Asset Pricing with Heterogeneous Agents and Long-Run Risk*

Walter Pohl
University of Zurich

Karl Schmedders
University of Zurich and Swiss Finance Institute

Ole Wilms
University of Zurich

December 20, 2016

PRELIMINARY AND INCOMPLETE

Abstract

This paper examines the effect of agent belief heterogeneity on long-run risk models. We find that for the long-run risk explanation to adequately explain the equity premium, it is not sufficient for long-run risk to merely *exist*: agents must all agree that it exists. Agents who believe in a lower persistence level come to dominate the economy rather quickly, even if their belief is wrong. This drives the equity premium down below the level observed in the data.

Keywords: asset pricing, long-run risk, recursive preferences, heterogeneous agents.

JEL codes: G11, G12.

*We are indebted to Jarda Borovička, Nicole Branger, Kenneth Judd, and Lars Hansen for helpful discussions on the subject. Karl Schmedders gratefully acknowledges financial support from Swiss Finance Institute.

1 Introduction

The Bansal-Yaron long-run risk model (Bansal and Yaron (2004)) has emerged as perhaps the premier consumption-based asset pricing model. It can generate many of the features of aggregate stock prices that have long been considered puzzles. The model generates a high equity premium by combining two mechanisms – investors with a taste for early resolution of uncertainty, and very persistent shocks to the growth rate of consumption. This persistence is hard to detect in the data, but its presence is enough to replicate many features of the stock market.

For long-run risk to generate a high equity premium, the level of persistence must be very close to a unit root. The amount of persistence is very hard to measure, and arguments for a range of estimates have appeared in the literature (Bansal, Kiku, and Yaron (2016), Schorfheide, Song, and Yaron (2016) or Grammig and Schaub (2014)). This suggests that there is considerable scope for disagreement over the true value.

In this paper, we consider the consequences if agents themselves disagree. Our results are quite surprising. We find that when agents disagree about the level of persistence that the agent who is more skeptical about long-run risk—who believes in that the level of persistence is lower—dominates the economy. This happens even if the beliefs of the skeptical agent are wrong. More surprisingly, the skeptical agent will dominate the economy in a short amount of time even for small belief differences. In turn, this drives the equity premium below the level seen in the data. Thus, for long-run risk to work as an explanation of the equity premium, it is not enough for long-run to exist—all agents in the economy must also believe in it.

This may seem surprising because it is well established that agents with constant relative risk aversion (CRRA), in the long run the agent with correct beliefs (Sandroni (2000), Blume and Easley (2006), Yan (2008)) always grows to dominate the economy. This analysis breaks down once you allow agents to have preferences for early or late resolution of risk (Borovička (2015)), which allows agents with incorrect beliefs to survive and even drive out agents with incorrect beliefs.

The disagreements can be very small. Disagreement could for example arise because investors use different estimation techniques or data samples to estimate the long-run risk component in consumption (see for example the estimation studies of Bansal, Kiku, and Yaron (2016), Schorfheide, Song, and Yaron (2016) or Grammig and Schaub (2014) who all report different estimates for the persistence of the long-run risk component) or because an investor does not believe in long-run risks at all. Kroenecke (2016), for example, argues that the persistence in consumption arises solely from the filtering procedures and that the unfiltered consumption series is in fact not persistent at all.

We find that investors who are more skeptical about long-run risks accumulate wealth on average. Even if they initially hold only a very small consumption share, their share increases dramatically after short time periods. As small differences in the beliefs about the long-run risk process have large effects on asset prices, we report a drop in the equity premium by 2% within a century for our baseline calibration. For slightly larger belief differences the drop increases to more than 3.5%. This result holds true irrespectively of whether the skeptical investors have the correct beliefs or not (if they also have the correct beliefs, the drop in the premium is even more severe).

While the difference in beliefs poses a puzzle for the explanation of expected returns, it significantly helps in explaining the volatility figures. Beeler and Campbell (2012) show that the long-run risk models of Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012) can not explain the large volatility of the price-dividend ratio observed in the data (a value of 0.45 compared to 0.18 in the models). Differences in the beliefs about long-run risk can generate large shifts in the wealth distributions. This in turn increases the volatility of the price-dividend ratio as the impact of the different agents on asset prices varies over time. We find that even a small difference can generate significant excess volatility close to the values observed in the data. This result also gives a model-based explanation for the empirical findings of Carlin, Longstaff, and Matoba (2014) who use data from the mortgage-backed security market and show that higher disagreement leads to higher volatility. They also show that, as in our model, disagreement is time-varying and correlated with macroeconomic variables.

Related Literature The study of agent belief heterogeneity begins with the market selection hypothesis of Alchian (1950) and Friedman (1953). In analogy with natural selection, the market selection hypothesis states that agents with systematically wrong beliefs will eventually be driven out of the market. The influence of agent heterogeneity on market outcomes under the standard assumption of discounted expected utility is well-understood, and consistent with market selection. Sandroni (2000) and Blume and Easley (2006) find strong support for this hypothesis under the assumption of time separable preferences in an economy without growth. Yan (2008) and Cvitanić, Jouini, Malamud, and Napp (2012) analyze the survival of investors in a continuous-time framework where there are not only differences in the beliefs but also potentially differences in the utility parameters of the investors. They show that it is always the investor with the lowest survival index[1] who survives in the long-run. However, the 'long-run' can be very long and hence, irrational investors can have significant effects on asset prices even under the assumption of discounted expected utility. David (2008) considers a

[1] Yan (2008) shows that the survival index increases with the belief distortion, risk aversion and subjective time discount rate of the investor.

similar model setup, where both agents have distorted estimates about the mean growth rate of the economy and shows, that—as agents with lower risk aversion undertake more aggressive trading strategies—the equity premium increases, the lower the risk aversion. Chen, Joslin, and Tran (2012) analyze how differences in the beliefs about the probability of disasters affect asset prices. They show that, even if there is only a small fraction of investors who are optimistic about disasters, they sell insurance for the disaster states and hence, eliminate most of the risk premium associated with disaster risk.

For non-expected utility equilibrium outcomes change fundamentally. However, there has been comparably little research in this area, as solving such models is anything but trivial.[2] Borovička (2015) shows that agents with fundamentally wrong beliefs can survive or even dominate in an economy with recursive utility.[3] So the inferences about market selection and equilibrium outcomes fundamentally differ under the assumption of general recursive utility compared to the special case of standard time separable preferences. While Borovička (2015) concentrates on the special case of i.i.d. consumption growth, Branger, Dumitrescu, Ivanova, and Schlag (2011) generalize the results to a model with long-run risks as a state variable.

However, most papers with heterogeneous investors and recursive preferences only consider an i.i.d. process for consumption growth. For example Gârleanu and Panageas (2015) analyze the influence of heterogeneity in the preference parameters on asset prices in a two agent OLG economy. Roche (2011) considers a model where the heterogeneous investors can only invest in a stock but there is no risk-free bond. Hence, as there is no savings trade-off, the impact of recursive preferences on equilibrium outcomes will be quite different.

Exceptions that relax the i.i.d. assumptions are for example the papers by Branger, Konermann, and Schlag (2015) or Collin-Dufresne, Johannes, and Lochstoer (2016). Both papers reexamine the influence of belief differences about disaster risk with Epstein-Zin instead of CRRA preferences as in Chen, Joslin, and Tran (2012). Branger, Konermann, and Schlag (2015) provide evidence that the influence of investors with more optimistic beliefs about disasters is less profound, when the disaster occurs to the growth rate of consumption and

[2] Dumas, Uppal, and Wang (2000) show how to solve continuous time asset pricing models with heterogeneous investors and recursive utility. In particular, they show how to characterize the equilibrium by a single value function instead of one value function for each agent. Bhamra and Uppal (2014) show how to solve models with heterogeneous investors that have habit preferences. Related to the method we use in this paper is the approach described in Collin-Dufresne, Johannes, and Lochstoer (2015), who show how to solve discrete time economies with heterogeneous investors and recursive preferences. They derive similar expressions for the characterization of the equilibrium by equation the intertemporal marginal rates of substitutions of the investors. However, their numerical methods to solve for the equilibrium functions numerically fundamentally differs from our approach. While they transform the infinite horizon economy to a finite horizon and solve the model by a backward recursion, we propose a solution method based on projection methods to actually solve the infinite horizon problem.

[3] Borovička (2015) describes four channels that affect equilibrium outcomes. We examine these channels in more detail in Section 4.1 and show how they affect equilibrium outcomes in the asset pricing model considered in this paper.

investors have recursive preferences. Collin-Dufresne, Johannes, and Lochstoer (2016) make a similar claim but for a different reason. They show that, if the investors can learn about the probability of disaster and if the investors have recursive preferences, the impact of the optimistic investor on asset prices decreases. Optimists are uncertain about the probability of disaster and hence will provide less insurance to the pessimistic investors. Collin-Dufresne, Johannes, and Lochstoer (2016) use an OLG model with two generations to model optimists and pessimists. Hence—in contrast to the results in this study—the consumption shares of the investors are fixed and the increasing influence of optimistic agents due to the risk aversion channel over time are not captured.

Also related to our work is the paper by Andrei, Carlin, and Hasler (2016). While in this paper, the agents agree to disagree about the long-run risks in the economy, Andrei, Carlin, and Hasler (2016) provide an explanation how this disagreement can arise from model uncertainty as market participants calibrate their models differently. They consider a setup with and find that uncertainty about long-run risks can explain many stylized facts of stock return volatilities like large volatilities during recessions and booms and persistent volatility clustering.

The paper is organized as follows. In Section 2.1 we describe the general equilibrium for the asset pricing model with heterogeneous investors and recursive preferences. A detailed derivation of the equilibrium is shown in Appendix A.1. Section 3 describes the long-run risk model with 2 investors and different beliefs about long-run risks. Results are shown in Section 4.

2 Theoretical Framework

2.1 The Heterogeneous Agents Economy

We consider a standard infinite-horizon discrete-time endowment economy, with a finite number of agents. We introduce a general setup for heterogenous agents. Agents can differ on both their utility functions and their subjective beliefs.

We restrict our attention to the complete-market setting, which allows us to reformulate the problem as a social planner's problem. Here we run into one critical difference with the representative agent problem – the social planner's problem is not recursive. This defeats most of the techniques to solving for equilibrium in an infinite-horizon model.

This failure of recursiveness occurs for essentially economic reasons – even if aggregate consumption does not contain a trend, the individual consumption allocations can. For example, Blume and Easley (2006) show that if agents have different beliefs, then the individual consumption of an agent with wrong beliefs will trend down over time. Yan (2008) shows that in an economy with growth and agents with differing risk aversion, the relative consumption of the more risk-averse agent tends downward.

Our theoretical contribution is to present a reformulation of the first-order conditions for equilibrium that is recursive. This reformulation involves introducing new state variables. Interestingly, the state variables have a clear interpretation in terms of time-varying weights in the social planner's problem. These weights capture the relative trend in an agent's consumption – an agent that has a declining share of consumption will have a decling weight.

We introduce some notation to state our result. Let $t \in \{0, 1, \ldots, \}$, and let y_t be the exogenous time t variable that determines the current state of the economy. Aggregate consumption is a purely a function of the exogenous state, $C(y^t)$.

The economy is populated by a finite number of infinitely-lived agents, H. Let $\mathbb{H} = 1 \ldots H$ and $\mathbb{H}^- = 2 \ldots H$. Agents choose individual consumption at time t as a function of the history of the exogenous state, y^t, where $y^t = (y_0, ..., y_t)$. Note that as will become clear, we cannot make individual consumption a function of the exogeneous state alone, even if y_t is a Markov process. Let $C_t^h = C^h(y^t)$ be the individual consumption for agent h.

Under market clearing,

$$\sum_{h=1}^{H} C^h(y^t) = C(y^t). \tag{1}$$

Agents have subjective beliefs about the probability distribution of the exogenous state variable. We denote the expectation operator for agent h at time t as E_t^h. Each agent has recursive utility. Let $\{C^h\}_t = \{C^h(y^t), C^h(y^{t+1}), \ldots\}$ denote the consumption stream of agent h from time t forward. Utility for agent h at time t, $U^h(\{C^h\}_t)$, is specified by an aggregator $F^h(c, x)$ and a certainty-equivalence function $G(x)$,

$$U^h(\{C^h\}_t) = F^h\left(C^h(y^t), R_t^h\left[U^h(\{C^h\}_{t+1})\right]\right) \tag{2}$$

where

$$R_t^h[x_{t+1}] = G_h^{-1}(E_t^h[G_h(x_{t+1})]). \tag{3}$$

We assume F^h and G^h are both differentiable. This framework includes both Epstein-Zin utility, and discounted expected utility, for the appropriate choices of F^h and G^h. We consider Epstein-Zin utility in section 2.2 and discounted expected utility in section 2.3.

To simplify the analysis, we ensure that agents never choose zero consumption in any state of the world. We do so by imposing an Inada condition on F^h: $F_1(c, x) \to \infty$ as $c \to 0$.

We also impose one condition on the agents' beliefs. Let $P_{t,t+1}^h$ be the subjective conditional distribution of y_{t+1} given y^t, and $P_{t,t+1}$ be the true conditional distribution. We assume that each agent's expectation can be written in terms of the true distribution as

$$E_t^h[x] = E_t\left[x\frac{\mathrm{dP}_{t,t+1}^h}{\mathrm{dP}_{t,t+1}}\right],$$

for some measurable function $dP_{t,t+1}^h/dP_{t,t+1}$. (In mathematical terms, every agent's conditional distribution is absolutely continuous with respect to the true distribution. Then by the Radon-Nikodym theorem such a $dP_{t,t+1}^h/dP_{t,t+1}$ must exist. Accordingly, $dP_{t,t+1}^h/dP_{t,t+1}$ is known as the Radon-Nikodym derivative of $P_{t,t+1}^h$ with respect to $dP_{t,t+1}$.)

To solve for equilibrium, we assume that markets are complete and reformulate as a social welfare problem. The social planner solves a weighted sum of the individual agent's utilities at $t = 0$. Let $\boldsymbol{\lambda} = \{\lambda^1, \ldots, \lambda^H\}$, and $\{\boldsymbol{C}\}_0 = \{\{C^1\}_0, \ldots, \{C^H\}_0\}$. Then the social planner problem is

$$SP(\{\boldsymbol{C}\}_0, \boldsymbol{\lambda}) = \sum_{h=1}^{H} \lambda^h U^h\left(\{C^h\}_0\right). \tag{4}$$

Theorem 1. *Let $\{\boldsymbol{C}\}_0$ be the solution to the social planner's problem for weights $\boldsymbol{\lambda}$. For each agent, let*

$$U_t^h = U^h(\{C^h\}_t),$$

where U^h is evaluated at the optimum.

A solution to the social planner's problem 4 solves the following first-order conditions for each state t,

$$\lambda_t^h F_1^h(C_t^h, R_t^h\left[U_{t+1}^h\right]) = \lambda_t^1 F_1^1(C_t^1, R_t^1\left[U_{t+1}^1\right]) \tag{5}$$

where the λ_t^h satisfy

$$\lambda_0^h = \lambda^h, \tag{6}$$

$$\frac{\lambda_{t+1}^h}{\lambda_{t+1}^1} = \frac{\Pi_{t+1}^h}{\Pi_{t+1}^1}\frac{\lambda_t^h}{\lambda_t^1}, \qquad t > 0, h \in \{2, \ldots H\}, \tag{7}$$

where Π_t^h is given by

$$\Pi_{t+1}^h = F_2^h\left(C_t^h, R_t^h[U_{t+1}^h]\right) \cdot \frac{G_h'(U_{t+1}^h)}{G_h'(R_t^h[U_{t+1}^h])}\frac{\mathrm{dP}_{t,t+1}^h}{\mathrm{dP}_{t,t+1}}. \tag{8}$$

The proof can be found in Appendix A.1.

The λ_t^h are only determined up to a scalar factor each period, so we are free to choose a normalization. For numerical purposes, a normalization so that the sum of the λ_t^h equal one every period is convenient. From a conceptual point of view, an attractive choice is to let

$\lambda^1_{t+1} = \Pi^1_{t+1}\lambda^1_t$, because then for all h, $\lambda^h_{t+1} = \Pi^h_{t+1}\lambda^h_t$

If F^h is additively separable, then the allocation of consumption in 5 depends only on the current value for the λ^h_t. Additive separability is the most common case in applications. Discounted expected utility is additively separable, while Epstein-Zin can be transformed to be so. In this particular case, the Negishi weights and individual agent consumption are closely linked. We can sharpen this to an asymptotic statement that relates the limit for λ^h_t and the limit for consumption.

Theorem 2. *Suppose that F^h is additively separable all h, and that aggregate endowment is bounded above a constant $\overline{C}$ and below by the constant $\underline{C}$. If $\lambda^j_t/\lambda^i_t \to \infty$, then $C^i_t \to 0$. If $C^i_t \to 0$, then for at least one agent j, $\limsup_t \lambda^j_t/\lambda^i_t = \infty$.*

Note that $\limsup_t \lambda_j/\lambda_i$ is a random variable – the outcome can depend on the history. This result generalizes a similar result in Blume and Easley (2006). We extend this result to a growth economy in the next section.

2.2 Growth Economy with Epstein-Zin Preferences

We specialize the results of the previous section to the case where agents have Epstein-Zin preferences (Epstein and Zin (1989) and Weil (1989)), and aggregate consumption is expressed in terms of growth rates.

