

Asset Pricing when 'This Time is Different'

Pierre Collin-Dufresne, Michael Johannes and Lars A. Lochstoer*

University of Lausanne and Columbia Business School

February 2014

Abstract

Recent empirical evidence suggest that the young update beliefs about macro outcomes more in response to aggregate shocks than the old. We embed this form of 'this time is different'-bias in an overlapping generations general equilibrium macro-finance model where agents have recursive preferences and are unsure about the specification of the exogenous aggregate stochastic process. In this economy, parameter/model learning persists indefinitely. Differences in beliefs and learning about the model specification induces additional, quantitatively significant priced risks in the economy, as well as significant time-variation in these risks, depending both on agents' the relative wealth and beliefs. Time-varying bias in agents' beliefs leads to over- and under-valuation, accompanied by over- and under-investment, that tends to be exacerbated in equilibrium as outcomes of the optimal risk-sharing in the economy. This feature also induces fat tails in market returns as in the data.

*Pierre Collin-Dufresne is at the University of Lausanne. Michael Johannes and Lars Lochstoer are at Columbia Business School. This is work in progress – comments are welcome! Contact info: Lars A. Lochstoer, 405B Uris Hall, Columbia Business School, Columbia University, 3022 Broadway, New York, NY 10027. E-mail: LL2609@columbia.edu. First draft: November 2013.

1 Introduction

Investors' beliefs regarding future macro economic outcomes are important determinants of aggregate asset prices. Recent empirical evidence suggest that agents update their beliefs about macro economic outcomes differently based on their own personal experience. For instance, Malmendier and Nagel (2011) argue that investors who lived through the Great Depression display more pessimistic beliefs due to their direct, personal exposure to this event than younger agents, who were born later and therefore did not personally experience such a severe downturn. Similarly, Nagel and Malmendier (2013) present evidence from survey data that agents' sensitivity of beliefs about inflation dynamics to a shock to inflation is decreasing with the age of the agent. Thus, when learning about the economic environment, the young update more in response to shocks than the old. Taken together, these findings imply that agents place more importance on personal experience relative to what a fully Bayesian agent would do when forming beliefs.

While it is clear that this departure from Rational Expectations implies that any one agent's beliefs about aggregate dynamics at times will be too optimistic or too pessimistic relative to the beliefs of a fully rational Bayesian agent, the general equilibrium asset pricing implications of such a bias are not clear. For instance, since aggregate asset prices are a function of the (weighted) average belief, differences in beliefs between young and old may approximately 'wash out' and therefore have only minor effects on prices, especially over the long run. In fact, Ang, Bekaert, and Wei (2007) show, using the same survey data on inflation expectations, that the median (or mean) inflation forecast outperforms pretty much any other forecast they construct from available macro and asset price data. Thus, the median (or mean) belief appears to be quite 'rational,' which restricts the amount of bias in beliefs agents can exhibit.¹

This paper proposes a general equilibrium, macro-finance asset pricing model that incorporates this behavioral bias in way consistent with the above mentioned evidence on macro expectations. In this model, mean beliefs about aggregate macro outcomes such as GDP growth over the next year are very close to being unbiased in a 'Rational Expectations' sense, but at the same time a 'this time is different'-bias leads to substantial extra risks in the economy, helping the model account for high Sharpe ratios

¹Given the nature of the survey data, the forecasting horizon is up to 12 months. Note that the evidence in Ang, Bekaert, and Wei (2007) also holds for the Michigan survey, where respondents are consumers, not professionals.

and risk premiums as observed in the data along with low investor risk aversion. The model also features extended periods of over- and under-valuation, accompanied by over- and under-investment, leading to return predictability patterns as those found in the data.

In the model, agents are uncertain about the specification of the exogenous aggregate stochastic process and update beliefs as new data arrives in a Bayesian manner. In particular, there are two generations alive at each point in time, young and old. Each generation lives for 40 years, so there is a 20 year overlap between generations. When born, agents inherit the mean beliefs about the model specification from their parent generation (who die and are the previously Old), but with a prior variance of beliefs that is higher than the posterior variance of their parent generation's beliefs. The latter is the source of the 'this time is different'-bias. We consider the particular case where agents are unsure about the mean growth rate of the economy. A fully rational, Bayesian agent would eventually learn the true growth rate, but due to the 'this time is different' OLG feature of the model parameter learning persists indefinitely in this economy.

