogeneous risk aversion is that this independent variation should predict returns. Unfortunately, we do not observe the independent variation in $\mathcal{V}(\tau)$. To this end, we note that agents with different risk aversion load differently onto the aggregate risk aversion, $\mathcal{R}_t$, and the cross-section variation $\mathcal{V}(\tau)$ as illustrated in Figure 3. From this figure, we see that the most risk tolerant agent, agent 1, loads negatively on the aggregate risk aversion. Moreover, both the high and the low risk aversion agent load positively on the cross-sectional variance, while the agents in the middle of the risk aversion distribution load negatively. Therefore, we can construct a "level" and "slope" factor. Specifically, we define the level factor as the consumption share of the most risk tolerant agent. This has a correlation with the aggregate risk aversion of -97.29%. We define the "slope" factor as the sum of the consumption share of the most and least risk tolerant agents minus the median agent (agent 5) in our model. This measure has a correlation with the orthogonal component of the cross-sectional variance of 95.25% in our simulation using $10,000$ years. Hence, given the model we can closely match the two factors, $\mathcal{R}_t$ and the independent variation in $\mathcal{V}(\tau)$, by using the level and slope factor. The benefit of this is that we do not need to know the actual risk tolerance of the agents.

To find a proxy in the data for the level and slope factor, we use for the level the log wealth share of Forbes 400 (the top 400 richest people in Forbes list) and for the slope the wealth share of Forbes 400 plus the wealth share of the top 10% group minus the wealth share of the top 0.1%.[10] We use the sample period from 1982 to 2019 because the wealth share of Forbes 400 is available since 1982.

The level factor is the same as used by Gomez (2019). He uses this to test the implications of a model with heterogeneous risk aversion, where the difference to our paper is that he considers two agents and, hence, the level factor is sufficient to capture the dynamics of risk premia in his model. As we showed above, this is no longer sufficient in a multi-agent overlapping generations model. Thus, we include the slope factor.

In Tables 4 and 5, we run univariate and bivariate regressions, respectively, with annual excess returns as dependent variable. From the tables, we can see that both the level and the slope have predictive power. Specifically, the higher the level the lower the returns which is consistent with the model, where a higher wealth share indicates goods times and low risk aversion. The slope produces a positive sign, also consistent with the model. Further, the wealth share shows similar economic significance relative to risk aversion: An increase in the wealth share by one standard deviation raises the annual excess returns by 29.4% of its standard deviation while an increase in risk aversion by one standard deviation decreases the annual excess return only by 24.2% of its standard deviation.

[10] These wealth share data are used in Saez and Zucman (2016). See http://gabriel-zucman.eu/uswealth/ for a detailed construction of the wealth share data.

4 Conclusion

Investors differ in their risk tolerance and it is well known that this has implications for asset prices and trading behavior. The extant literature focuses mainly on the case of infinitely lived investors or more recently on agents with finite life in overlapping generations models. However, the overlapping generations models in the literature have focused on the two agent setting. Models with infinitely lived agents or two-agent type in overlapping generation models have in common that they can be described by a single state variable – the aggregate risk aversion. However, once we go beyond two agents in an overlapping generations setting, the number of state variables is given by $N-1$ for an economy with N agents. We then show that even with a very high number of state variables, most of the variation in consumption shares, and hence asset prices, can be captured by the aggregate risk aversion and the cross-sectional variance of the risk tolerance. Next, we show that the agent with the highest and the lowest risk tolerance respond differently to this factor than agents with intermediate risk tolerance. Moreover, variations in the cross-sectional variance impact the excess returns in the economy, where a higher cross-sectional variance is associated with a higher risk premium on the aggregate market. We verify this implication of our model by using data on the wealth distribution and show that not only the level, but also the slope predicts returns.

References

Bhamra, H. S., and R. Uppal, 2014, "Asset Prices with Heterogeneity in Preferences and Beliefs," *Review of Financial Studies*, 27, 519–580.

Blanchard, O., 1985, "Debt, Deficits, and Finite Horizons," *Journal of Political Economy*, 93, 223–247.

Chan, Y. L., and L. Kogan, 2002, "Catching Up with the Joneses: Heterogeneous Preferences and the Dynamics of Asset Prices," *Journal of Political Economy*, 110, 1255–1285.

Cox, J., and C. Huang, 1989, "Optimal Consumption and Portfolio Policies when Asset Prices Follow a Diffusion Process," *Journal of Economic Theory*, 83, 33–83.

Cvitanic, J., E. Jouini, S. Malamud, and C. Napp, 2012, "Financial Markets Equilibrium with Heterogeneous Agents," *Review of Finance*, 16, 285–321.