If agent h has Epstein-Zin preferences, then

$$F^h(c,x) = \left[(1-\delta^h)c^{\rho^h} + \delta^h x^{\rho^h}\right]^{1/\rho^h} \tag{9}$$

$$G_h(x) = x^{\alpha^h} \tag{10}$$

In this case, the equations are all homogeneous, so we can divide through by aggregate consumption, and express the equilibrium allocations in terms of individual consumption shares, $s^h_t = C^h_t/C_t$. Market clearing implies that

$$\sum_{h=1}^{H} s^h_t = 1. \tag{11}$$

Let V^h_t be agent h's value function. We also normalize this by aggregate consumption, $v^h_t = V^h_t/C_t$. Let $c_t = \log C_t$. The normalized value functions satisfy the following fixed-point equation,

$$v^h_t = \left[(1-\delta^h)(s^h_t)^{\rho^h} + \delta^h R^h_t \left(v^h_{t+1}e^{\Delta c_{t+1}}\right)^{\rho^h}\right]^{\frac{1}{\rho^h}}, \qquad h \in \mathbb{H}. \tag{12}$$

where $R_t^h(x) = \left(E_t^h\left[x^{\alpha^h}\right]\right)^{\frac{1}{\alpha^h}}$. The parameter δ^h is the discount factor, $\rho^h = 1 - \frac{1}{\psi^h}$ determines the intertemporal elasticity of substitution ψ^h and $\alpha^h = 1 - \gamma^h$ determines the relative risk aversion γ^h of agent h.

To accompany the normalized value function we introduce a normalized Negishi weight, $\underline{\lambda}_t^h = \frac{\lambda_t^h}{(v_t^h)^{\rho^h - 1}}$. In Appendix A.1 we show that the consumption share s_t^h of agent h is given by

$$\underline{\lambda}_t^h (1 - \delta^h)(s_t^h)^{\rho^h - 1} = \underline{\lambda}_t^1 (1 - \delta^1)(s_t^1)^{\rho^1 - 1}. \tag{13}$$

Below we consider the case where agents all have identical ρ^h, so the aggregate consumption term cancels out.

Finally, the equations for $\underline{\lambda}_t^h$ simplify to

$$\frac{\underline{\lambda}_{t+1}^h}{\underline{\lambda}_{t+1}^1} = \frac{\underline{\Pi}_{t+1}^h}{\underline{\Pi}_{t+1}^1} \frac{\underline{\lambda}_{t+1}^h}{\underline{\lambda}_{t+1}^1} \tag{14}$$

$$\underline{\Pi}_{t+1}^h = \delta^h e^{\rho^h \Delta c_{t+1}} \frac{\mathrm{dP}_{t,t+1}^h}{\mathrm{dP}_{t,t+1}} \frac{\left(v_{t+1}^h e^{\Delta c_{t+1}}\right)^{\alpha^h - \rho^h}}{R_t^h\left(v_{t+1}^h e^{\Delta c_{t+1}}\right)^{\alpha^h - \rho^h}}, , \qquad h \in \mathbb{H}^-. \tag{15}$$

This gives us $H - 1$ nonlinear equations for equilibrium. In our numerical calculation, we complete the system by requiring that $\sum \lambda_t^h = 1$. If we solve for the weights, $\underline{\lambda}_t^h$ given by

$$\begin{aligned} \underline{\lambda}_{t+1}^h &= \frac{\underline{\lambda}_t^h \underline{\Pi}_{t+1}^h}{\sum_{h=1}^H \underline{\lambda}_t^h \underline{\Pi}_{t+1}^h} \\ \underline{\Pi}_{t+1}^h &= \underbrace{\delta^h e^{\rho^h \Delta c_{t+1}} \frac{\mathrm{dP}_{t,t+1}^h}{\mathrm{dP}_{t,t+1}}}_{\text{CRRA-Term}} \underbrace{\frac{\left(v_{t+1}^h e^{\Delta c_{t+1}}\right)^{\alpha^h - \rho^h}}{R_t^h\left(v_{t+1}^h e^{\Delta c_{t+1}}\right)^{\alpha^h - \rho^h}}}_{\text{Additional EZ-Term}}, \qquad h \in \mathbb{H}^-. \end{aligned} \tag{16}$$

Unlike the discounted expected utility case, the dynamics of the weights $\underline{\lambda}_t^h$ depend on the value functions (12) that in turn depend on the consumption decisions (13). Hence, to compute the equilibrium we need to jointly solve equations (12), (13), (11) and (16). As there are—to the best of our knowledge—no closed-form solutions for the general model, we present in Appendix A.4 a solution approach based on projection methods to compute for the equilibrium functions numerically.

In this setting, we can derive a considerable improvement over theorem 2 – the limiting behavior for λ_t^h drives the limiting behavior for an agent's share of aggregate consumption. This requires no assumptions on aggregate consumption, only that agents have utility in the Epstein-Zin family.

Theorem 3. *If $\underline{\lambda}_t^j / \underline{\lambda}_t^i \to \infty$, then $s_t^i \to 0$. If $s_t^i \to 0$, then for at least one agent j,*

$\limsup_t \underline{\lambda}_t^j / \underline{\lambda}_t^i = \infty$.

2.3 Discounted Expected Utility

The bulk of our paper is concerned with Epstein-Zin preferences, which requires numerical solution, but in this section we consider classical discounted expected utility, which allows precise theorems. In this special case, the problem simplifies considerably, which allows us to replicate several results in the literature. This allow us to reproduce several classical results (Blume and Easley (2006), Yan (2008)) in our framework.

Discounted expected utility correspond to the case where F^h and G_h take the special forms

$$\begin{aligned} F^h(c,x) &= u^h(c) + \delta^h x \\ G_h(x) &= x. \end{aligned}$$

where u^h is agent h's one-period utility function. u^h must satisfy the Inada condition, $(u^h)'(c) \to \infty$ as $c \to 0$. Then

$$\begin{aligned} F_1^h(C,x) &= u^{h'}(c) \\ F_2^h(C,x) &= \delta^h \\ G_h'(x) &= 1, \end{aligned}$$

and equation 5 becomes

$$\lambda_t^h u^{h'}(C_t^h) = \lambda_t^1 u^{1'}(C_t^1) \tag{17}$$

and equation 8 simplifies dramatically to

$$\Pi_{t+1}^h = \delta^h \frac{dP_{t,t+1}^h}{dP_{t,t+1}}. \tag{18}$$

In the special case where agents have correct beliefs and equal δ^h, this specializes to Lemma 1 from Judd, Kubler, and Schmedders (2003).

Theorem 4. *If aggregate consumption is a Markov process, and all agent have identical δ^h, then individual agent consumptions are also Markov.*

Once we allow the agents to have different δ^h or different beliefs, individual consumption can have a trend, even if the aggregate consumption is Markov (Sandroni (2000), Blume and Easley (2006)). This leads to a genuinely new phenomenon – asymptotically, an agent can become a negligible part of the economy. This phenomenon is known as "survival" – agents survive if they still matters in the long run.

There are competing definitions of survival in the literature. Sandroni (2000) defines survival in terms of wealth, while Blume and Easley (2006) uses consumption. The Blume and Easley (2006) definition fits naturally into our framework. Those authors define agent i as surviving if $\limsup_{t\to\infty} C_t^i > 0$, and vanishing if $\lim C_t^i \to 0$. If $(u^h)'(C) \to \infty$ as $C \to 0$ then for an economy with aggregate consumption bounded above and below, theorem 2 links these directly to the behavior of the λ_t^h.

In turn, the behavior of λ_t^h is governed by the Π_t^h. In the special case of discounted expected utility, Π_t^h only depends on the characteristics of the agents, their utility function and subjective beliefs, and not on any endogenous variables such as individual consumption. This means that an agent's tendency to survive is an intrinsic property of the agent. If Π_t^i tends to be higher than Π_t^j, then in some sense agent i tends to survive relative to j. (This does not hold in the general case, since Π_i^h depends on the agent's value function, which depends on equilibrium consumption allocations, which in turn depends on the other agents.)

If we assume that consumption is i.i.d., then we can give a succinct expression for this intuition, in terms of the Kullback-Leibler divergence, $D(P\|Q)$. (For two probability distributions, P, and Q, the expectation of $\log dP/dQ$ with respect to P is known as Kullback-Leibler divergence, or the relative entropy.)

Theorem 5. *Suppose that consumption is i.i.d., with distribution Q. Suppose agents agree that consumption is i.i.d., and their subjective distribution is Q_i. Then*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \log \Pi_t^h = \log \delta^h - D(Q^h\|Q)$$

by the law of large numbers, so

$$\lim_{T\to\infty} \frac{1}{T} \left(\lambda_T^i - \lambda_T^j\right) = \left(\log \delta^i - D(Q^i\|Q)\right) - \left(\log \delta^j - D(Q^j\|Q)\right).$$

The previous theorem is in Blume and Easley (2006), which refers to $\log \delta^h - D(Q^h\|Q)$ as the survival index of the agent. Agents with higher survival indices will survive relative to agents with lower survival indices.

For general dependent processes, some general results are possible (following Blume and Easley (2006).)

Theorem 6. *1. If agents have identical beliefs but differ on δ, then only the agent with the highest δ survives.*

2. Suppose agents have identical utility functions, but differing beliefs. If one agent has correct beliefs, then only agents with correct beliefs survive.

The previous results depend on the unrealistic assumption that consumption is bounded for all time. For CRRA utility, we can exploit the homogeneity of the utility function to generalize to arbitrary consumption processes. In this case, the appropriate measure of survival is in terms of individual *share* of aggregate consumption, rather than the individual level of consumption. We say that an agent survives if $\limsup_{t\to\infty} s_t^i > 0$.

CRRA is a special case of Epstein-Zin where $\alpha^h = \rho^h$, so we can specialize the results in the previous section. In that case, the equation for $\underline{\Pi}_{t+1}^h$ simplifies to

$$\underline{\Pi}_{t+1}^h = \delta^h e^{\rho^h \Delta c_{t+1}} \frac{\mathrm{dP}_{t,t+1}^h}{\mathrm{dP}_{t,t+1}}.$$

Again each agent's $\underline{\Pi}_t^h$ only depends on exogenous variables, not endogenous, which makes an agent's tendency to survive an intrinsic quality of the agent.

In the i.i.d consumption growth case, we have a result similar to Theorem 5.

Theorem 7. *Suppose that consumption growth is i.i.d., with distribution Q. Suppose agents agree that consumption growth is i.i.d., and their subjective distribution is Q_i. Let μ be the expected log growth rate of consumption under Q. Then*

$$\lim_{T\to\infty} \sum_{t=1}^{T} \log \underline{\Pi}_t^h = \log \delta^h + \rho^h \mu - D(Q^h \| Q)$$

and thus

$$\lim_{T\to\infty} \frac{1}{T} \left(\lambda_T^i - \lambda_T^j\right) = \left(\log \delta^i + \rho^i \mu - D(Q^i \| Q)\right) - \left(\log \delta^j + \rho^j \mu - D(Q^j \| Q)\right).$$

Similar to the no-growth case, $\log \delta^h + \rho^h \mu - D(Q^h \| Q)$ serves as a survival index – the agent with the highest survival index is the one who survives. One interesting wrinkle is that the survival results depend on the sign of the expected log groth rate of the economy. If it is positive, the agent with the highest ρ^h survives, everything else being equal, while if it is negative the lowest ρ^h survives.

A form of the previous theorem can be found in Yan (2008) in continuous time with geometric Brownian motion. If we assume that log consumption growth is normally distributed, and each agent knows the true variance, σ^2, but believes the mean to be μ_i, then the survival index specializes to $\log \delta^h + \rho^h \mu - 1/2(\mu_i - \mu)^2/\sigma^2$, which is essentially identical to Yan. (Yan consider the negative of this quantity, so the agent with the lowest survival index survives.)

Again for general dependent processes, some general results are possible. We need to account for whether the economy is a growth economy in the general case. We use the simple criterion that $E_t(C_{t+1}) \geq C_t$.

Theorem 8. *1. If agents have identical beliefs and ρ^h but differ on δ^h, then only the agent with the highest δ^h survives.*

2. Suppose that agents have identical beliefs and identical δ^h, but differ on ρ^h. Also assume that $E_t(C_{t+1}) \geq C_t$. then the agent with the highest ρ^h survives.

3. Suppose agents have identical CRRA utility parameters, but differing beliefs. If one agent has correct beliefs, then that agent always survives.

Items 1 and 3 of the theorem are the natural generalizations of Blume and Easley (2006) to the growth setting, but item 2 for general dependent processes is new.

3 A Long-Run Risk Model with Differences in Belief

We consider a standard long-run risk model as in Bansal and Yaron (2004) where log aggregate consumption growth Δc_{t+1} and log aggregate dividend growth Δd_{t+1} are given by

$$\begin{aligned} \Delta c_{t+1} &= \mu_c + x_t + \sigma \eta_{c,t+1} \\ x_{t+1} &= \rho_x x_t + \phi_x \sigma \eta_{x,t+1} \\ \Delta d_{t+1} &= \mu_d + \Phi x_t + \phi_d \sigma \eta_{d,t+1} + \phi_{d,c} \sigma \eta_{c,t+1}. \end{aligned} \tag{19}$$

x_t captures the long-run variation in the mean of consumption and dividend growth and $\eta_{c,t+1}$, $\eta_{x,t+1}$ and $\eta_{d,t+1}$ are i.i.d. normal shocks. A key feature of long-run risk models are highly persistent shifts in the growth rate of consumption. Together with a preference for the early resolution of risks ($\gamma > \frac{1}{\psi}$) investors will dislike shocks in x_t and require a large premium for bearing those risks. Hence, the results in the long-run risk literature rely on a highly persistent state process x_t, or put differently ρ_x needs to be very close one (0.979 in the original calibration of Bansal and Yaron (2004)).

In this paper we analyze the equilibrium implications of differences in beliefs about the long-run risk process. As x_t is not directly observable from the data, it is reasonable to assume that investors disagree—at least slightly—about the data generating process of x_t. However, the majority of investors needs to belief in a highly persistent long-run risk process, as otherwise asset prices would be determined by the investors not (or less) believing in long-run risks and hence, the model outcomes would not be consistent with the data. Therefore, we assume that a majority investors beliefs in a highly persistent long-run risk process and address the question what happens if there is a small fraction of investors who beliefs in less persistent shocks, or put differently, who is skeptical about long-run risks.

For this we consider a setup with $H = 2$ agents where the first agent believes that ρ_x is close to one while the second agent believes that ρ_x is slightly smaller. We do not make a specific assumption about which agent has the correct beliefs and we show in our results, that for small belief differences, the true distribution has a negligible influence on equilibrium outcomes. We denote by ρ_x^h the belief of agent h about ρ_x. As x_{t+1} conditional on time t information is normally distributed with mean $\rho_x x_t$ and variance $\phi_x^2\sigma^2$, $\mathrm{dP}_{t,t+1}^h$ is given by

$$\mathrm{dP}_{t,t+1}^h = \frac{1}{\sqrt{2\pi}\phi_x\sigma}\exp\left(-\frac{1}{2}\left(\frac{x_{t+1}-\rho_x^h x_t}{\phi_x\sigma}\right)^2\right)dx_{t+1}.$$

We can think of this model as an extension of Borovička (2015) who considers a two agent setup with different beliefs about the mean growth rate of the economy. For Epstein-Zin preferences, Borovička (2015) shows that the agent with the more optimistic beliefs (a larger belief about the mean growth rate) will dominate the economy in the long-run as long as the risk aversion in the economy is large enough. This result stands in stark contrast to the case of CRRA preferences, where the agent with the more correct beliefs will always dominate independent of the choice of preference parameters (see for example Yan (2008)).[4]

In the model with different beliefs about the persistence of long-run risks, the beliefs about the mean growth rate of the economy change over time. Consider the example where $\rho_x = \rho_x^1 > \rho_x^2$. The time t expectation of agent h about the mean growth rate is given by $\rho_x^h x_t$. This implies that for a negative realization of x_t, $\rho_x^2 x_t > \rho_x^1 x_t$ and hence the second agent is more optimistic. For $x_t > 0$, we have that $\rho_x^2 x_t < \rho_x^1 x_t$ and hence the first agent is more optimistic about the mean growth rate (the second agent is more pessimistic). Hence we can think of this model as a time-varying version of Borovička (2015) where the beliefs about the growth rate change over time. In Section 4.1 we analyze in detail, how the time variation induced by the long-run risks influences equilibrium outcomes.

Most long-run risk models calibrate the underlying cash-flow parameters in order to match asset pricing data. For example Bansal and Yaron (2004) use a value of $\rho_x = 0.979$, Bansal, Kiku, and Yaron (2012) use $\rho_x = 0.975$, and Drechsler and Yaron (2011) assume $\rho_x = 0.976$. They obtain high values of ρ_x by construction, as otherwise the models would not be consistent with the high equity premium observed in the data. The study by Bansal, Kiku, and Yaron (2016) uses cash flow and asset pricing data to estimate the long-run risk model parameters and reports a value of $\rho_x \approx 0.98$ with a standard error of 0.01. For our baseline calibration we assume that the first agent beliefs that $\rho_x^1 = 0.985$. This implies an equity premium of 6.53% for the representative agent economy, which is consistent with the value observed in the data.