We assume agents have Epstein-Zin preferences with a preference for early resolution of uncertainty, following Bansal and Yaron (2004). With such preferences, updates in beliefs about the model specification are priced and can substantially amplify the effect of macro shocks on marginal utility. In particular, when a Bayesian agent updates her beliefs about a fixed quantity (the mean growth rate), this update is permanent and therefore affects the subjective consumption distribution for all future dates. Thus, even a small update in beliefs can have a large impact on the continuation utility. Effectively, parameter/model learning generates subjective long-run consumption risks (see Collin-Dufresne, Johannes, and Lochstoer (2013a)).

The fact that the young and old 'agree to disagree' means the more optimistic agent will hold a relatively high portfolio weight in assets that decrease in value in bad times, thus making it harder to match the standard asset pricing moments. On the other hand, the 'this time is different'-bias leads to perpetual model uncertainty, which itself is a significant source of risk given the chosen investor preferences. In fact, in a benchmark model with i.i.d. consumption growth and learning about the mean growth rate in the economy, the 'this time is different' Epstein-Zin model can account for the standard asset pricing moments, as well as time-varying Sharpe ratios and risk premiums due to periods of optimistic and pessimistic beliefs. In contrast, power utility

preferences with the same level of risk aversion leads to an equity Sharpe ratio that is on average close to zero and otherwise overall worse asset pricing performance when 'this time is different'-learning is introduced.

When considering a standard representative firm production economy, consumption and capital accumulation are endogenous and the agents learn about the mean growth rate of technology growth. Again Sharpe ratios are high and strongly time-varying as in the exchange economy case. Further, we show that the 'this time is different'-bias in this economy increases the volatility of equity returns by an order of magnitude due to the fluctuations in aggregate beliefs. Similar to the analysis in Hirshleifer and Yu (2013), who consider a single agent production economy with extrapolative beliefs about technology growth, agents have a strong desire to increase (decrease) investment when beliefs are optimistic (pessimistic). Therefore, the model allows for higher capital adjustment costs while still matching aggregate investment dynamics, which helps with the otherwise much too low volatility of the equity claim in these models (see Jermann (1998), and Kaltenbrunner and Lochstoer (2010)). In the model, as in the data, aggregate investment negatively forecasts future excess equity returns.

In the production economy output is endogenous and a function of capital accumulation as well as the technology level. If investors currently are, say, optimistic, meaning that their mean belief about the mean technology growth rate is higher than truth, they will tend to invest more aggressively, leading to high future output growth. Thus, since agents make investment decisions based on their beliefs, high (low) expected output growth under the agents' beliefs is over the relatively near future associated with on average high (low) output growth also under the true probability measure. Of this reason, the mean beliefs of agents in this economy is a good forecaster of future output growth over the horizons considered in Ang, Bekaert, and Wei (2007). We show that, in fact, this forecast is as good as or better than the forecasts one would obtain from running, say, an AR(4) on output growth. We document that this is also true in the data using actual survey forecasts for real output growth (combining forecasts of inflation and nominal output).²

The time-varying conditional Sharpe ratio of the equity claim the model generates is due to time-varying volatility of the stochastic discount factor under the true probabil-

²The studies mentioned initially analyze inflation dynamics, but the standard production economy model we analyze is concerned with real quantities, which is why we consider real output growth, for which there are also survey forecasts available.

ity measure. Thus, the learning mechanism with a 'this time is different'-bias leads to what an econometrician would identify as time-varying and high prices of macro risks, even though agents have the same preferences with relatively low levels of risk aversion, utility is isoelastic and shocks are homoskedastic. This is notable as the model is also consistent with the average belief about output growth over the next year being quite accurate, as in the data. However, if one could measure *long-run* beliefs about real output growth, both in surveys and accurately as an econometrician, this is where one would see that agents get it wrong. Unfortunately, survey data on longer-run beliefs (5 and 10 year outlooks) are so far quite limited.

In terms of the dynamics that arise from having heterogeneous agents, the optimal risk-sharing in the Epstein-Zin model tends to exacerbate the impact of biased beliefs on asset prices and investment as the more optimistic (pessimistic) agent holds more (less) stock. A positive (negative) shock is therefore amplified in terms of the wealth-weighted average belief in this model where, unlike the power utility case, prices are pro-cyclical. This endogenous amplification of shocks is much stronger when agents have Epstein-Zin preferences as there is a larger difference in the impact of model risk on utility across the generations when agents are very averse to model uncertainty (when they have a preference for early resolution of uncertainty). This means the average difference in portfolio holdings across generations is also large. This is opposed to the case of power utility where model uncertainty in general has much less impact on utility. This amplification feature leads to fat tails in market returns, similar to those found in the data, even though the exogenous shocks are Normal.