Cvitanic, J., and S. Malamud, 2011, "Price Impact and Portfolio Impact," *Journal of Financial Economics*, 100, 201–225.

Detemple, J. B., and F. Zapatero, 1991, "Asset Pricing in an Exchange Economy with Habit Formation," *Econometrica*, 59, 1633–1657.

Dumas, B., 1989, "Two-Person Dynamic Equilibrium in the Capital Market," *Review of Financial Studies*, 2, 157–188.

Ehling, P., A. Graniero, and C. Heyerdahl-Larsen, 2018, "Asset Prices and Portfolio Choice with Learning from Experience," *Review of Economic Studies*, 85(3), 1752–1780.

Ehling, P., J. Guo, C. Heyerdahl-Larsen, and P. Illeditsch, 2021, "Asset Prices with Overlapping Generations of Heterogeneous Agents and Firms," *Working Paper*.

Ehling, P., and C. Heyerdahl-Larsen, 2017, "Correlations," *Management Science*, 63(6), 1919–1937.

Garleanu, N., and S. Panageas, 2007, "Young, Old, Conservative, and Bold: The Implications of Heterogeneity and Finite Lives for Asset Pricing," *Working Paper*, 670–685.

Garleanu, 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.

Garleanu, N., and S. Panageas, 2020a, "Heterogeneity and Asset Prices: An Intergenerational Approach," *Working Paper*, 1–42.

Garleanu, N., and S. Panageas, 2020b, "What to expect when everyone is expecting: Self-fulfilling expectations and asset-pricing puzzles," *Journal of Financial Economics*, forthcoming, 1–20.

Gomez, M., 2019, "Asset Prices and Wealth Inequality," *Working Paper*.

Heyerdahl-Larsen, C., and P. Illeditsch, 2020, "Demand Disagreement," *Working Paper*.

Kimball, M. S., C. R. Sahm, and M. D. Shapiro, 2008, "Imputing Risk Tolerance from Survey Responses," *Journal of the American Statistical Association*, 103, 1028–1038.

Krusell, P., and A. A. Smith, 1998, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106(5), 867–896.

Longstaff, F. A., and J. Wang, 2013, "Asset Pricing and the Credit Market," *Review of Financial Studies*, 25, 3169–3215.

Malamud, S., 2008a, "Long run forward rates and long yields of bonds and options in heterogeneous equilibria," *Finance and Stochastics*, 12, 245–264.

Malamud, S., 2008b, "Universal Bounds for Asset Prices in Heterogenous Economies," *Finance and Stochastics*, 12, 411–422.

Martin, I., 2013, "The Lucas Orchard," *Econometrica*, 81, 55–111.

Nualart, D., 1995, *The Malliavin Calculus and Related Topics*, Springer Verlag.

Saez, E., and G. Zucman, 2016, "Wealth inequality in the United States since 1913: Evidence from capitalized income tax data," *Quarterly Journal of Economics*, 131(2), 519–578.

Santos, T., and P. Veronesi, 2018, "Leverage," *Working Paper*, 1–50.

Veronesi, P., 2019, "Heterogeneous Households under Uncertainty," *Working Paper*, 1–61.

Wang, J., 1996, "The Term Structure of Interest Rates in a Pure Exchange Economy with Heterogeneous Investors," *Journal of Financial Economics*, 41, 75–110.

Yaari, M. E., 1965, "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer," *Review of Economic Studies*, 32, 137–150.

Yan, H., 2008, "Natural Selection in Financial Markets: Does It Work?" *Management Science*, 54, 1935–1950.

Zapatero, F., and C. Xiouros, 2010, "The Representative Agent of an Economy with External Habit Formation and Heterogeneous Risk Aversion," *Review of Financial Studies*, 23, 3017–3047.

Table 4: **Univariate Regressions.** The table shows the standardized estimated coefficients, t-statistics (Newey-West corrected with 12 lags) and R^2 for the univariate regression. The second row is obtain by regressing annual excess return on level while the third row is obtained by regressing annual excess return on the slope factor. The sample is from 1982 to 2019.

	b	t-stat	R^2
level	-0.305	-4.556	0.067
slope	0.346	2.939	0.095

Table 5: **Bivariate Regressions.** The table shows the standardized estimated coefficients, t-statistics (Newey-West corrected with 12 lags) and R^2 for the joint regression, where we regress annual excess return on level and on the slope factor. The sample is from 1982 to 2019.

	b_1	t-stat	b_2	t-stat	R^2
level	-0.242	-2.816			0.127
slope			0.294	2.146	