[4] We assume that the only difference between the agents is their beliefs about the state processes and they share the same utility parameter specifications.

The second agent has slightly smaller beliefs about the persistence with $\rho_x^2 = 0.975$. Both values lie well within the confidence interval provided by Bansal, Kiku, and Yaron (2016). The effect of a small change in ρ_x has large effects on asset prices. For $\rho_x = 0.975$ the equity premium decreases to 2.76%. For $\rho_x = 0.95$ it already collapses to 0.26% and the influence of x_t on asset prices is negligible. Therefore, we consider a second set of results where agent 1 beliefs that $\rho_x^1 = 0.985$ and agent two beliefs that $\rho_x^2 = 0.95$.

Except from the differences in beliefs, the two agents are the same and share the properties of the representative investor of Bansal and Yaron (2004) with $\psi^1 = \psi^2 = 1.5$, $\gamma^1 = \gamma^2 = 10$, $\delta^1 = \delta^2 = 0.998$. For the remaining parameters of the state processes (19) we also use the calibration from Bansal and Yaron (2004) with $\mu_c = \mu_d = 0.0015, \sigma = 0.0078, \Phi = 3, \phi_d = 4.5, \phi_{d,c} = 0$ and $\phi_x = 0.044$. (This calibration will be used for all results in this paper, unless otherwise stated.)

4 Results

We begin with the analysis of the equilibrium dynamics of the consumption shares of the individual agents. Figure 1 shows the consumption share of the second, skeptical agent ($\rho_x^2 = 0.975$) over time for different initial shares $s_0^2 = \{0.01, 0.05, 0.5\}$. We report the median, 5% and 95% quantile paths using 1000 samples each consisting of 500 years of simulated data. To minimize the influence of the initial value of x_t, we initialize each simulated path by running a burn-in period of 1000 years before using the output. The left panel shows the results for $\rho_x = \rho_x^1 = 0.985$ (the first agent has correct beliefs) and the right panel for ($\rho_x = \rho_x^2 = 0.975$ (the second agent has correct beliefs).

We observe that in all cases the consumption share of the skeptical agent 2 strongly increases over time. While it occurs faster if agent 2 has the correct beliefs (right panel) the increase is almost as strong if agent 1 has the correct beliefs (left panel). Hence, given a small difference in the beliefs, independent of whether agent 1 or agent 2 has the correct beliefs, in the long-run the agent with the lower beliefs about ρ_x will dominate the economy. Most importantly, even if the economy is initially almost entirely populated by agent 1 ($s_0^2 = 0.01$), his consumption share decreases sharply and he loses significant shares in a short amount of time. Table 1 reports the corresponding median consumption shares for different time horizons for $s_0^2 = \{0.01, 0.05, 0.5\}$. We observe that for $s_0^2 = 0.01$ the consumption share of agent 1 has decreased by 27% after 100 years, 63% after 200 years and almost 93% after 500 years.

Figure 2 shows the corresponding results for $\rho_x^2 = 0.95$ and an initial allocation of $s_0^2 = 0.01$. The left panel shows the results for $\rho_x = 0.985$ (agent 1 has the correct beliefs). We observe that the initial increase in the consumption share is stronger, compared to the case with

$\rho_x^2 = 0.975$ but the median share does not become as large in the long-run (the median shares of the second agent after 100, 200 and 500 years are given by 32.59%, 37.82% and 40.19% respectively). Also the 5% and 95% quantile paths show that there is significantly more variation in the shares. The figure also shows a sample paths (grey line). We observe that there are large drops and recoveries in the consumption share. The large drops occur, because the second agent assigns 'wrong' probabilities to extreme states and hence bets on states, that turn out to occur less often in the long-run. This effect works in favor of agent 2, once he has the correct beliefs and therefore he is more likely to bet on the correct states. This case is shown in the right panel ($\rho_x = \rho_x^2 = 0.95$) where we indeed observe that the increase in the consumption share is much stronger and the large drops in consumption are not present anymore. The recoveries in the left panel occur because the second agent is less afraid of long-run risks and hence, sells insurance against this risks to the first agent. As the first agent beliefs that $\rho_x^1 = 0.985$, he strongly dislikes shocks in x_t and is willing to pay a high premium to insure against these risks. So there are two interacting effects that affect equilibrium outcomes. We later provide a detailed analysis of the two effects is in Section 4.1.

What does the change in the consumption shares imply for asset prices and aggregate financial market statistics? We assume that the economy is initially almost entirely populated by agent 1 to generate a high equity premium consistent with the data. But the consumption share of the first agent decreases rapidly and so will his influence on asset prices. In Table 2 we show the annualized equity premium in the years 0, 100, 200 and 500 assuming an initial share of $s_0^2 = 0.01$.[5] The left panel shows the results for $\rho_x^2 = 0.975$ and agent 1 has the correct beliefs. For the initial allocation $s_t^2 = 0.01$, where agent 1 dominates the economy, the aggregate risk premium is 6.42%. A value very close to the representative agent economy populated only by the first agent which implies a premium of 6.53%. After 100 years, when the share of agent 1 has decreased from 99% to 72%, the premium decreases to 4.59%. Hence, even if agent 1 holds almost all wealth initially, which implies a high risk premium, the premium will drop by almost 2% within a century. After 200 years, the premium decreases by almost 3% and after 500 years it is almost at the level of the representative agent economy populated only by agent 2 with a premium of 2.89%. The right panel shows the corresponding results for $\rho_x^2 = 0.95$. We observe that the sharp increase in the consumption share decreases the premium from 5.42% initially to 1.84% after 100 years—a decrease of more than 3.5% in a century. In Table 4 in Appendix B we show the corresponding results for the case where agent 2 has the correct beliefs instead of agent 1. We observe that the drop in the equity premium is

[5]Note that Table 2 does <u>not</u> report the premium starting with a given value for s_0^2 and simulating a long time series, but we report the average premium for a given consumption share $s_t^2 = \bar{s}$. Hence, we take the expectation over all x_t while keeping the consumption share constant at $\bar{s}$. The population moment for 500 years of simulated data is given in Table 3.

even more severe. Hence, the difference in beliefs brings down the equity premium well below the levels observed in the data even if the agent who is skeptical about long-run risks does not have the correct beliefs.

Table 3 shows the population moments from the 1000 sample paths starting with an initial share of $s_0^2 = 0.01$. We report the mean and the standard deviation of the annualized log price-dividend ratio, the annualized equity premium and the risk-free return. Results are shown for the case where agent 1 has the correct beliefs. In addition to the two agent economy, the table also shows the two representative agent cases where the economy is populated only by agent 1 ($s_t^2 = 0$) or agent 2 ($s_t^2 = 1$). While the mean statistics of the two agent economy lie well within the bands of the two representative agent economies and depict the wealth shift towards the second agent, we observe that the volatility of the log price-dividend ratio is significantly larger for the two agent economy compared to both representative agent economies. This effect is especially strong for $\rho_x^2 = 0.95$ where the volatility is 0.48 compared to 0.25 and 0.14 for the two representative agent economics.

Beeler and Campbell (2012) argue that one of the major issues of the long-run risk models of Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012) is that they significantly underestimate the volatility of the price-dividend ratio (they report values of 0.18 compared to 0.45 observed in the financial market data). Our results show, that differences in beliefs can potentially resolve this puzzle leading to a significant increase in the volatility figures. The strong increase can be explained by the large variation in the consumption shares for the case of $\rho_x^2 = 0.95$ (see Figure 2). Variation in the shares implies that the influence of each agent on asset prices varies of time. As both agents have significantly different price-dividend ratios in the representative agent economies (a mean value of 2.68 for agent 1 compared to 6.27 for agent 2), the variation in the consumption shares generates excess volatility for the price-dividend ratio.

To sum up, if there are different investors that all believe in long-run risks but use slightly different estimates for the long-run risk process, the investor who is more skeptical about ρ_x will dominate the economy. The investor with a larger belief about ρ_x will rapidly lose wealth, independent of whether his beliefs are correct or not. But a large ρ_x is needed to obtain a high risk premium in the long-run risk model. Even if the investor with the high belief about ρ_x almost entirely populates the economy initially, his consumption share decreases so fast, that the equity premium in the economy drops tremendously in a short amount of time. Hence, with differences in beliefs—that are likely present for real-world investors—there must be mechanisms that shift wealth to the investors with the higher beliefs about ρ_x as otherwise risk premia in the economy collapse even in small samples. On the other hand, different beliefs about ρ_x introduce variations in the consumption shares, which can in turn increase the volatility of the price-dividend ratio and brings the values closer to the level observed in the data.

Table 1: Consumption Shares: Summary Statistics

	$\rho_x = 0.985$			$\rho_x = 0.975$		
Years	100	200	500	100	200	500
$s_0^2 = 0.5$	0.7429	0.8515	0.9628	0.8904	0.9683	0.9995
	(0.0500)	(0.0481)	(0.0212)	(0.0118)	(0.0072)	(4.2e-5)
$s_0^2 = 0.05$	0.4507	0.7143	0.9393	0.5140	0.7947	0.9787
	(0.0589)	(0.0636)	(0.0293)	(0.0301)	(0.0252)	(0.0049)
$s_0^2 = 0.01$	0.2824	0.6376	0.9278	0.3404	0.7249	0.9732
	(0.0509)	(0.0681)	(0.0326)	(0.0294)	(0.0293)	(0.0060)

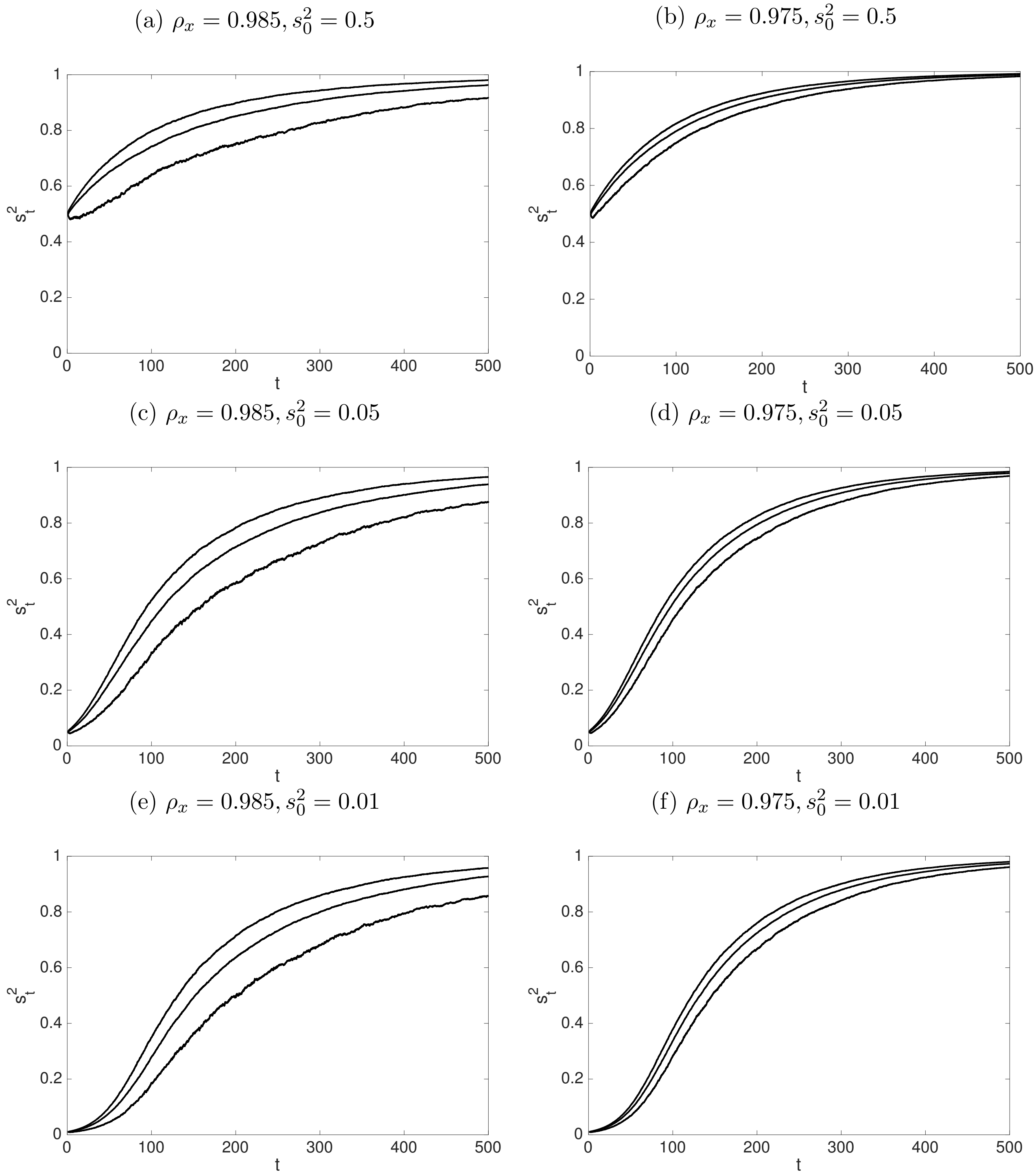
The table shows the median and the standard deviation (in parenthesis) of the consumption share of agent 2 using 1000 samples each consisting of 500 years of simulated data. Agent 2 believes that $\rho_x^2 = 0.975$ and agent 1 believes that $\rho_x^1 = 0.985$. Summary Statistics are shown for different initial consumption shares ($s_0^2 = \{0.01, 0.05, 0.5\}$) and different time periods $T = \{100, 200, 500\}$ years. The left panel depicts the case where the pessimistic agent has the right beliefs about the long-run risk process ($\rho_x = 0.985$) and in the right panel, the optimistic agent has the right beliefs ($\rho_x = 0.975$).

4.1 Optimal Consumption Decisions and Equilibrium Dynamics

In this section we analyze the different effects that determine the equilibrium allocations of the agents. For this purpose we set our results in relation to the findings of Borovička (2015). Borovička (2015) considers a simple two-agent economy with identical preferences of Epstein-Zin type and different beliefs about the mean growth rate of the economy.[6] Our model can be viewed as a generalized version of his model with time-varying beliefs about the mean growth rate in the economy. Borovička (2015) describes four channels through which the individual choices influence long-run equilibrium dynamics: the *speculative bias channel*, the *risk premium channel*, the *savings channel* and the *speculative volatility channel*. The *speculative volatility channel* only influences equilibrium outcomes for small degrees of risk aversion and has therefore a negligible influence for the results obtained in this paper. In the following we argue, that the *speculative bias channel* and the *risk premium channel* can explain the equilibrium dynamics of the long-run risk model considered in the previous section, while the *savings channel* is rather irrelevant for our model specification.

[6]In Borovička (2015) there is no long-run risk and log aggregate consumption growth is normally distributed.

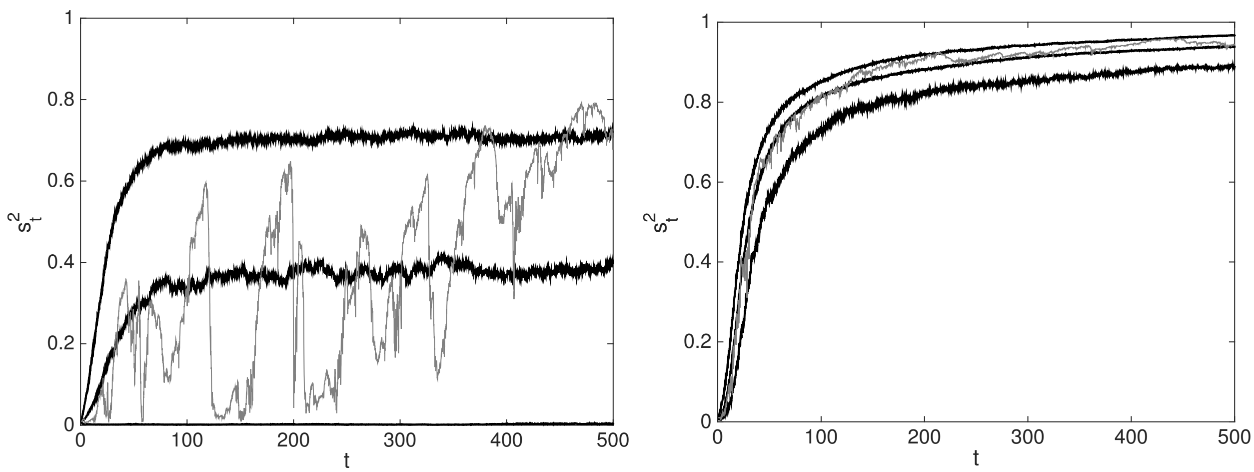
Figure 1: Consumption Shares: Simulations



The figure shows the median, 5% and 95% quantile paths of the consumption share of agent 2 for 1000 samples each consisting of 500 years of simulated data. Agent 2 believes that $\rho_x^2 = 0.975$ and agent 1 believes that $\rho_x^1 = 0.985$. Results are shown for different initial consumption shares ($s_0^2 = \{0.01, 0.05, 0.5\}$). The left panel depicts the case where the pessimistic agent has the right beliefs about the long-run risk process ($\rho_x = 0.985$) and in the right panel, the optimistic agent has the right beliefs ($\rho_x = 0.975$).