There are four state-variables in the exchange economy model and five state variables in the production economy model, which also has capital as a state variable. Solving the endogenous risk sharing problem is non-trivial when agents have Epstein-Zin preferences. We solve the model using a new robust numerical solution methodology developed by Collin-Dufresne, Johannes, and Lochstoer (2013b) for solving risk-sharing problems in complete markets when agents have recursive preferences. This numerical method does not rely on approximations to the actual economic problem (e.g., it does not rely on an expansion around a non-stochastic steady-state) and therefore provides an arbitrarily accurate solution (depending of course on the chosen coarseness of grids, quadratures, etc.).³

³Accurate solutions do require efficient coding in a fast programming language, such as C++ or Fortran, and extensive use of the multiprocessing capability of high-performance desktops. Such

Related literature. There is a large literature on the effects of differences in beliefs on asset prices. Harrison and Kreps (1978) and Scheinkman and Xiong (2003) show how over-valuation can arise when agents have differences in beliefs and there are short sale constraints. Dumas, Kurshev, and Uppal (2009) consider a general equilibrium, complete markets model where two agents with identical power utility preferences disagree about the dynamics of the aggregate endowment. Baker, Hollifield, and Osambela (2014) consider a general equilibrium production economy where two agents have heterogeneous beliefs about the mean productivity growth rate, where the agents agree-to-disagree and do not update their beliefs. Agents have power utility preferences, and the authors show that speculation leads to a counter-cyclical risk premium and that the investment and stock return volatility dynamics are counter-cyclical when agents have high elasticity of intertemporal substitution. These papers are close to ours along many dimensions, with the most important exceptions being that our agents have Epstein-Zin preferences and the overlapping generations feature of our model, which determines the learning dynamics. A number of the properties of equilibrium are qualitatively similar. We therefore focus our analysis on the particular implications of agents with recursive preferences, as compared to the standard time separable CRRA preferences.

In contemporaneous work, Ehling, Graniero, and Heyerdahl-Larsen (2013) consider a similar learning bias in an OLG endowment economy framework, but with log utility preferences. These authors present empirical evidence that the expectations of future stock returns are more highly correlated with recent past returns for the young than the old, consistent with the overall evidence given by Malmendier and Nagel (2011, 2013). In other recent work, Choi and Mertens (2013) solve a model with two sets of infinitely-lived agents with Epstein-Zin preferences, portfolio constraints, where one set of agents has extrapolative beliefs, in an incomplete markets setting. These authors estimate the size of the belief bias by backing it out from standard asset price moments, whereas we calibrate the bias to available micro estimates as given by Malmendier and Nagel (2011, 2013), as well as solve an OLG model. Further, our numerical model solution method is exact, whereas they rely on higher-order expansions around a steady-state solution. Both these authors and Dumas, Kurshev, and Uppal (2009) have only one set of agents with biased beliefs, while the generational 'this time is different'-bias leads to multiple agents with biased beliefs. Thus, the nature of the OLG problem we solve

technology is, however, easily available.

has more state variables as we need to keep track of the individual beliefs of multiple generations (the young and the old in our case). In terms of heterogeneous agent models with Epstein-Zin preferences, Garleanu and Panageas (2012) solve an OLG model with Epstein-Zin agents with different preference parameters, while Borovicka (2012) shows long-run wealth dynamics in a two-agent general equilibrium setting where agents have Epstein-Zin preferences and differences in beliefs. Finally, Marcet and Sargent (1989), Sargent (1999), Orphanides and Williams (2005a), and Milani (2007) are prominent examples of the effects of perpetual, non-Bayesian learning in macro economics.

2 The Model

Solving a model with parameter learning where agents have differences in beliefs quickly becomes infeasible as the state space becomes prohibitively large. Of that reason, we will focus on parsimonious learning problems in an as simple as possible, but still quantitatively interesting, overlapping generations framework. In this Section, we consider an exchange economy setting so the beliefs about aggregate consumption (which drives the asset pricing implications) are explicit, while we in Section 4 consider the implications of the 'this time is different'-bias in a production economy where consumption, output and investment are distinct endogenous quantities.

We assume there are two sets of agents alive at any point in time, young and old. A generation last for T periods, and each agent lives for $2T$ periods. Thus, there is no uncertainty about life expectancy. All young agents currently alive were born at the same time, all old agents currently alive were born at the same time, and all agents born at the same time have the same prior beliefs. These assumptions imply that (a) there are no hedging demands related to uncertain life span, and (b) that there effectively are only two agents in the economy. The latter is important in order to keep the number of state variables manageable. The former is a necessary assumption for the latter to be true given our learning problem, as will be clear below.