Figure 2: Consumption Shares for $\rho_x^2 = 0.95$: Simulations

(a) $\rho_x = 0.985$ (b) $\rho_x = 0.95$



The figure shows the median, 5% and 95% quantile paths of the consumption share of agent 2 for 1000 samples each consisting of 500 years of simulated data as well as a sample path (grey line). Agent 2 believes that $\rho_x^2 = 0.95$ and agent 1 believes that $\rho_x^1 = 0.985$. Results are shown for an initial consumption share of $s_0^2 = 0.01$. The left panel depicts the case where the pessimistic agent has the right beliefs about the long-run risk process ($\rho_x = 0.985$) and in the right panel, the optimistic agent has the right beliefs ($\rho_x = 0.95$).

Table 2: Equity Premium for Different Consumption Shares

	$\boldsymbol{\rho_x^2 = 0.975}$		$\boldsymbol{\rho_x^2 = 0.95}$	
	s_t^2	Equity Premium	s_t^2	Equity Premium
Rep. Agent 1	0	6.53	0	6.53
0 Years	0.01	6.42	0.01	5.42
100 Years	0.2824	4.59	0.3259	1.84
200 Years	0.6376	3.49	0.3782	1.64
500 Years	0.9278	2.89	0.4019	1.56
Rep. Agent 2	1	2.76	1	0.26

The table shows the annualized equity premium for a specific consumption share $s_t^2 = \bar{s}$. The premium is reported for the equilibrium allocations after 0, 100, 200 and 500 years of simulated data assuming an initial share of $s_0^2 = 0.01$ (see Table 1). Agent 1 has the correct beliefs with $\rho_x^1 = \rho_x = 0.985$. The left panel depicts the case for $\rho_x^2 = 0.975$ and the right panel for $\rho_x^2 = 0.95$.

Table 3: Annualized Asset Pricing Moments

	$E(p_t - d_t)$	$\sigma(p_t - d_t)$	$E\left(r_t^m - r_t^f\right)$	$E\left(r_t^f\right)$	$\sigma(r_t^m)$	$\sigma\left(r_t^f\right)$
			$\boldsymbol{\rho_x^2 = 0.975}$			
$s_t^2 = 0$	2.68	0.25	6.53	2.32	17.84	1.50
Two Agent Economy	3.10	0.29	3.98	2.58	17.19	1.51
$s_t^2 = 1$	3.29	0.20	2.83	2.71	16.55	1.53
			$\boldsymbol{\rho_x^2 = 0.95}$			
$s_t^2 = 0$	2.68	0.25	6.53	2.32	17.84	1.50
Two Agent Economy	3.60	0.48	2.63	2.47	20.37	1.58
$s_t^2 = 1$	6.27	0.14	0.26	2.93	14.80	1.52

The table shows the population moments from 1000 samples each containing 500 years of simulated data starting with an initial share of $s_0^2 = 0.01$. It shows the mean and the standard deviation of the annualized log price-dividend ratio, the annualized market over the risk-free return and the risk-free return. Agent 1 has the correct beliefs with $\rho_x^1 = \rho_x = 0.985$. The left panel depicts the case for $\rho_x^2 = 0.975$ and the right panel for $\rho_x^2 = 0.95$. All returns are shown in percent, so a value of 1.5 is a 1.5% annualized figure.

4.1.1 The Speculative Bias Channel

The *speculative bias channel* solely determines equilibrium outcomes in the special case of CRRA preferences. The investors assign different subjective probabilities to future states and buy assets that pay off in states they belief are more likely. Hence, for CRRA utility the agent with the more correct beliefs will accumulate wealth in the long-run, as the investor with the more distorted beliefs bets on states that have a vanishing probability under the true probability measure.

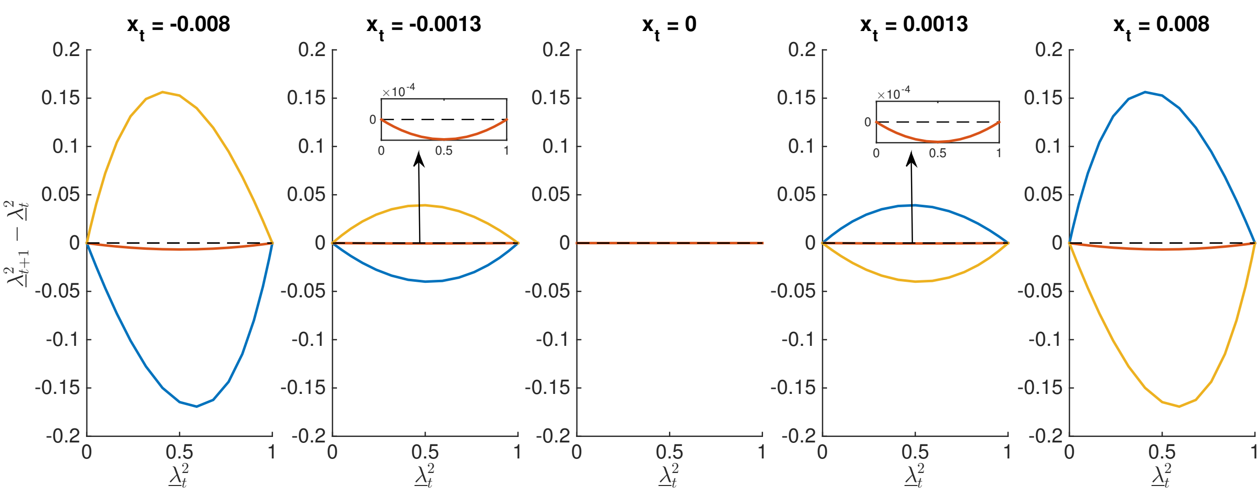
To demonstrate how the *speculative bias channel* affects equilibrium outcomes in the long-run risk model with different beliefs, we first consider the special case of CRRA preferences. In Figure 3 we show the change in the Pareto weights $\underline{\lambda}_{t+1}^2 - \underline{\lambda}_t^2$ as a function of $\underline{\lambda}_t^2$. The blue and yellow lines depict the cases of a negative shock ($x_{t+1} - \rho_x x_t = -0.001$) and a positive shock ($x_{t+1} - \rho_x x_t = 0.001$) in x_{t+1} respectively. The red line shows the average over all shocks. From left to right, the results are shown for for $x_t = -0.008, x_t = -0.0013, x_t = 0, x_t = 0.0013$ and $x_t = 0.008$. Agent 1 has the correct beliefs $\rho_x^1 = \rho_x = 0.985$ while agent 2 believes that $\rho_x^2 = 0.975$.

The second agent believes that x_t converges faster to its long-run mean compared to agent 1. Hence, if $x_t < 0$ he assigns larger probabilities to large x_{t+1} and bets on those states as $\rho_x^2 x_t > \rho_x^1 x_t$ (left panels). The opposite holds true for $x_t > 0$. So agent 2 loses wealth if x_t is low and the shock in x_t is negative (blue line in the left figures) or if x_t is high and the shock in x_t is also high (yellow line in the right figures). Taking the average over all future realization of x_{t+1} (red line), agent 2 loses wealth on average (red line). For $x_t = 0$ both agents share the same beliefs ($\rho_x^2 x_t = \rho_x^1 x_t$) and hence they assign the same probabilities to x_{t+1} (red and blue line coincide with the red line). As agent 2 loses wealth on average for all x_t except for $x_t = 0$, he will eventually diminish in the long-run. Note that the influence of the *speculative bias channel* becomes stronger, the larger $|x_t|$, as the belief dispersion grows the more x_t deviates from its unconditional mean $E(x_t) = 0$.

The *speculative bias channel* can be directly related to the two sets of results in Section 4. Results are shown for the case where agent 1 has the correct beliefs ($\rho_x = \rho_x^1$) as well as for the case where agent 2 has the correct beliefs ($\rho_x = \rho_x^2$). In the first case, the *speculative bias channel* works in favor of agent 1, while in the second case, it works in favor of agent 2. Hence, in case two, the consumption share of agent 2 increases more rapidly, as the *speculative bias channel* works in his favor (see Figure 1).

The *speculative bias channel* entirely determines the equilibrium in the standard case of CRRA preferences. For general Epstein-Zin preferences equilibrium dynamics become more complex. In the following we first describe the general effects of the *risk premium channel* and then analyze how the two effects interact and influence equilibrium outcomes.

Figure 3: Changes in the Wealth-Distribution—The CRRA Case



The figure shows the change in the optimal weights $\lambda_{t+1}^2 - \lambda_t^2$ as a function of λ_t^2. From left to right, the change is shown for $x_t = \{-0.008, -0.0013, 0, 0.0013, 0.008\}$ ($\pm$ 4 standard deviations). The blue line depicts the case of a negative shock in x_{t+1} ($x_{t+1} - \rho_x x_t = -0.001$) and the yellow line of a positive shock in x_{t+1} ($x_{t+1} - \rho_x x_t = 0.001$). The red line shows the average over all shocks. Baseline calibration with $\rho_x = \rho_x^1$ and CRRA preferences.

4.1.2 The Risk Premium Channel

With Epstein-Zin preferences, risk-return trade-offs are not the same among agents and optimistic agents are willing to take larger risks (see Borovička (2015)). So if risk aversion, and hence risk premia, are high, more optimistic agents will profit from investing in a portfolio with a higher average return. Borovička (2015) calls this the *risk premium channel*. In our model we can't specify optimists or pessimists as the beliefs about the mean growth rate change over time (see Section 3). We rather refer to agents who are skeptical about long-run risks, that is, they have a lower belief about ρ_x. Skepticism implies that the agent is less afraid of long-run risks. An investor who beliefs in a large ρ_x is afraid of large negative realizations of x_t and would therefore like buy insurance against these risks. As risk premia in the economy are high due to the combination of high risk aversion, the preference of early resolution of risks and highly persistent shocks to x_t, the premium the investor is willing to pay, will be high. The skeptical investor on the other hand, will be willing to provide this insurance as he is less afraid of the long-run risks.

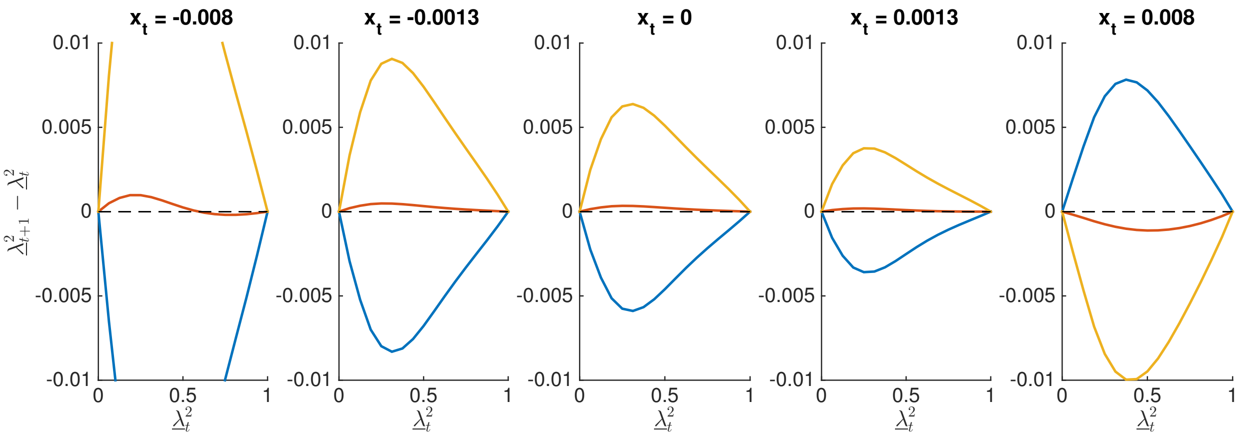
In Figures 4 we demonstrate how this channel affects model outcomes. It shows the corresponding results to Figure 3 but for the general case of Epstein-Zin preferences. First, consider the center panel where $x_t = 0$ and hence the speculative bias channel has no effect on equilibrium outcomes (see Figure 3). Agent 1 is more afraid of negative shocks to x_{t+1} compared to agent 2. Therefore he buys insurance against the long-run risks which pays of in bad times when there is a negative shock to x_{t+1} (the blue line is negative which implies an increase in the weights of the first agent for all $\underline{\lambda}_t^2$). Therefore he has to pay a premium in good times. So for a positive shock in x_{t+1} the results reverse (yellow line). The average over all shocks (red line) is positive, so he pays a positive premium to insure against long-run risk which is why he loses wealth on average. The effect is stronger for small $\underline{\lambda}_t^2$ and decreases for large $\underline{\lambda}_t^2$. A small value of $\underline{\lambda}_t^2$ implies that there is a large share of agents who wants to buy insurance against long-run risks. Hence, they are willing to pay a higher price. The larger the share of the skeptical investors becomes, the lower becomes the demand for the insurance and hence, also the increase in the Pareto weights becomes less pronounced.

Decreasing x_t has two effects. First of all, agent 1 becomes more afraid of long-run risks (given a negative value of x_t a large negative realization of x_{t+1} becomes more likely due to high persistence of ρ_x), which is why he wants to buy more insurance against long-run risks and is willing to pay a higher premium. We observe this effect in the second panel from the left ($x_t = -0.0013$) where the average increase in the Pareto weight of the second agent (red line) increases compared to the results for $x_t = 0$. Additionally, the belief difference and hence the difference between the subjective probabilities becomes more pronounced for large

$|x_t|$. So the influence of the *speculative bias channel* becomes stronger, the further x_t is away from its unconditional mean. This potentially shifts wealth to the first agent who has the correct beliefs about ρ_x. We observe this pattern in the left panel ($x_t = -0.008$) where for large $\underline{\lambda}_t^2$ the average change in the weights $\underline{\lambda}_{t+1}^2 - \underline{\lambda}_t^2$ becomes negative. For positive x_t agent 1 becomes less afraid of long-run risks and hence, he is only willing to pay less to insure against these risk. Therefore, the average increase in the weights of agent 2 decrease for $x_t = 0.0013$ compared to $x_t = 0$. For very large x_t (right panel) the influence of the *speculative bias channel* dominates and hence the results reverse. The second agent wins, if there is a negative shock (blue line), but loses, if there is a positive shock (yellow line). The *risk premium channel* becomes negligible and the second investor loses on average as he bets on states, that have a vanishing probability under the true measure (see Figure 3). So the *risk premium channel* dominates the *speculative bias channel* for x_t close to its unconditional mean, and only for very large x_t the *speculative bias channel* dominates and agent 2 potentially loses wealth on overage. However, values of $x_t = 0.008$ (+4 standard deviation of x_t) occur only very rarely and most of the time, the process stays within the range where the *risk premium channel* clearly dominates the *speculative bias channel* and hence, the consumption share of agent 2 increases on average.

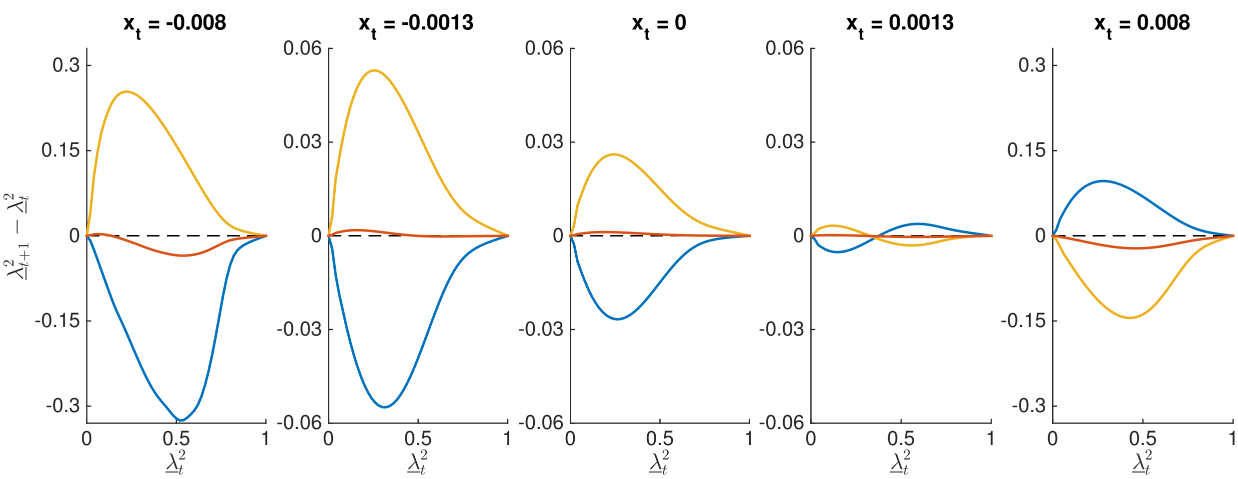
In Figure 5 we show the corresponding results for $\rho_x^2 = 0.95$ instead of $\rho_x^2 = 0.975$. This increases the influence of the *speculative bias channel* as the beliefs of the second agent are 'more wrong' on average and hence will shift wealth to the first investor. Furthermore, the second investor is less afraid of long-run risks and therefore will be willing to sell more insurance. So also the influence of *risk premium channel* increases, which on the other hand shifts wealth to the second investor. Looking at the aggregate effects, we observe that for $x_t = 0$ the change in the weights $\underline{\lambda}_{t+1}^2 - \underline{\lambda}_t^2$ becomes larger on average (note the different scale. For a better visualization we show the average change separately in Figure 10 in Appendix B). This reflects the increasing influence of the *risk premium channel* compared to the case with $\rho_x^2 = 0.975$. However, for larger $|x_t|$, the influence of the *speculative bias channel* quickly increases and only for small $\underline{\lambda}_t^2$ where there is a large share of investors who want to buy insurance against long-run risks, the *risk premium channel* dominates. This explains, why the median consumption share in Figure 2 only increase to a certain level and does not converge further towards 1. The magnitude of the change in the weights explains the large drops and recoveries that we observe in Figure 2. For example for the extreme case with $x_t = -0.008$, a large negative shock implies a drop in the weights of more than 0.3 for $\underline{\lambda}_t^2 = 0.5$. This implies a decrease in the consumption share of the second agent of more than 0.3. But as the influence of the *risk premium channel* increases for small $\underline{\lambda}_t^2$ the second investor recovers rather quickly as can be observed from Figure 2.