When an 'old' generation dies, the previously young generation becomes the new old generation and a new young generation is born. The old leave their assets for their offspring, the new young.⁴ We model the bequest motive as that in a Dynasty model.

⁴The labels 'old' and 'young' in this model refer to the two generations currently alive. A new generation could be born, say, every 20 years, which implies the investors in this economy live for 40 years. When the old 'die' they give life to new 'young,' and so 'death' may be better thought of as

I.e., the parents care as much about their offspring as themselves (with the caveat that there is time-discounting in the utility function).

2.1 Belief formation and the 'This Time is Different'-bias

The agents in the economy do not know the true model specification for the dynamics of aggregate consumption. They are Bayesian, however, and learn about these dynamics by updating their beliefs from observed consumption growth realizations during their lifetime. We assume that the young inherit the beliefs from their parent generation in the following manner. The young inherit the mean beliefs about each parameter or model probability from their parents (the dying old). However, the young are assumed to be incapable of learning all of what their parents know, or simply do not trust their parents fully, and therefore they are, when born, endowed with more dispersed initial beliefs than their parents had. This is the source of the 'this time is different'-bias.

A concrete example is useful here. Assume log aggregate consumption growth is given by:

$$\Delta c_{t+1} = \mu + \sigma \varepsilon_{t+1}, \quad (1)$$

where σ is known and $\varepsilon_{t+1} \stackrel{i.i.d.}{\sim} N(0, 1)$, but where μ is unknown. Assume further that the time t beliefs of agent i are given by $\mu \sim N(m_{i,t}, A_{i,t}\sigma^2)$, where, from Bayes rule:

$$m_{i,t+1} = m_{i,t} + \frac{A_{i,t}}{1 + A_{i,t}} (\Delta c_{t+1} - m_{i,t}) \quad (2)$$

and

$$A_{i,t+1}^{-1} = A_{i,t}^{-1} + 1. \quad (3)$$

First, note that if agent i were to live forever, the variance of her subjective beliefs, $A_{i,t}\sigma^2$, would go to zero, per Equation (3). Assume the posterior beliefs of the old at death are $m_{o,t}$ and $A_{o,t}$. The new young will then be born and consume at time $t + 1$ with prior beliefs $m_{y,t} = m_{o,t}$ and $A_{y,t} = kA_{o,t}$, where $k \geq 1$. The parameter k determines to what extent the agents in the economy suffer from a 'this time is different'-bias. If k is set such that the prior variance of each young generation is

retirement age and the new 'young' as post-graduates, perhaps in their mid twenties. In other words, the model is stylized in order to in a transparent manner capture a 'this time is different'-bias related to personal experience in a quantitatively interesting setting. It is not designed to explain all aspects of observed life-cycle patterns in endowments or consumption-saving decisions.

constant (or increasing), parameter learning persists indefinitely. We assume that the young and old generations that live concurrently do not update from each others beliefs—i.e., they 'agree to disagree.'

In a fully rational setting each agent would learn from the full history of consumption realizations in a Bayesian manner. As should be clear from the preceding discussion, the updating scheme with the 'this time is different'-bias implies that the Young update 'too much' from personal experience relative to the Rational Expectations benchmark case.

2.2 The utility function with bequest motive

We assume agents have recursive preferences a la Epstein and Zin (1989). In particular, the value function $V_{i,t}$ of agent i alive at time t who will 'die' at time $\tau > t + 1$ is:

$$V_{i,t}^\rho = (1 - \beta) C_{i,t}^\rho + \beta E_t^i [V_{i,t+1}^\alpha]^{\rho/\alpha}. \quad (4)$$

Here, $\rho = 1 - 1/\psi$ where ψ is the elasticity of intertemporal substitution (EIS) and $\alpha = 1 - \gamma$, where γ is the 'risk aversion'-parameter. As shown in Collin-Dufresne, Johannes, and Lochstoer (2013a), a preference for early resolution of uncertainty, which the Epstein-Zin preference allow for, greatly magnifies the impact of model learning on equilibrium asset prices. The agent's last date consuming is τ . At $\tau + 1$ the agent's offspring j comes to life and starts consuming at time $\tau + 1$. The offspring will have different beliefs about the aggregate endowment, per the above discussion. We consider a bequest function of the form:

$$B_i(W_{i',\tau+1}) = \phi_{i'}(X_{\tau+1}) W_{i',\tau+1}, \quad (5)$$

where X_t is a vector of state variables while W_t is the agent's wealth at time t . State variables will include all agent's beliefs as well as a measure of the time each class of agent has been alive (or equivalently, how long until end of life). The index i' denotes agent i 's offspring.

With this bequest function, we have that:

$$V_{i,\tau}^\rho = (1 - \beta) C_{i,\tau}^\rho + \beta E_\tau^i [\phi_{i'}(X_{\tau+1})^\alpha W_{i',\tau+1}^\alpha]^{\rho/\alpha}. \quad (6)$$

Substituting in the usual budget constraint, we have:

$$V_{i,\tau}^\rho = (1 - \beta) C_{i,\tau}^\rho + \beta (W_{i,\tau} - C_{i,\tau})^\rho E_\tau^i [\phi_{i'} (X_{\tau+1})^\alpha R_{w_{i,\tau+1}}^\alpha]^\rho. \quad (7)$$

The first order condition over consumption implies that:

$$\begin{aligned} \rho (1 - \beta) C_{i,\tau}^{\rho-1} &= \rho \beta (W_{i,\tau} - C_{i,\tau})^{\rho-1} \mu_{i,\tau}^\rho \Leftrightarrow \\ \mu_{i,\tau} &= \left(\frac{1 - \beta}{\beta} \right)^{1/\rho} \left(\frac{W_{i,\tau}}{C_{i,\tau}} - 1 \right)^{(1-\rho)/\rho}, \end{aligned} \quad (8)$$

where the certainty equivalent is denoted $\mu_{i,\tau} = E_\tau^i [\phi_{i'} (X_{\tau+1})^\alpha R_{w_{i,\tau+1}}^\alpha]^{1/\alpha}$. Inserting this back into the value function, we have that:

$$\frac{V_{i,\tau}}{W_{i,\tau}} = (1 - \beta)^{1/\rho} \left(\frac{W_{i,\tau}}{C_{i,\tau}} \right)^{1/\rho-1}. \quad (9)$$

Since preferences are homothetic, the W/C ratio is a function of the state variables X_t . Let:

$$\phi_i (X_t) = (1 - \beta)^{1/\rho} \left(\frac{W_{i,\tau}}{C_{i,\tau}} \right)^{1/\rho-1}. \quad (10)$$

Then, $V_{i,t} = \phi_i (X_t) W_{i,t}$ for each t during the life of agent i . Since i was a general agent, it follows that $B_i (W_{i',\tau+1}) = V_{i',\tau+1}$. In this sense, the bequest function is of a Dynasty form, where the agent cares 'as much' about their offspring as themselves. Note, however, that the expectation of the offspring's indirect utility is taken using the parent generation's beliefs.

This specification of the bequest motive has the following nice implication: when there is no model/parameter uncertainty, the model reduces to a representative agent model where the representative agent is an infinitely-lived Epstein-Zin agent with the same preference parameters as those assumed above (β, γ, ψ). This agent, together with the maintained assumption of i.i.d. consumption growth, implies that the risk premium and price of risk are constant. In particular, the Sharpe ratio and risk premium of the consumption claim in this benchmark economy are approximately $\gamma\sigma$ and $\gamma\sigma^2$, respectively. Thus, the benchmark economy without learning and the associated 'this time is different'-bias fails along all the usual dimensions (equity premium, excess volatility, high Sharpe ratio, predictability of excess returns).

2.3 The consumption sharing rule

We assume markets are complete, so each agent's intertemporal marginal rates of substitution are equal for each state ω . The two (sets of) agents alive at each time t are the young and the old. It is useful, in light of results to come, to index the two agents in the economy as belonging to Dynasty A or Dynasty B, where a Dynasty consists of a lineage of parent-child relations. That is, when an 'old' agent leaves wealth to a new 'young' agent, who in the process inherits the mean beliefs of the 'old,' the two are related in a Dynasty. Thus, the current living representative agent from each Dynasty is either 'young' or 'old' at each point in time. Further, if agent A is 'old,' agent B is 'young,' and vice versa.

Given Equations (4), (5), and (10), and the assumption of complete markets, we have that the two representative agents' ratios of marginal utilities are equalized in each state (i.e., the stochastic discount factor is unique and both agents' IMRS price assets given the respective agent's subjective beliefs):

$$\begin{aligned} \pi^A(\omega_{t+1}|X_t) \left(\frac{c_{A,t+1}}{c_{A,t}} \right)^{\rho-1} \left(\frac{v_{A,t+1}}{\mu_{A,t}(v_{A,t+1}C_{t+1}/C_t)} \right)^{\alpha-\rho} &= \dots \\ \pi^B(\omega_{t+1}|X_t) \left(\frac{1-c_{A,t+1}}{1-c_{A,t}} \right)^{\rho-1} \left(\frac{v_{B,t+1}}{\mu_{B,t}(v_{B,t+1}C_{t+1}/C_t)} \right)^{\alpha-\rho} &. \end{aligned} \quad (11)$$