Figure 4: Changes in the Wealth-Distribution—The Epstein-Zin Case



The figure shows the change in the optimal weights $\underline{\lambda}_{t+1}^2 - \underline{\lambda}_t^2$ as a function of $\underline{\lambda}_t^2$. From left to right, the change is shown for $x_t = \{-0.008, -0.0013, 0, 0.0013, 0.008\}$ (± 4 standard deviations). The blue line depicts the case of a negative shock in x_{t+1} ($x_{t+1} - \rho_x x_t = -0.001$) and the yellow line of a positive shock in x_{t+1} ($x_{t+1} - \rho_x x_t = 0.001$). The red line shows the average over all shocks. Baseline calibration with $\rho_x = \rho_x^1$.

Figure 5: Changes in the Wealth-Distribution—The Epstein-Zin Case ($\rho_x^2 = 0.95$)



The figure shows the change in the optimal weights $\underline{\lambda}_{t+1}^2 - \underline{\lambda}_t^2$ as a function of $\underline{\lambda}_t^2$. From left to right, the change is shown for $x_t = \{-0.008, -0.0013, 0, 0.0013, 0.008\}$ (± 4 standard deviations). The blue line depicts the case of a negative shock in x_{t+1} ($x_{t+1} - \rho_x x_t = -0.001$) and the yellow line of a positive shock in x_{t+1} ($x_{t+1} - \rho_x x_t = 0.001$). The red line shows the average over all shocks. Calibration with $\rho_x = \rho_x^1 = 0.985$ and $\rho_x^2 = 0.95$.

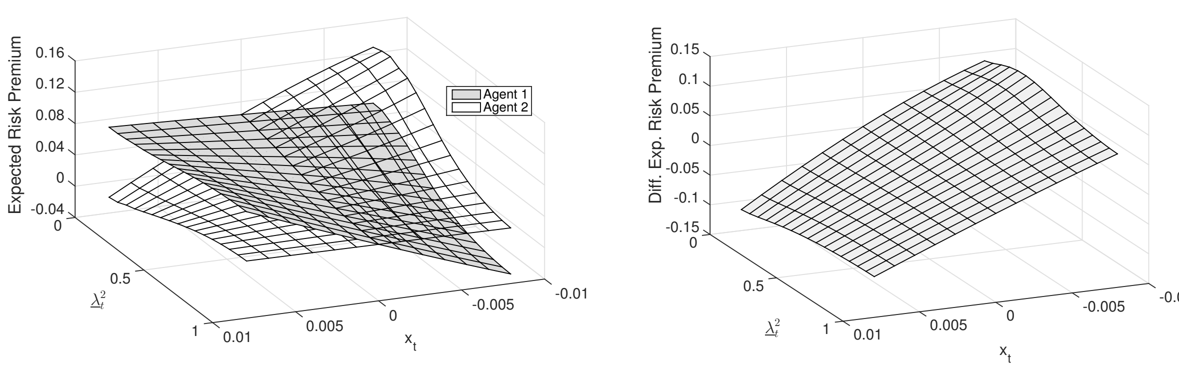
4.1.3 The Savings Channel

The third channel that influences equilibrium outcomes for Epstein-Zin preferences is the *savings channel*. It states that agents with high subjective beliefs about expected returns will choose a high (low) savings rate if the IES is large (small). In the long-run risk model the IES needs to be significantly larger than 1 in order to model a strong preference for the early resolution of risks. Hence, the agent with the higher subjective expected returns chooses a higher savings rate and therefore—all other things being equal—his consumption share increases relative to the agent with the lower expected returns.

Figure 6 shows the subjective expected risk premia of the two agents as a function of the states (Figure 6a as well as the difference between the two (Figure 6b). Agent 2 has higher subjective risk premia for small x_t and the opposite is true for large x_t. Therefore, for small (large) x_t, agent 2 will choose a higher (lower) savings rate compared to agent 1 that in turn increases (decreases) his consumption share. However, we find that in the aggregate, the influence of the *savings channel* is rather small compared to the *risk premium channel* and the *speculative bias channel*. In Figure 7 we show the corresponding results to Figure 4 but with $\psi^1 = \psi^2 = 1.1$ instead of $\psi^1 = \psi^2 = 1.5$ and hence, a smaller influence of the *savings channel* channel. We observe that the quantitative change is rather small and the qualitative conclusions stay the same.

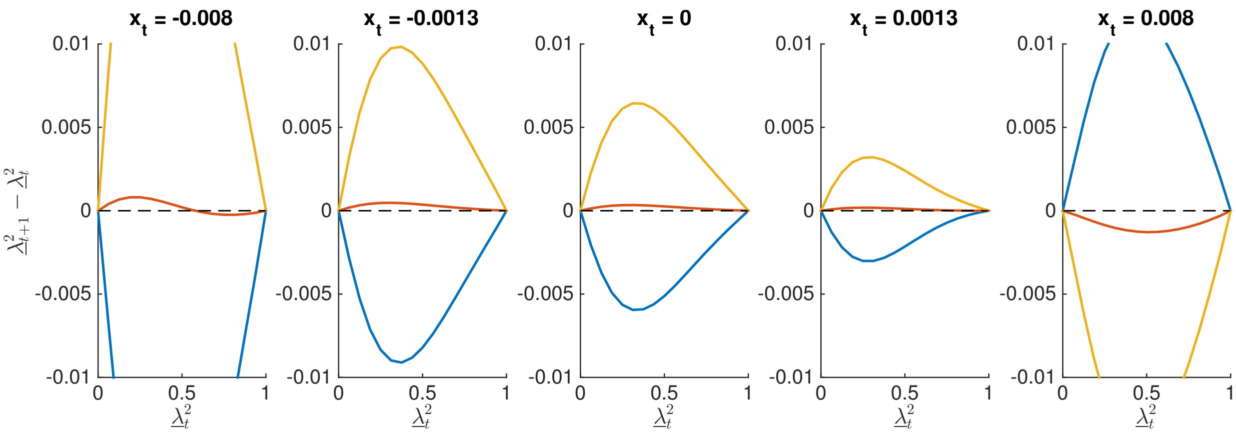
Figure 6: Expected Subjective Risk Premia

(a) (b)



The figure shows the expected subjective risk premium of the two agents as a function of the states $\underline{\lambda}_t^2$ and x_t. Panel (a) show the absolute values for the two agents and Panel (b) shows the difference between the subjective risk premium of agent 2 and agent 1.

Figure 7: Changes in the Wealth-Distribution—The Epstein-Zin Case Sensitivity ψ^h



The figure shows the change in the optimal weights $\underline{\lambda}^2_{t+1} - \underline{\lambda}^2_t$ as a function of $\underline{\lambda}^2_t$. From left to right, the change is shown for $x_t = \{-0.008, -0.0013, 0, 0.0013, 0.008\}$ ($\pm$ 4 standard deviations). The blue line depicts the case of a negative shock in x_{t+1} ($x_{t+1} - \rho_x x_t = -0.001$) and the yellow line of a positive shock in x_{t+1} ($x_{t+1} - \rho_x x_t = 0.001$). The red line shows the average over all shocks. Baseline calibration with $\rho_x = \rho^1_x$ and $\psi^1 = \psi^2 = 1.1$ (instead of 1.5 as in the baseline model).

4.2 Examination of the Risk Premium Channel (Robustness of the Results)

In this section we examine the influence of the *risk premium channel* in more detail. We have argued that, if risk premia are high, the influence of the *risk premium channel* is strong. This will in turn shift wealth to the investors who are skeptical about long-run risks. In Figure 8a we show the median consumption share of agent 2 (as in Figure 1) for different degrees of risk aversion $\gamma^h = \{2, 5, 10\}$. For $\gamma^1 = \gamma^2 = 10$ the equity premium for the representative agent economies either populated only by agent 1 or 2 are 6.53% and 2.76% respectively (see Table 2). For a risk aversion of $\gamma^1 = \gamma^2 = 5$ they decrease to 2.71% and 0.72% and for $\gamma^1 = \gamma^2 = 2$ the premia are only -0.61% and -0.68%. So for $\gamma^h = 5$ and $\gamma^h = 2$, we expect the impact of the *risk premium channel* to decrease significantly. For $\gamma^h = 10$ (yellow line) the influence of the *risk premium channel* is strong. Hence, agent 2 profits from selling the insurance against long-run risks and rapidly accumulates wealth. For $\gamma^h = 5$ (red line) this effect becomes less severe and his consumption share increases less quickly. For $\gamma^h = 2$ (blue line), where risk premia are negative the *risk premium channel* has no influence and the *speculative bias channel* becomes dominates equilibrium outcomes. As $\rho_x = \rho^1_x$ the *speculative bias channel* works in favor of agent 1 (agent 2 bets on states, that have a vanishing probability under the true probability measure) and agent 1 dominates the economy in the long-run. If agent two has the correct beliefs $\rho_x = \rho^2_x$, the *speculative bias channel* works in favor of agent 2. We show this case in Figure 8b. The blue line shows the consumption shares for $\rho_x = \rho^1_x$ and the red line for $\rho_x = \rho_x^2$. So in the absence of the *risk premium channel* the *speculative bias channel* determines equilibrium outcomes.

In Figure 8c we analyze the robustness of our findings with regard to the level of the persistences of x_t. We show the consumption paths for $\rho_x^2 = 0.6$, $\rho_x^1 = 0.5$ instead of 0.975 and 0.985. Lowering the persistence will—similarily to the decrease in risk aversion—bring down the equity premium to -0.74%. Consequently, we observe that in this setup the dynamics of the consumption shares strongly depend on the true value of ρ_x as the *speculative bias channel* dominates—that is, the agent with the correct beliefs will dominate the economy.

But long-run risk models require a high degree of risk aversion and a high persistence level of the long-run risk process in order to obtain an equity premium consistent with the data. Consequently, the impact of the *risk premium channel* will be strong and the investors who are skeptical about long-run risks will dominate the economy. The qualitative implications also hold irrespectively of the true value of the underlying persistence of the long-run risk process. In Figure **??** we show the consumption paths with $\rho_x^1 = 0.985$ and $\rho_x^2 = \rho_x = 0.975$ for different values of $\rho_x = \{0, 0.9, 0.99\}$. A lower persistence of ρ_x implies that x_t will remain closer to its unconditional mean (given the same standard deviation). As Figure 4 shows, for x_t close to 0, the consumption share of the second agent increases on average. Hence, the lower the true persistence, the faster the increase in the consumption share. But even for the very large value of ρ_x of 0.99, the risk premium effect still dominates and the second agent dominates the economy in the long-run.

4.3 Correcting for the Difference in Mean Consumption Growth

Different beliefs about the persistence of the long-run risks process imply that—all other equal—the agent also has different beliefs about the mean of the gross growth rate of consumption $E\left(\frac{C_{t+1}}{C_t}\right)$ due to Jensen's inequality. In this section we argue that our results are not driven by this simple mean effect, but rather by the time varying *risk premium channel* as demonstrated in the previous section. In fact, when we correct for the belief difference in the mean growth rate of consumption, the consumption share of the skeptical investor increases even faster. For the long-run risks model (19), the mean growth rate of consumption is given by

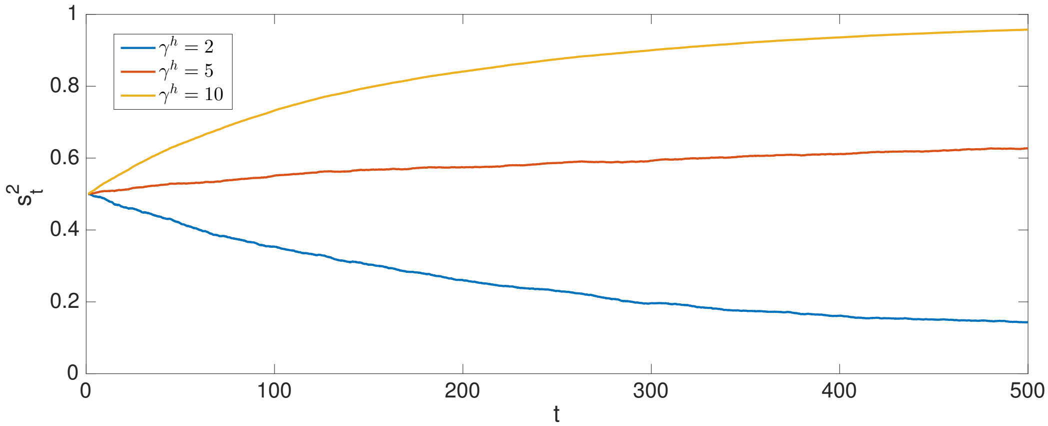
$$E\left(\frac{C_{t+1}}{C_t}\right) = E\left(e^{\Delta c_{t+1}}\right) = e^{\mu_c + 0.5\sigma^2 + 0.5\frac{\phi_x^2\sigma^2}{1-(\rho_x)^2}}. \tag{20}$$

For $\rho_x^2 < \rho_x^1 = \rho_x$ we have that

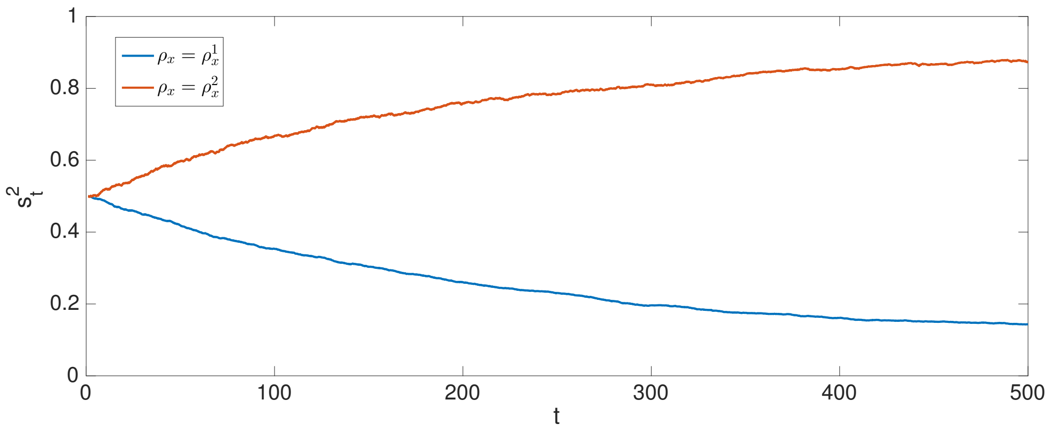
$$E^2\left(\frac{C_{t+1}}{C_t}\right) = e^{\mu_c + 0.5\sigma^2 + 0.5\frac{\phi_x^2\sigma^2}{1-(\rho_x^2)^2}} < E\left(\frac{C_{t+1}}{C_t}\right). \tag{21}$$

Figure 8: The Risk Premium and Speculative Bias Channels

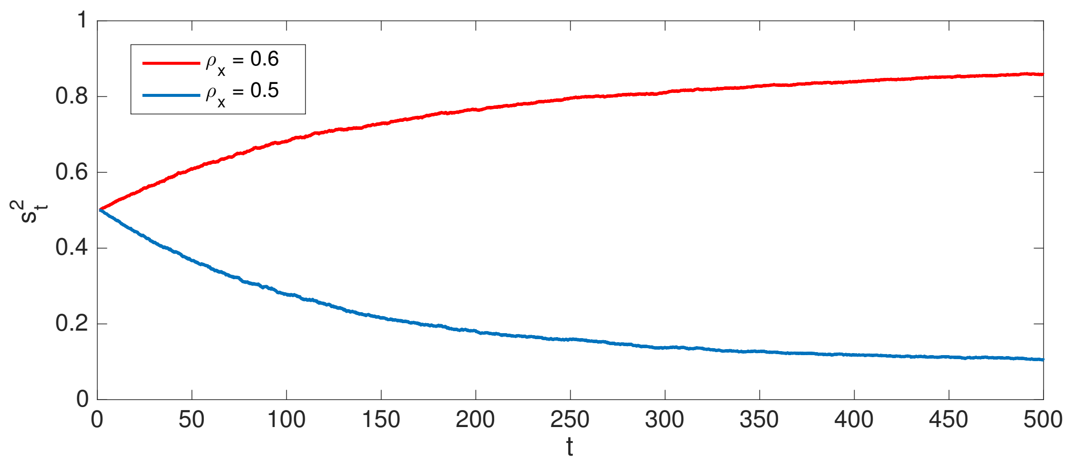
(a) $\rho_x^1 = 0.985$, $\rho_x^2 = \rho_x = 0.975$



(b) $\gamma^h = 2$, $\rho_x^1 = 0.985$, $\rho_x^2 = 0.975$

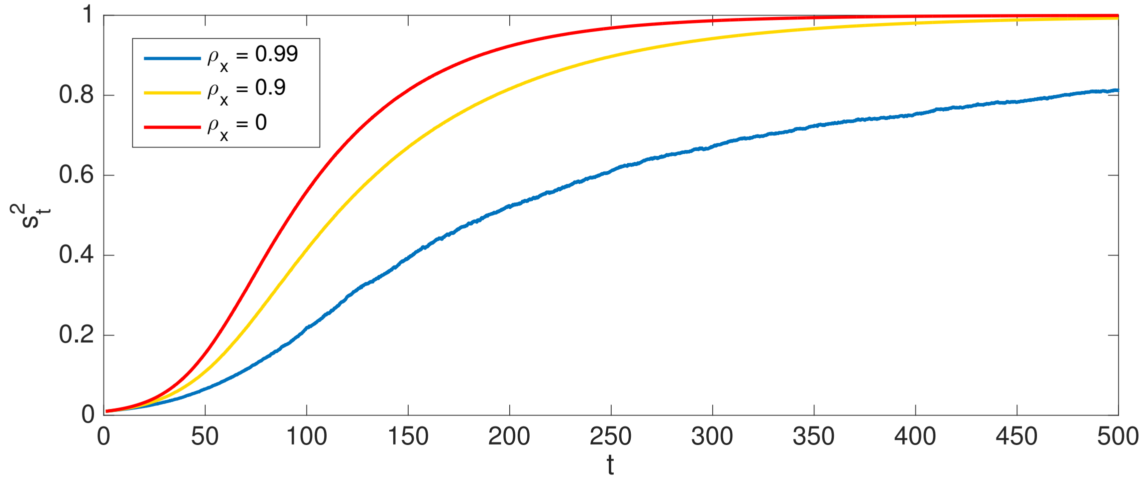


(c) $\gamma^h = 10$, $\rho_x^2 = 0.6$, $\rho_x^1 = 0.5$



The figure shows the median consumption share of agent 2 for 1000 samples each consisting of 500 years of simulated data. The Panel (a) shows the time series for different degrees of risk aversion $\gamma^h = \{2, 5, 10\}$. Agent 2 believes that $\rho_x^2 = 0.975$ and agent 1 has the correct beliefs with $\rho_x^1 = \rho_x = 0.985$. Panel (b) shows the time series for $\gamma^h = 2$, $\rho_x^1 = 0.985$ $\rho_x^2 = 0.975$ for the two cases where either agent 1 (blue line) or agent 2 (red line) has the correct beliefs. Panel (c) shows the time series for $\gamma^h = 10$, $\rho_x^1 = 0.6$ and $\rho_x^2 = 0.5$ for the two cases where either agent 1 (blue line) or agent 2 (red line) has the correct beliefs.

Figure 9: The Risk Premium and Speculative Bias Channels (2)



The figure shows the median consumption share of agent 2 for 1000 samples each consisting of 500 years of simulated data. Both agents have a risk aversion of $\gamma = 10$. The results are shown for $\rho_x^2 = 0.975$ and $\rho_x^1 = 0.985$ for different values of $\rho_x = \{0, 0.9, 0.99\}$.

So the second agent beliefs in a lower mean growth rate of consumption as he beliefs in a lower persistence and hence a lower unconditional volatility of the long-run risk process. We correct for this belief differences by setting the subjective belief of the second investor about mean log consumption growth to $\mu_c^2 = \mu_c + 0.5\frac{\phi_x^2\sigma^2}{1-(\rho_x)^2} - 0.5\frac{\phi_x^2\sigma^2}{1-(\rho_x^2)^2}$. Once we correct for this difference, the consumption shares of the skeptical investor increase even faster. For the original specification with an initial allocation of $s_0^2 = 0.01$ the consumption shares of the skeptical investor increased to 0.2824, 0.6376 and 0.9278 after 100,200 and 500 years respectively (see Table 1). With the corrected mean we obtain values of 0.2827, 0.6379 and 0.9281. Hence, our results are not driven by the effect of different mean beliefs about consumption growth. This result is also in line with Borovička (2015) who shows that an underestimation of the mean growth rate lowers the chances of survival while the overestimation has the opposite effect due to the positive risk premium channel. Consequently, in our model specification, the effect from the mean growth rate should lead to lower consumption shares of the skeptical investor. And indeed, once we correct the mean of the skeptical investor we obtain a faster increase in the consumption shares of the skeptical investor.

5 Conclusion

We have performed a detailed study of heterogenity in agent beliefs for the long-run risk model of Bansal and Yaron (2004). In particular, we consider agents with different beliefs about the level of persistence in long-run risk. We find that as long as the level of heterogeneity is not too large, agents who believe in a lower level of persistence come to dominate the economy rather quickly relative to agents who believe in a higher level of persistence. This holds even if the agent with the higher level of persistence holds the correct belief. This suggests that for long-run risk to work as an explanation of the equity premium, it is not sufficient for consumption suffer from long-run risk— all agents must also agree on the amount of long-run risk the economy experiences.

Appendix

A Proofs and Details

A.1 Solution Method for Asset Pricing Models with Heterogeneous Agents and Recursive Preferences

Proof of Theorem 1. Let $\boldsymbol{\lambda} = \{\lambda^1, \ldots, \lambda^H\}$, and $\{\boldsymbol{C}\}_0 = \{\{C^1\}_0, \ldots, \{C^H\}_0\}$. The optimal decision of the social planner in the initial period takes into account all future consumption streams of the individual agents and the optimal decisions must satisfy the market clearing condition (1). For the ease of notation we abbreviate the state dependence in the following, so we use C_t^h for $C^h(y^t)$ and $U_{\{t\}}^h$ for $U^h\left(\{C^h\}_t\right)$.

To derive the first-order conditions, we borrow a technique from the calculus of variations. For any function f_t we can vary the consumption of two agents by

$$\begin{aligned} C_t^h &\rightarrow C_t^h + \epsilon f_t \\ C_t^l &\rightarrow C_t^l - \epsilon f_t. \end{aligned} \tag{22}$$

It is sufficient to consider the variation with agent $l = 1$. Since we have an optimal allocation it must be true that

$$\left. \frac{\mathrm{d}SP(\{\boldsymbol{C}\}_0, \boldsymbol{\lambda})}{\mathrm{d}\epsilon} \right|_{\epsilon=0} = 0. \tag{23}$$

This gives us

$$\lambda^h \hat{U}_{0,t}^h = \lambda^1 \hat{U}_{0,t}^1, \qquad h \in \mathbb{H}^-, \tag{24}$$

where $\hat{U}_{t,t+k}^h$ is defined as

$$\hat{U}_{t,t+k}^h = \left. \frac{\mathrm{d}U^h(C_t^h, \ldots, C_{t+k}^h + \epsilon f_{t+k}, \ldots)}{\mathrm{d}\epsilon} \right|_{\epsilon=0}. \tag{25}$$

$\hat{U}_{t,t+k}^h$ satisfies a recursive equation with the initial condition

$$\hat{U}_{t,t}^h = \left. \frac{\mathrm{d}U^h(C_t^h + \epsilon f_t, \ldots)}{\mathrm{d}\epsilon} \right|_{\epsilon=0} = F_1^h\left(C_t^h, R_t[U_{\{t+1\}}^h]\right) \cdot f_t \tag{26}$$

where $F_k^h\left(C_t^h, R_t^h[U_{\{t+1\}}^h]\right)$ denotes the derivative of $F^h\left(C_t^h, R_t^h[U_{\{t+1\}}^h]\right)$ with respect to its

kth argument. The recursive step is given by

$$
\begin{aligned}
\hat{U}^h_{t,t+k} &= \left.\frac{\mathrm{d}F^h\left(C^h_t, R^h_t\left[U^h(C^h_{t+1},\ldots C^h_{t+k}+\epsilon f_{t+k},\ldots)\right]\right)}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \\
&= F^h_2\left(C^h_t, R^h_t[U^h_{\{t+1\}}]\right)\cdot\left.\frac{\mathrm{d}R^h_t\left[U^h(\cdot)\right]}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \\
&= F^h_2\left(C^h_t, R^h_t[U^h_{\{t+1\}}]\right)\cdot\frac{\mathrm{d}G^{-1}_h\left(E^h_t G_h\left[U^h(\cdot)\right]\right)}{\mathrm{d}E^h_t G_h[U^h(\cdot)]}\cdot\left.\frac{\mathrm{d}E^h_t G_h[U^h(\cdot)]}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \\
&= F^h_2\left(C^h_t, R^h_t[U^h_{\{t+1\}}]\right)\cdot\frac{1}{G'_h(G^{-1}_h(E^h_t G_h[U^h_{\{t+1\}}]))}\cdot E^h_t\left(G'_h(U^h_{\{t+1\}})\cdot\hat{U}^h_{t+1,t+k}\right) \\
&= F^h_2\left(C^h_t, R^h_t[U^h_{\{t+1\}}]\right)\cdot\frac{E^h_t\left(G'_h(U^h_{\{t+1\}})\cdot\hat{U}^h_{t+1,t+k}\right)}{G'_h(R^h_t[U^h_{\{t+1\}}])} \qquad (27)
\end{aligned}
$$

where we use $\frac{\partial G^{-1}(x)}{\partial x}=\frac{1}{G'(G^{-1}(x))}$ and abbreviate $U^h(C^h_{t+1},\ldots C^h_{t+k}+\epsilon f_{t+k},\ldots)$ by $U^h(\cdot)$. We can recast this recursion into a useful form. Therefore we define a second recursion $U^h_{t,t+k}$ by

$$U^h_{t,t} = F^h_1\left(C^h_t, R^h_t[U^h_{\{t+1\}}]\right) \qquad (28)$$

and

$$U^h_{t,t+k} = \Pi^h_{t+1}\cdot U^h_{t+1,t+k} \qquad (29)$$

where

$$\Pi^h_{t+1} = F^h_2\left(C^h_t, R^h_t[U^h_{\{t+1\}}]\right)\cdot\frac{G'_h(U^h_{\{t+1\}})}{G'_h(R^h_t[U^h_{\{t+1\}}])}\frac{\mathrm{dP}^h_{t,t+1}}{\mathrm{dP}_{t,t+1}}. \qquad (30)$$

A simple induction shows that

$$\hat{U}^h_{t,t+k} = E_t(U^h_{t,t+k}f_t). \qquad (31)$$

Plugging (31) into the optimality condition (24) we get

$$E_0((\lambda^h U^h_{0,t} - \lambda^1 U^1_{0,t})f_t) = 0, \qquad h\in\mathbb{H}^-. \qquad (32)$$

Under a broad range of regularity conditions, this implies that

$$\lambda^h U^h_{0,t} = \lambda^1 U^1_{0,t}, \qquad h\in\mathbb{H}^-. \qquad (33)$$

For example, if $\lambda^h U^h_{0,t} - \lambda^1 U^1_{0,t}$ has finite variance, then this holds for the Riesz Representation Theorem for L^2 random variables.

We can then split expression (33) into two parts. First define $\lambda^h_0 \equiv \lambda^h$ to obtain

$$\frac{\lambda_0^h}{\lambda_0^1} = \frac{U_{0,t}^1}{U_{0,t}^h} = \frac{\Pi_0^1}{\Pi_0^h}\frac{U_{1,t}^1}{U_{1,t}^h} = \frac{\Pi_0^1}{\Pi_0^h}\frac{\lambda_1^h}{\lambda_1^1}, \qquad h \in \mathbb{H}^-,$$

where λ_1^h denotes the Negishi weight of the social planner's optimum in $t = 1$. Generalizing this equation for any period t, we obtain the following dynamics for the optimal weight λ_{t+1}^h

$$\frac{\lambda_{t+1}^h}{\lambda_{t+1}^1} = \frac{\Pi_{t+1}^h}{\Pi_{t+1}^1}\frac{\lambda_t^h}{\lambda_t^1}, \qquad h \in \mathbb{H}^-. \tag{34}$$

Note that we can either solve the model in terms of the ratio $\frac{\lambda_t^h}{\lambda_t^1}$ (this is equal to setting $\lambda_t^1 = 1 \forall t$ as done in Judd, Kubler, and Schmedders (2003)) or we can normalize the weights so that they are bounded between $(0, 1)$. We later propose a solution method that uses the latter approach as it obtains better numerical properties.

The second expression is obtained by inserting the initial condition (28) into (33) for $t = 0$ and generalizing it for any social planner's optimum at time t:

$$\lambda_t^h F_1^h\left(C_t^h, R_t^h[U_{\{t+1\}}^h]\right) = \lambda_t^1 F_1^1\left(C_t^1, R_t^1[U_{\{t+1\}}^1]\right), \qquad h \in \mathbb{H}^-. \tag{35}$$

Equation (35) states the optimality conditions for the individual consumption choices at any time t. Note that for time separable utility, $F_1^h\left(C_t^h, R_t^h[U_{\{t+1\}}^h]\right)$ is simply marginal utility of agent h at time t, and so we obtain the same optimality condition as for example in Judd, Kubler, and Schmedders (2003) (compare equation (7) on page 2209). In this special case the Negishi weights can be pinned down in the initial period and thereafter remain constant. For general recursive preferences this is not true. The optimal weights vary over time following the law of motion described by equation (34).

□

We can use the two equations (34) and (35) together with the market clearing condition (1) to compute the social planner's optimum. We therefore define $\boldsymbol{\lambda}_t^- = \{\lambda_t^2, \lambda_t^3, \ldots, \lambda_t^H\}$ and let V^h denote the value function of agent $h \in \mathbb{H}$. We are looking for model solutions of the form $V^h(\boldsymbol{\lambda}_t^-, y^t)$. So additional to the exogenous states y^t, the model solution depends on the time varying Negishi weights $\boldsymbol{\lambda}_t^-$. An optimal allocation is then characterized by the following four equations:

- the market clearing condition (1)

$$\sum_{h=1}^{H} C^h(\boldsymbol{\lambda}_t^-, y^t) = C(y^t) \tag{36}$$

- the optimality conditions (35) for the individual consumption decisions

$$\lambda_t^h F_1^h \left(C^h(\boldsymbol{\lambda_t^-}, y^t), R_t^h[V^h(\boldsymbol{\lambda_{t+1}^-}, y^{t+1})] \right) = \lambda_t^1 F_1^1 \left(C^1(\boldsymbol{\lambda_t^-}, y^t), R_t^1[V^1(\boldsymbol{\lambda_{t+1}^-}, y^{t+1})] \right) \quad (37)$$

  for $h \in \mathbb{H}^-$

- the value functions (2) of the individual agents

$$V^h(\boldsymbol{\lambda_t^-}, y^t) = F^h \left(C^h(\boldsymbol{\lambda_t^-}, y^t), R_t^h[V^h(\boldsymbol{\lambda_{t+1}^-}, y^{t+1})] \right), \qquad h \in \mathbb{H} \quad (38)$$

- the equations (34) for the dynamics of $\boldsymbol{\lambda_t^-}$

$$\frac{\lambda_{t+1}^h}{\lambda_{t+1}^1} = \frac{\Pi_{t+1}^h}{\Pi_{t+1}^1} \frac{\lambda_t^h}{\lambda_t^1}, \qquad h \in \mathbb{H}^-, \quad (39)$$

  where

$$\Pi_{t+1}^h = F_2^h \left(C^h(\boldsymbol{\lambda_t^-}, y^t), R_t^h[V^h(\boldsymbol{\lambda_{t+1}^-}, y^{t+1})] \right) \cdot \frac{G_h'(V^h(\boldsymbol{\lambda_{t+1}^-}, y^{t+1}))}{G_h'(R_t^h[V^h(\boldsymbol{\lambda_{t+1}^-}, y^{t+1})])} \frac{\mathrm{dP}_{t,t+1}^h}{\mathrm{dP}_{t,t+1}} \quad (40)$$

This concludes the general description of the equilibrium obtained from the social planner's optimization problem.

To prove theorem 2, we derive a variant of lemma 1 in Blume and Easley (2006).

Lemma 1. *Let X_t^i be a family of H non-negative random variables for each t, such that $A \leq \sum_i X_i \leq B$. Let f and g be decreasing functions such that $f(x), g(x) \to \infty$ as $x \to 0$. If $f(X_t^j)/g(X_t^i) \to \infty$ then $X_t^i \to 0$. If $X_t^i \to 0$, then for at least one j, $\limsup_t f(X_t^j)/g(X_t^i) = \infty$.*

Proof. Since X_i is positive, $X_i \leq B$. By assumption, $g(B) \leq g(X_i)$, and $1/g(X_i) \leq 1/g(B)$. So $f(X_j)/g(X_i) \to \infty$ if and only if $f(X_j) \to \infty$ which happens when $X_j \to 0$.