Here X_t holds the total set of state variables in the economy, including the sufficient statistics for each agent's beliefs, $c_{i,t} \equiv C_{i,t}/C_t$, $v_{i,t} \equiv V_{i,t}/C_t$, and where we have imposed the goods market clearing condition $c_{A,t} + c_{B,t} = 1 \iff C_{A,t} + C_{B,t} = C_t$ for all t . Note that the beliefs within each Dynasty are given by the agent currently alive at time t (see Equation (6)). Thus, if in, say, Dynasty A there is a change from 'old' to a new 'young,' the beliefs relevant for the IMRS between the two periods are that of the 'old.' In the following period, however, the beliefs will be those of the new young, who have higher prior dispersion when $k > 1$.

With recursive preferences the value functions appear in the intertemporal marginal rates of substitution. Thus, unlike in the special case of power utility, Equation (11) does not provide us with an analytical solution for the consumption sharing rule. This complicates significantly the model solution. We solve the model using a backwards recursion algorithm that solves numerically for the consumption sharing rule starting at a distant terminal date for the economy, \tilde{T} . The solution corresponds to the in-

finite horizon economy when the transversality condition is satisfied and \tilde{T} is chosen sufficiently far into the future (e.g., 500+ years). The state variables in this model are $m_{A,t}$, $m_{B,t}$, $\omega_{A,t}$ (the relative wealth of agent A), and t . Time t is a sufficient statistic for $A_t^{A,B}$ as the variance dynamics are deterministic. Details regarding the numerical solution technique are given in Collin-Dufresne, Johannes, and Lochstoer (2013b). We note that for general ρ and α , the prior distribution for the mean growth rate μ must be bounded in order to have existence of equilibrium.

2.4 Model Discussion

With Epstein-Zin preferences and a preference for early resolution of uncertainty ($\gamma > 1/\psi$), the agents are averse to long-run risks (see Bansal and Yaron (2004)). Learning about the mean growth rate induces subjective long-run consumption risks as both the conditional mean and volatility of consumption growth are slowly time-varying for the agent that is learning about the mean consumption growth rate (see Equations (2) and (3)). Collin-Dufresne, Johannes, and Lochstoer (2013a) show in the case of a representative agent that parameter and model learning can be a tremendous amplifier of macro shocks in terms of their impact on marginal utility with such preferences.

The same amplification mechanism is at work in the model at hand, but, importantly, there is (a) risk-sharing across generations with respect to the model uncertainty and (b) learning persists indefinitely if k is set appropriately. The former arises as agents that differ in their assessment of probabilities of future states will trade with each other to decrease perceived risk in the economy. That is, bad states that are perceived as less likely from the perspective of agent A relative to that of agent B will be states for which agent B will buy insurance from agent A and vice versa, thus making each agent better off given their beliefs. These effects serve to decrease the risk premium and undo some of the asset pricing effects of the long-run risks that arise endogenously from parameter learning. However, the 'this time is different'-bias works in the opposite direction in that the magnitude of belief updates from parameter learning remains high indefinitely.

2.5 Model Calibration

Both sets of agents have the same preference parameters, which we set to the same values as in Bansal and Yaron (2004), though in a quarterly calibration: $\gamma = 10$,

$\psi = 1.5$ and $\beta = 0.994$.⁵ We will refer to this as the 'EZ' calibration and assume, as previously stated, that aggregate consumption growth is truly i.i.d., with mean and volatility set to the same variables as in Bansal and Yaron (2004), in order to facilitate easy comparison with this well-known model. To this end, calibrating the model at the quarterly frequency, we set $\mu = 0.0045$ and $\sigma = 0.0165$. This implies annual time-averaged consumption growth with mean 1.8% and standard deviation 2.7%. We also consider a power utility version of this economy where $\gamma = 1/\psi = 10$ (see Table 1). We refer to this as the 'Power' calibration.

Table 1 - Parameter values for Exchange Economy

Table 1: The top half of this table gives the preference parameters used in the two calibrations of the model. 'EZ' refers to the calibration where agents prefer early resolution of uncertainty, while 'Power' refers to the calibration where agents are indifferent to the timing of the resolution of uncertainty. The bottom half of the table gives the value for the parameter that governs the dispersion of initial beliefs of the Young about the mean growth rate of the economy, A_0 , as well as the upper and lower truncation bounds for the otherwise Normal prior. The numbers correspond to the quarterly frequency of the model calibration.