Conversely, assume $X_t^i \to 0$. Every period, for at least one j, $X_j \geq A/H$ (otherwise $\sum X_{j=1}^H < A$). Since there are only finitely many random variables, for at least one j we have $X_j \geq A/H$ infinitely often. By assumption, $f(X_j) > f(A/H)$, so $\limsup f(X_j)/g(X_i) = \infty$. □

Proof of Theorem 2. By the first-order condition, equation ((5)), $\lambda_t^j/\lambda_t^i = F_1^j(C_1^j, R_1^j)/F_1^i(C_1^j, R_1^j)$. Since F^h is additively separable, F_1^h is a function of consumption alone. Let $f = F_1^j, g = F_1^i$, $A = \underline{C}$ and $B = \overline{C}$, and apply lemma 1. $\square$

A.2 The Case of Epstein-Zin Preferences

In this section we provide the specific expressions for V^h, F_1^h, F_2^h and Π^h when the heterogeneous investors have recursive preferences as in Epstein and Zin (1989) and Weil (1989). The value function for Epstein-Zin (EZ) preferences is given by[7]

$$V_t^h = \left[(1-\delta^h)(C_t^h)^{\rho^h} + \delta^h R_t^h\left(V_{t+1}^h\right)^{\rho^h}\right]^{\frac{1}{\rho^h}} \tag{41}$$

where

$$\begin{aligned} R_t^h\left(V_{t+1}^h\right) &= G_h^{-1}\left(E_t^h\left[G_h(V_{t+1}^h)\right]\right) \\ G_h(V_{t+1}^h) &= \left(V_{t+1}^h\right)^{\alpha^h}. \end{aligned}$$

The parameter δ^h is the discount factor, $\rho^h = 1-\frac{1}{\psi^h}$ determines the intertemporal elasticity of substitution ψ^h and $\alpha^h = 1-\gamma^h$ determines the relative risk aversion γ^h of agent h.

Using this notation the derivatives of $F^h\left(C_t^h, R_t^h[U_{t+1}^h]\right) = V_t^h$ with respect to its first and second argument are then given by

$$F_{1,t}^h = (1-\delta^h)(C_t^h)^{\rho^h-1}(V_t^h)^{1-\rho^h} \tag{42}$$

and

$$F_{2,t}^h = \delta^h R_t^h\left(V_{t+1}^h\right)^{\rho^h-1}(V_t^h)^{1-\rho^h}. \tag{43}$$

In this paper we focus on growth economies. Therefore we introduce the following normalization to obtain a stationary formulation of the model. We define the consumption share of agent h by $s_t^h = \frac{C_t^h}{C_t}$ and we define the normalized value functions $v_t^h = \frac{V_t^h}{C_t}$. The value function (41) is then given by

$$v_t^h = \left[(1-\delta^h)(s_t^h)^{\rho^h} + \delta^h R_t^h\left(v_{t+1}^h e^{\Delta c_{t+1}}\right)^{\rho^h}\right]^{\frac{1}{\rho^h}}. \tag{44}$$

By inserting (42) into (37) we obtain the optimality condition for the individual consumption decisions:

[7] For the ease of notation we again abbreviate the dependence on the exogenous state y^t and the endogenous state $\boldsymbol{\lambda}_t^-$. Hence we write V_t^h for $V^h(\boldsymbol{\lambda}_t^-, y^t)$ or C_t^h for $C^h(\boldsymbol{\lambda}_t^-, \boldsymbol{y}^t)$.

$$\begin{aligned} \lambda_t^h F_1^h\left(C^h(\boldsymbol{\lambda_t^-}, y^t), R_t^h[V^h(\boldsymbol{\lambda_{t+1}^-}, y^{t+1})]\right) &= \lambda_t^1 F_1^1\left(C^1(\boldsymbol{\lambda_t^-}, y^t), R_t^1[V^1(\boldsymbol{\lambda_{t+1}^-}, y^{t+1})]\right) \\ \lambda_t^h(1-\delta^h)(C_t^h)^{\rho^h-1}(V_t^h)^{1-\rho^h} &= \lambda_t^1(1-\delta^1)(C_t^1)^{\rho^1-1}(V_t^1)^{1-\rho^1} \end{aligned} \tag{45}$$

In the following we define the detrended weights, $\underline{\lambda}_t^h = \frac{\lambda_t^h}{(v_t^h)^{\rho^h-1}}$. From equation (45) we get

$$\underline{\lambda}_t^h(1-\delta^h)(s_t^h)^{\rho^h-1} = \underline{\lambda}_t^1(1-\delta^1)(s_t^1)^{\rho^1-1}. \tag{46}$$

This is the optimality condition for the individual consumption decisions we are going to use for the model with Epstein-Zin preferences. Inserting the detrended weight $\underline{\lambda}_t^h$ into the dynamics for the weights (39) we obtain

$$\frac{\lambda_{t+1}^h}{\lambda_{t+1}^1} = \frac{\underline{\lambda}_{t+1}^h(v_{t+1}^h)^{\rho^h-1}}{\underline{\lambda}_{t+1}^1(v_{t+1}^1)^{\rho^1-1}} = \frac{\underline{\lambda}_t^h(v_t^h)^{\rho^h-1}}{\underline{\lambda}_t^1(v_t^1)^{\rho^1-1}}\frac{\Pi_{t+1}^h}{\Pi_{t+1}^1}, \qquad h \in \mathbb{H}^- \tag{47}$$

Plugging the expressions for Epstein-Zin preferences (41)–(43) into equation (40), we obtain the following expression for Π_{t+1}^h:

$$\begin{aligned} \Pi_{t+1}^h &= \delta^h R_t^h\left(V_{t+1}^h\right)^{\rho^h-1}(V_t^h)^{1-\rho^h}\frac{\left(V_{t+1}^h\right)^{\alpha^h-1}}{R_t^h\left(V_{t+1}^h\right)^{\alpha^h-1}}\frac{\mathrm{dP}_{t,t+1}^h}{\mathrm{dP}_{t,t+1}} \\ &= \delta^h(V_t^h)^{1-\rho^h}\frac{\left(V_{t+1}^h\right)^{\alpha^h-1}}{R_t^h\left(V_{t+1}^h\right)^{\alpha^h-\rho^h}}\frac{\mathrm{dP}_{t,t+1}^h}{\mathrm{dP}_{t,t+1}}. \end{aligned} \tag{48}$$

Using the normalized value function $v_t^h = \frac{V_t^h}{C_t}$ we have

$$\Pi_{t+1}^h = \delta^h(v_t^h)^{1-\rho^h}\frac{\left(v_{t+1}^h e^{\Delta c_{t+1}}\right)^{\alpha^h-1}}{R_t^h\left(v_{t+1}^h e^{\Delta c_{t+1}}\right)^{\alpha^h-\rho^h}}\frac{\mathrm{dP}_{t,t+1}^h}{\mathrm{dP}_{t,t+1}}. \tag{49}$$

Equation (47) can then be written as

$$\frac{\underline{\lambda}_{t+1}^h}{\underline{\lambda}_{t+1}^1} = \frac{\underline{\lambda}_t^h}{\underline{\lambda}_t^1}\frac{\underline{\Pi}_{t+1}^h}{\underline{\Pi}_{t+1}^1}, \qquad h \in \mathbb{H}^- \tag{50}$$

where

$$\underline{\Pi}_{t+1}^h = \underbrace{\delta^h e^{\rho^h\Delta c_{t+1}}\frac{\mathrm{dP}_{t,t+1}^h}{\mathrm{dP}_{t,t+1}}}_{\text{CRRA-Term}}\underbrace{\frac{\left(v_{t+1}^h e^{\Delta c_{t+1}}\right)^{\alpha^h-\rho^h}}{R_t^h\left(v_{t+1}^h e^{\Delta c_{t+1}}\right)^{\alpha^h-\rho^h}}}_{\text{New EZ-Term}}. \tag{51}$$

For $\alpha^h = \rho^h$, we obtain the standard term for CRRA preferences where the dynamics of $\underline{\lambda}_{t+1}^h$ only depend on the subjective discount factor, the IES and the subjective beliefs of the investors. For Epstein-Zin preferences, we obtain an extra term that reflects the time trade-off. Using the normalization $\sum_{h=1}^H \underline{\lambda}_t^h = 1$, the dynamics for $\underline{\lambda}_{t+1}^h$ are then given by

$$\underline{\lambda}_{t+1}^h = \frac{\underline{\lambda}_t^h \underline{\Pi}_{t+1}^h}{\sum_{h=1}^H \underline{\lambda}_t^h \underline{\Pi}_{t+1}^h} \tag{52}$$

Hence, for Epstein-Zin preferences we obtain the following system for the first-order conditions (36)-(40):

The market clearing condition:

$$\sum_{h=1}^{H} s_t^h = 1. \tag{MC}$$

The optimality condition for the individual consumption decisions:

$$\underline{\lambda}_t^h(1-\delta^h)(s_t^h)^{\rho^h-1} = \underline{\lambda}_t^1(1-\delta^1)(s_t^1)^{\rho^1-1}, \quad h \in \mathbb{H}^- \tag{CD}$$

with $\sum_{h=1}^{H} \underline{\lambda}_t^h = 1$.

The value functions of the individual agents:

$$v_t^h = \left[(1-\delta^h)(s_t^h)^{\rho^h} + \delta^h R_t^h \left(v_{t+1}^h e^{\Delta c_{t+1}}\right)^{\rho^h}\right]^{\frac{1}{\rho^h}}, \qquad h \in \mathbb{H}. \tag{VF}$$

The equation for the dynamics of $\underline{\lambda}_t^h$:

$$\begin{aligned}
\underline{\lambda}_{t+1}^h &= \frac{\underline{\lambda}_t^h \underline{\Pi}_{t+1}^h}{\sum_{h=1}^{H} \underline{\lambda}_t^h \underline{\Pi}_{t+1}^h} \\
\underline{\Pi}_{t+1}^h &= \delta^h e^{\rho^h \Delta c_{t+1}} \frac{\mathrm{dP}_{t,t+1}^h}{\mathrm{dP}_{t,t+1}} \frac{\left(v_{t+1}^h e^{\Delta c_{t+1}}\right)^{\alpha^h-\rho^h}}{R_t^h \left(v_{t+1}^h e^{\Delta c_{t+1}}\right)^{\alpha^h-\rho^h}}, \qquad h \in \mathbb{H}^-.
\end{aligned} \tag{Dλ}$$

Note that equation (13) and hence the individual consumption decisions s_t^h only depend on time t information and there is no intertemporal dependence. This allows us to first solve for s_t^h given the current state of the economy and in a second step solve for the dynamics of the Negishi weights. Hence, we can separate solving the optimality conditions (11)-(16) into two steps in order to reduce the computational complexity. In the following section we describe this approach in detail.

Proof of Theorem 3. Let $f(s) = s^{\rho^j}$, $g(s) = s^{\rho^i}$, and $A = B = 1$, and apply lemma 1. □

A.3 The Case of Discounted Expected Utility

Proof of Theorem 4. Under the assumptions of the theorem, $\Pi_t^h = \delta$, and the ratio of λ's simplifies to

$$\frac{\lambda_{t+1}^h}{\lambda_{t+1}^1} = \frac{\lambda_t^h}{\lambda_t^1},$$

so they are not time-varying.

If aggregate consumption is Markov, then the first-order condition is purely a function of the aggregate state. □

Proof of Theorem 5. In logs,

$$\log \Pi_t^h = \log \delta^h + \log \frac{dQ_t^h}{dQ_t},$$

so

$$\frac{1}{T}\sum_{t=1}^{T} \log \Pi_t^h = \log \delta^h + \frac{1}{T}\sum_{t=1}^{T} \log \frac{dQ_t^h}{dQ_t} \to \log \delta^h - D(Q^h \| Q),$$

by the law of large numbers.

$$\log \lambda_T^i - \log \lambda_T^j = \left(\log \lambda_0^i - \log \lambda_0^j\right) + \sum_{t=1}^{T} \Pi_t^i - \sum_{t=1}^{T} \Pi_t^j,$$

so the result follows. □

For general dependent processes, survival results depend on a deceptively simple technical result.

Lemma 2. *Suppose that $E_t(\Pi_t^i/\Pi_t^j) \leq 1$. Then agent j survives.*

Proof. We have

$$E_t\left(\frac{\lambda_{t+1}^i}{\lambda_{t+1}^j}\right) = E_t\left(\frac{\Pi_{t+1}^i}{\Pi_{t+1}^j}\right)\left(\frac{\lambda_t^i}{\lambda_t^j}\right) \leq \frac{\lambda_t^i}{\lambda_t^j},$$

so λ_t^i/λ_t^j is a supermartingale. By the martingale convergence theorem, $\lim \lambda_t^i/\lambda_t^j$ exists as a random variable that is finite almost surely. Thus by the contrapositive of the second part of theorem 2, C_t^i does not go to zero. □

An identical result holds for $\underline{\Pi}_t^h$.

Proof of Theorem 6.

1. In this case, $\Pi^i/\Pi^j = \delta^i/\delta^j$, so $\lambda_t^i/\lambda_t^j \to \infty$. The result follows from theorem 2.
2. Suppose agent j has the correct belief. Then $\Pi^i/\Pi^j = dP_i/dP$. dP_i/dP is a martingale, so by theorem 2 the result follow.

□

Proof of Theorem 7.

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \log \underline{\Pi}_t^h = \log \delta^h + rho^h \mu - D(Q^h \| Q),$$

by the law of large numbers. The proof is otherwise identical to 5. □

Proof of Theorem 8. 1. The proof is identical to the proof of part 1 of theorem 6.

2. In this case $\Pi_t^i/\Pi_t^j = (C_t)^{\rho^i-\rho^j}$. If $\rho^i - \rho^j < 0$, then

$$E_t(C_{t+1}^{\rho^i-\rho^j}) \leq E_t(C_{t+1})^{\rho^i-\rho^j} \leq C_t^{\rho^i-\rho^j}.$$

   (The first inequality follows from Jensen's inequality, while the second follows from the fact that x^a is order reversing for $a < 0$.)

   Thus Π_t^i/Π_t^j is a supermartingale, and by Lemma 2 the result follows.

3. The proof is identical to the proof of part 2 of theorem 6.

□

A.4 Computational Procedure - A Two Step Approach

For the ease of notation the following procedures are described for $H = 2$ agents and a single state variable $y_t \in \mathbb{R}^1$. However, the approach can analogously be extended to the general case of $H > 2$ agents and multiple states. We solve the social planner's problem using a collocation projection. For this we perform the usual transformation from an equilibrium described by the infinite sequences (with a time index t) to the equilibrium being described by functions of some state variable(s) x on a state space X. We denote the current exogenous state of the economy by y and the subsequent state in the next period by y' with the state space $Y \in \mathbb{R}^1$. $\underline{\lambda}_2$ denotes the current endogenous state of the Negishi weight and $\underline{\lambda}_2'$ denotes the corresponding state in the subsequent period with $\underline{\Lambda}_2 \in (0, 1)$.

We approximate the value functions of the two agents $v^h(\underline{\lambda}_2, y), h = \{1, 2\}$ by two dimensional cubic splines and we denote the approximated value functions $\hat{v}^h(\underline{\lambda}_2, y)$. For the collocation projection we have to choose a set of collocation nodes $\{\underline{\lambda}_{2_k}\}_{k=0}^n$ and $\{y_l\}_{l=0}^m$ at which we evaluate $\hat{v}^h(\underline{\lambda}_2, y)$. The individual consumption shares only depend on the endogeneous state $\underline{\lambda}_{2_k}$. So in the following we show how to first solve for the individual consumption shares at the collocation nodes $s_k^h = s^h(\underline{\lambda}_{2_k})$ that are then used to solve for the value functions v^h and the dynamics of the endogenous state $\underline{\lambda}_2$.

Step 1: *Computing Optimal Consumption Allocations*

Equation (13) has to hold at each collocation node $\{\underline{\lambda}_{2_k}\}_{k=0}^n$:

$$\underline{\lambda}_{2_k}(1 - \delta^2)\left(s_k^2\right)^{\rho^2-1} = (1 - \underline{\lambda}_{2_k})(1 - \delta^1)\left(s_k^1\right)^{\rho^1-1}.$$

Together with the market clearing condition (11) we get

$$\underline{\lambda}_{2_k}(1-\delta^2)\left(s_k^2\right)^{\rho^2-1} = (1-\underline{\lambda}_{2_k})(1-\delta^1)\left(1-s_k^2\right)^{\rho^1-1}. \tag{53}$$

So for each node $\{\underline{\lambda}_{2_k}\}_{k=0}^n$ the optimal consumption choice s_k^2 can be computed by solving equation (53) and s_k^1 is obtained by the market clearing condition (11).[8] For the special case of $\rho^2 = \rho^1$ we can solve for s^2 as function of $\underline{\lambda}_2$ analytically and hence, we don't have to solve the system of equations for each node.