Preference parameters:	EZ	Power
γ (risk aversion parameter)	10	10
ψ (elasticity of intertemporal substitution)	1.5	1/10
β (quarterly time discounting)	0.994	0.994
Priors:	EZ	Power
A_0 (dispersion parameter for prior at birth)	0.025	0.025
\bar{m} (upper truncation point of prior)	1.45%	1.2%
\underline{m} (lower truncation point of prior)	-0.55%	-0.3%

Malmendier and Nagel (2013) estimate the sensitivity of the young (at age 30) to updates in beliefs from model learning to be about 2.5% of the size of the macro shock

⁵Collin-Dufresne, Johannes, and Lochstoer (2013a) show that uncertainty about parameters governing the persistence of bad states is especially risky to the Epstein-Zin agent. Thus, it is likely that with such more complicated learning problem, the coefficient of relative risk aversion could be decreased substantially. However, this comes at the cost of several additional state variables and we leave this for future research.

(in their case, quarterly inflation). Towards the end of their life (at age 70), the old have a sensitivity of about 1%. We set the initial value of A_t for each young generation to 0.025. This implies that the initial size of the update in beliefs of the young in response to a macro shock is about 2.5% of the size of the macro shock, consistent with the estimate of Malmendier and Nagel.⁶ The old then have a posterior sensitivity to shocks of 0.5% of the size of the shock, somewhat lower than that estimated by Malmendier and Nagel. Note that a different learning problem, for instance learning about a persistence parameter, would lead to slower learning relative to the simple learning about the mean case that we consider here. The micro estimates from Malmendier and Nagel do not correspond directly to the learning problem we consider, both since they allow for a non-Bayesian learning scheme and because they consider a different model (not just learning about a mean parameter). These issues notwithstanding, we view our calibration of the bias in beliefs as reasonable relative to the available micro estimates of the generational 'this time is different'-bias.

In order to ensure existence of equilibrium for the benchmark model, we truncate the prior for the mean growth rate for all agents to have an upper (lower) bound of 1.45% (-0.55%) per quarter, while the true mean is 0.45% per quarter.⁷ Given that the prior standard deviation about the quarterly mean growth rate when born is $\sqrt{0.025} \times 1.65\% = 0.26\%$ these bounds are quite wide and not too tight to materially affect the update in mean beliefs relative to the case of an untruncated Normal prior.

The equity claim is a claim to an exogenous dividend stream specified as in Campbell and Cochrane (1999):

$$\Delta d_{t+1} = \lambda \Delta c_{t+1} + \sigma_d \eta_{t+1}, \quad (12)$$

⁶This calculation is based on the following. With a prior $\mu \sim N(m_0, A_0 \sigma^2)$, the subjective consumption dynamics for the next period are:

$$\Delta c_1 = m_0 + \sqrt{A_0 + 1} \sigma \tilde{\varepsilon}_{t+1},$$

and the update in belief can be written:

$$m_1 = m_0 + \frac{A_0}{\sqrt{A_0 + 1}} \sigma \tilde{\varepsilon}_{t+1}.$$

Thus, the sensitivity of the update in mean beliefs to the macro shock when $A_0 = 0.025$ is $\frac{0.025}{\sqrt{0.025+1}} \approx 0.025$.

⁷For the power utility specification, we use 1.20% (-0.05%). This difference in the bounds is necessary in order for equilibrium to exist given the much lower EIS in the power utility model.

where $\lambda = 3$ is a leverage parameter and $\sigma_d = 5\%$ is the volatility of idiosyncratic dividend growth.

3 Results from the Exchange Economy Model

We first describe the dynamic portfolio allocations and implied risk-sharing of the two agents and thereafter focus on the asset pricing implications of the model.

3.1 Portfolio allocation and risk-sharing

An unsurprising outcome of the 'this time is different'-bias is that the more optimistic agent will tend to hold a large portfolio share in risky assets that pay off in good states. Before we describe the portfolio allocations a couple of definitions are in order. First, total wealth in the economy is the value of the claim to aggregate consumption. Second, since we solve a discrete-time, complete markets problem where the shock has a continuous support, the complete portfolio choice decisions of agents involve positions in principle in an infinite set of Arrow-Debreu securities. To convey the portfolio decisions of the agents in a simple (first-order) manner, we define the weight implicit in the total wealth portfolio of agent i as the local sensitivity of the return to agent i 's wealth to a small shock to total wealth (as arising from an aggregate consumption shock close to zero). In the continuous-time limit, this local sensitivity is exactly the current portfolio allocation of agent i in the total wealth portfolio (because in this case, markets would be dynamically complete with two assets).