Step 2: *Solving for the Value Function and the Dynamics of the Negishi Weights*

Solving for the value function is not straight-forward as it depends on the dynamics of the endogenous state $\underline{\lambda}_2$ that are unknown and follow equation (16). We compute the expectation over the exogenous state by a Gauss-Quadrature with Q quadrature nodes. This implies that the values for y' at which we evaluate v^h are given by the quadrature rule. We denote the corresponding quadrature nodes by $\{y_{l,g}^{'}\}_{l=0,g=1}^{m,Q}$ and the weights by $\{\omega_g\}_{g=1}^{Q}$.[9] We can then solve equation (16) for a given pair of collocation nodes $\{\underline{\lambda}_{2_k}, y_l\}_{k=0,l=0}^{n,m}$ and the corresponding quadrature nodes $\{y_{l,g}^{'}\}_{l=0,g=1}^{m,Q}$ to compute a vector $\vec{\underline{\lambda}}_2'$ of size $(n+1)\times(m+1)\times Q$ that consists of the corresponding values $\underline{\lambda}'_{2_{k,l,g}}$ for each node. For each $\underline{\lambda}'_{2_{k,l,g}}$ equation (16) then reads

$$\begin{aligned}
\underline{\lambda}'_{2_{k,l,g}} &= \frac{\underline{\lambda}_{2_k}\underline{\Pi}^2}{(1-\underline{\lambda}_{2_k})\underline{\Pi}^1 + \underline{\lambda}_{2_k}\underline{\Pi}^2} \\
\underline{\Pi}^h &= \delta^h e^{\rho^h \Delta c(y'_{l,g})}\left(\frac{v^h(\underline{\lambda}'_{2_{k,l,g}}, y'_{l,g})e^{\Delta c(y'_{l,g})}}{R^h\left[v^h(\underline{\lambda}'_2, y')e^{\Delta c(y')}|\underline{\lambda}_{2_k}, y_l\right]}\right)^{\alpha^h-\rho^h}\frac{\mathrm{dP}^h(y'_{l,g}|y_l)}{\mathrm{dP}(y'_{l,g}|y_l)}
\end{aligned} \tag{54}$$

where

$$R^h\left[v^h(\underline{\lambda}'_2, y')e^{\Delta c(y')}|\underline{\lambda}_{2_k}, y_l\right] = G_h^{-1}\left(E\left[G_h\left(v^h(\underline{\lambda}'_2, y')e^{\Delta c(y')}\right)\frac{\mathrm{dP}^h(y')}{\mathrm{dP}(y')}\middle|\underline{\lambda}_{2_k}, y_l\right]\right).$$

Note that $\underline{\lambda}'_{2_{k,l,g}}$ depends on the full distribution of $\underline{\lambda}'_2$ through the expectation operator. By applying the Gauss-Quadrature to compute the expectation we get

[8]Note that in the case of H agents we have to solve a system of $H-1$ equations that pin down the $H-1$ individual consumption choices $s^h \in \mathbb{H}^-$.

[9]Note that the quadrature nodes $\{\{y_{l,g}^{'}\}_{g=0}^{G}\}_{l=0}^{m}$ depend on the state today $\{y_l\}_{l=0}^{m}$.

$$E\left[G_h\left(v^h(\underline{\lambda}_2', y')e^{\Delta c(y')}\right)\frac{\mathrm{dP}^h(y')}{\mathrm{dP}(y')}\middle|\underline{\lambda}_{2_k}, y_l\right] \approx \sum_{g=1}^{Q} G_h\left(v^h(\underline{\lambda}_{2_{k,l,g}}', y_{l,g})e^{\Delta c(y_{l,g})}\right)\cdot \omega_g.$$

So by computing the expectation with the quadrature rule, we do not need the full distribution of $\underline{\lambda}_2'$ but only have to evaluate v^h at the values $\underline{\lambda}_{2_{k,l,g}}'$ that can be obtained by solving (54) for each pair of collocation nodes $\{\underline{\lambda}_{2_k}, y_l\}_{k=0,l=0}^{n,m}$ and the corresponding quadrature nodes $\{y_{l,g}'\}_{l=0,g=1}^{m,G}$. So at the end we have a square system of equations with $(n+1)\times(m+1)\times G$ unknowns $\underline{\lambda}_{2_{k,l,g}}'$ and as many equations (54) for each $\{k, l, g\}$.

The value function is in general not known so we have to compute it simultaneously when solving for $\underline{\lambda}_{2_{k,l,g}}'$. Plugging the approximation $\hat{v}^h(\underline{\lambda}_2, y)$ into the value function (12) yields

$$\hat{v}^h(\underline{\lambda}_{2_k}, y_l) = \left[(1-\delta^h)\left(s_k^h\right)^{\rho^h} + \delta^h R^h\left(\hat{v}^h(\underline{\lambda}_2', y')e^{\Delta c(y')}\middle|\underline{\lambda}_{2_k}, y_l\right)\right]^{\frac{1}{\rho^h}}. \tag{55}$$

The collocation projection conditions require that the equation has to hold at each collocation node $\{\underline{\lambda}_{2_k}, y_l\}_{k=0,l=0}^{n,m}$. So we obtain a square system of equations with $(n+1)\times(m+1)\times 2$ equations (55) and as many unknowns for the spline interpolation at each collocation node that we solve simultaneously with the system for $\underline{\lambda}_{2_{k,l,g}}'$ described above.

A.5 Properties of the Value Function

In the case of heterogeneous agents the approximation of the value function is a delicate computational task as an agent can die out over time. Marginal utility of the agent at this limiting case is infinity which makes it difficult to obtain accurate approximations for the value function close to the singularity. To obtain information about the properties of the singularity, we formally derive the limiting behavior of the value function for the special case of an economy with no uncertainty. We then include this information in the value function approximation for the stochastic economy. From equation (13) we know that

$$s^2(\underline{\lambda}_2) = \left(\frac{1-\delta^1}{1-\delta^2}\right)^{-\psi^2}(\underline{\lambda}_2)^{\psi^2}(1-\underline{\lambda}_2)^{-\psi^2}(s^1(\underline{\lambda}_2))^{\frac{\psi^2}{\psi^1}}. \tag{56}$$

We are interested in the properties of $s^2(\underline{\lambda}_2)$ for $\underline{\lambda}_2$ close to 0. For $\underline{\lambda}_2 \approx 0$ agent 1 obtains all consumption so $s^1(\underline{\lambda}_2) \approx 1$ and the Negishi weight of the first agent becomes 1. Therefore we obtain

$$s^2(\underline{\lambda}_2) \approx \left(\frac{1-\delta^1}{1-\delta^2}\right)^{\psi^2}(\underline{\lambda}_2)^{\psi^2} \tag{57}$$

for $\underline{\lambda}_2$ close to 0. The value function (41) for the deterministic economy at the steady state $y = y', \underline{\lambda}_2 = \underline{\lambda}'_2$ is given by

$$v^2(\underline{\lambda}_2, y) = s^2(\underline{\lambda}_2). \tag{58}$$

Inserting the behavior of $s^2(\underline{\lambda}_2)$ for $\underline{\lambda}_2$ close to 0, we obtain[10]

$$v^2(\underline{\lambda}_2) \approx \left(\frac{1-\delta^1}{1-\delta^2}\right)^{\psi^2} (\underline{\lambda}_2)^{\psi^2} \equiv \Upsilon^0(\underline{\lambda}_2). \tag{59}$$

We denote by $\Upsilon^0(\underline{\lambda}_2)$ the zero basis functions which we add to the cubic spline value function approximation to obtain accurate approximations close to the singularity. We find that for all solutions reported in this paper, including the zero basis functions improves the accuracy of the solution. This concludes the description of the methodology for solving the heterogeneous agent model with recursive preferences. In the following section we apply the approach to solve the long-run risk model of Bansal and Yaron (2004) with heterogeneous agents.

A.6 Computational Details

For the projection method outlined above, we need to choose certain collocation nodes. In this paper we use 17 uniform nodes for the $\underline{\lambda}^2$ dimension and 13 uniform nodes for the x_t dimension for the results with $\rho^2 = 0.975$ and $\rho^1 = 0.985$. For the results with $\rho^2 = 0.95$ and $\rho^1 = 0.985$, we use 51 uniform nodes for the $\underline{\lambda}^2$ dimension and 23 uniform nodes for the x_t dimension. For $\underline{\lambda}^2$ the minimum and maximum values are given by 0 and 1. For x_t we choose the approximation interval to cover ± 4 standard deviations around the unconditional mean of the process. We approximate the value functions using two-dimensional cubic splines with not-a-knot end conditions. We provide the solver with additional information that we can formally derive for the limiting cases. For example we know that for $\underline{\lambda}_t^2 = 1$ ($\underline{\lambda}_t^2 = 0$) agent 2 (1) consumes everything, so it corresponds to the representative agent economy populated only by agent 2 (1). Hence, we require that the value function for these cases equals the value function for the corresponding representative agent economy. We also know that for $\underline{\lambda}_t^2 = 0$ ($\underline{\lambda}_t^2 = 1$) consumption of agent 2 (1) is 0 and hence the value function is also zero. As the shocks in the model are normally distributed, we compute the expectations over the exogenous states by Gauss-Hermite quadrature using 5 nodes for the shock in x_{t+1} and 3 nodes for the shock in Δc_{t+1}. Euler errors for the value function approximations evaluated on a 200×200 uniform grid for both states are less than 1×10^{-6} suggesting a high accuracy of

[10] For the first agent we obtain a similar expression for $\underline{\lambda}_2$ close to 1 given by $v^1(\underline{\lambda}_2) \approx \left(\frac{1-\delta^2}{1-\delta^1}\right)^{\psi^1} (1-\underline{\lambda}_2)^{\psi^1}$.

our results. We double checked the accuracy by increasing the approximation interval as well as the number of collocation nodes with no significant change is the results.

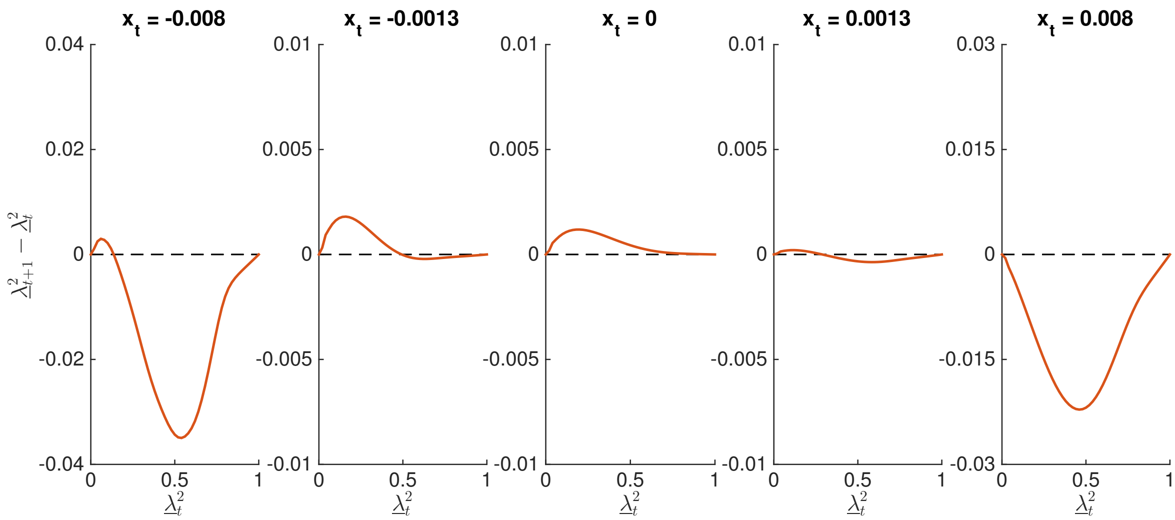
B Additional Figures and Tables

Table 4: Equity Premium for Different Consumption Shares ($\rho_x = \rho_x^2$)

	$\boldsymbol{\rho_x^2 = 0.975}$		$\boldsymbol{\rho_x^2 = 0.95}$	
	s_t^2	Equity Premium	s_t^2	Equity Premium
Rep. Agent 1	0	6.49	0	6.50
0 Years	0.01	6.38	0.01	5.36
100 Years	0.3404	4.39	0.8147	0.49
200 Years	0.7249	3.33	0.8810	0.41
500 Years	0.9732	2.83	0.9388	0.34
Rep. Agent 2	1	2.75	1	0.25

The table shows the annualized equity premium for a specific consumption share $s_t^2 = \bar{s}$. The premium is reported for the equilibrium allocations after 0, 100, 200 and 500 years of simulated data assuming an initial share of $s_0^2 = 0.01$ (see Table 1). Agent 1 beliefs that $\rho_x^1 = 0.985$ and agent 2 has the correct beliefs ($\rho_x = \rho_x^2$). The left panel depicts the case for $\rho_x^2 = 0.975$ and the right panel for $\rho_x^2 = 0.95$.

Figure 10: Changes in the Wealth-Distribution—The Epstein-Zin Case 2 ($\rho_x^2 = 0.95$)



The figure shows the change in the optimal weights $\underline{\lambda}_{t+1}^2 - \underline{\lambda}_t^2$ as a function of $\underline{\lambda}_t^2$. From left to right, the change is shown for $x_t = \{-0.008, -0.0013, 0, 0.0013, 0.008\}$ ($\pm$ 4 standard deviations). The red line shows the average over all shocks in x_{t+1}. Calibration with $\rho_x = \rho_x^1 = 0.985$ and $\rho_x^2 = 0.95$.

References

Alchian, A. A. (1950): "Uncertainty, Evolution, and Economic Theory," *Journal of Political Economy*, 58, 211–221.

Andrei, D., B. Carlin, and M. Hasler (2016): "Model Calibration And Uncertainty," Working Paper.

Bansal, R., D. Kiku, and A. Yaron (2012): "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 1, 183–221.

——— (2016): "Risks For the Long Run: Estimation with Time Aggregation," *Journal of Monetary Economics*, 82, 52–69.

Bansal, R. and A. Yaron (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *The Journal of Finance*, 59, 1481–1509.

Beeler, J. and J. Y. Campbell (2012): "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment," *Critical Finance Review*, 1, 141–182.

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

Blume, L. and D. Easley (2006): "If You're so Smart, why Aren't You Rich? Belief Selection in Complete and Incomplete Markets," *Econometrica*, 74, 929–966.

Borovička, J. (2015): “Survival and Long-Run Dynamics with Heterogeneous Beliefs Under Recursive Preferences,” Available at SSRN: `http://ssrn.com/abstract=2023501`.

Branger, N., I. Dumitrescu, V. Ivanova, and C. Schlag (2011): “Preference Heterogeneity and Survival in Long-Run Risk Models,” Working Paper, Goethe University.

Branger, N., P. Konermann, and C. Schlag (2015): “Optimists, Pessimists, and the Stock Market: The Role of Preferences and Market (In)Completeness,” Working Paper, Available at SSRN: `http://ssrn.com/abstract=2356502`.

Carlin, B. I., F. A. Longstaff, and K. Matoba (2014): “Disagreement and asset prices,” *Journal of Financial Economics*, 114, 226–238.

Chen, H., S. Joslin, and N.-K. Tran (2012): “Rare Disasters and Risk Sharing with Heterogeneous Beliefs,” *Review of Financial Studies*, 25, 2189–2224.

Chien, Y., H. Cole, and H. Lustig (2011): “A Multiplier Approach to Understanding the Macro Implications of Household Finance,” *The Review of Economic Studies*, 78, 199–234.

——— (2012): “Is the Volatility of the Market Price of Risk Due to Intermittent Portfolio Rebalancing?” *American Economic Review*, 102, 2859–96.

——— (2016): “Implications of heterogeneity in preferences, beliefs and asset trading technologies in an endowment economy,” *Review of Economic Dynamics*, 20, 215–239.

Collin-Dufresne, P., M. Johannes, and L. Lochstoer (2015): “A Robust Numerical Method for Solving Risk-Sharing Problems with Recursive Preferences,” Working Paper.

——— (2016): “Asset Pricing when ”This Time is Different”,” *The Review of Financial Studies*, Forthcoming.

Cvitanić, J., E. Jouini, S. Malamud, and C. Napp (2012): “Financial Markets Equilibrium with Heterogeneous Agents,” *Review of Finance*, 16, 285–321.

David, A. (2008): “Heterogeneous Beliefs, Speculation, and the Equity Premium,” *The Journal of Finance*, 63, 41–83.

Drechsler, I. and A. Yaron (2011): “What’s Vol Got to Do with It,” *Review of Financial Studies*, 24, 1–45.

DUMAS, B., R. UPPAL, AND T. WANG (2000): "Efficient Intertemporal Allocations with Recursive Utility," *Journal of Economic Theory*, 93, 240–259.

EPSTEIN, L. AND S. ZIN (1989): "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

FRIEDMAN, M. (1953): *Essays in Positive Economics*, Chicago: University of Chicago Press.

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.

GRAMMIG, J. AND E. M. SCHAUB (2014): "Give Me Strong Moments and Time: Combining GMM and SMM to Estimate Long-Run Risk Asset Pricing Models," Available at SSRN: `http://ssrn.com/abstract=2386094`.

JUDD, K., F. KUBLER, AND K. SCHMEDDERS (2003): "Asset Trading Volume with Dynamically Complete Markets and Heterogeneous Agents," *The Journal of Finance*, 58, 2203–2217.

KROENECKE, T. A. (2016): "Asset Pricing without Garbage," *The Journal of Finance*, Forthcoming.

ROCHE, H. (2011): "Asset prices in an exchange economy when agents have heterogeneous homothetic recursive preferences and no risk free bond is available," *Journal of Economic Dynamics and Control*, 35, 80–96.

SANDRONI, A. (2000): "Do Markets Favor Agents Able to Make Accurate Predictions?" *Econometrica*, 68, 1303–1341.

SCHORFHEIDE, F., D. SONG, AND A. YARON (2016): "Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach," Available at: `http://www.economics.illinois.edu/seminars/documents/Frank.pdf`.

WEIL, P. (1989): "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 24, 401–421.

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