Figure 1 plots for the 'EZ' calibration the portfolio weight of the Old in the total wealth portfolio versus the mean belief about the mean growth rate of the Young and the Old, summarized by $m_{Old,t}$ and $m_{Young,t}$. The support plotted for the mean beliefs correspond roughly to \pm three standard deviation bounds of their beliefs over their lives. Note that the beliefs of the two agents tend to move together, so the corners of the surfaces that correspond to large differences in beliefs are quite unlikely. In the left-hand plot, the current wealth of the two agents is set equal. Further, the agents are both assumed to be in the 'middle' of their respective generations. Thus, the Young are have been living for 10 years and the Old for 30 years (remember, the financial life of an agent is 40 years from birth to death). The right-hand plot shows the case where the current wealth of the Old is only 10% of aggregate wealth.

Two clear patterns emerge. First, the portfolio weight is higher the more optimistic an agent is and the more pessimistic the other agent is, as one might have expected. Second, the fluctuations in the portfolio share resulting from differences in beliefs is larger for the agent with less wealth. The latter is intuitive as well—for example, as the wealth of an agent approaches all the wealth in the economy, the portfolio weight must approach 1.

Figure 1 - Portfolio allocation

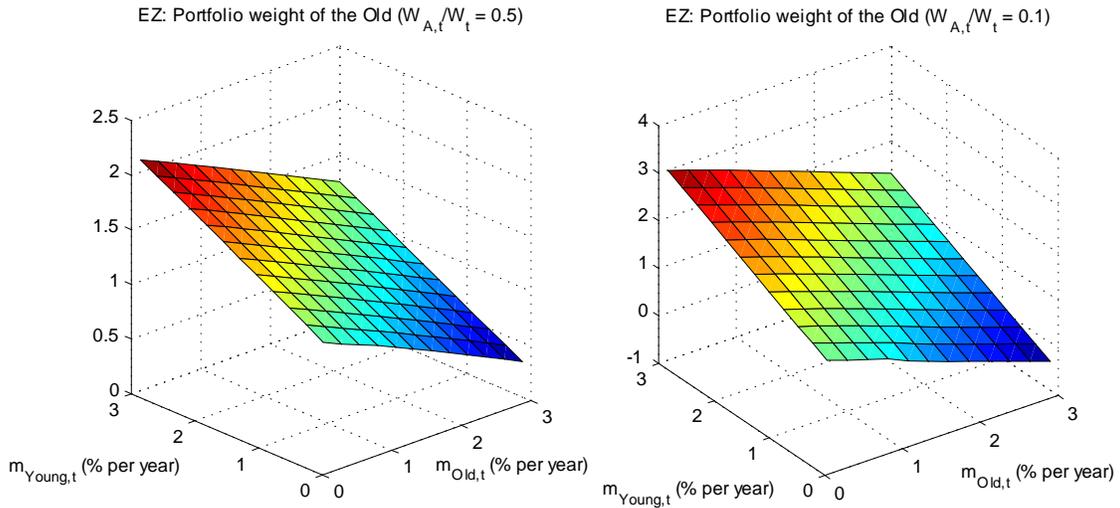


Figure 1: The plots show the portfolio weight that the Old allocates to the total wealth portfolio as a function of the mean belief of the Young as well as the Old. Both plots are given when the agents are in the middle of their respective generations (i.e., at 10 years (the Young) and 30 years (the Old) since 'birth'). In the left hand plot the current level of wealth is equal across agents, while in the right hand plot the wealth of the old are only 10% of total wealth. The mean beliefs about the growth rate in the economy are annualized and the portfolio weight is expressed as a regular number (i.e., not as a percentage—thus, 1.5 means a 150% portfolio weight in the wealth portfolio. As explained in the paper, the allocations are calculated as the local sensitivity of the wealth of the 'Old' to a shock to total wealth.).

The risk-sharing between the agents is not just a function of their mean beliefs, however, but also the posterior variance of beliefs. Differences in the latter induces interesting differences between the 'EZ' and the 'Power' calibrations. The left plots in Figure 2 shows how the risk-sharing operates in this economy. In particular, the change in the relative wealth share of the Young is plotted against realizations of the aggregate shock (log aggregate consumption growth). Again, the current wealth of the

two agents is assumed equal and both agents are in the middle age of their respective generations, as described earlier. In addition, the current mean beliefs of the Old are assumed to be unbiased, $m_{Old,t} = \mu$.

Two features of the model stand out. First, the case where the beliefs of the Young also are unbiased (the solid line) shows that the Old are in fact insuring the Young against bad states even when the mean beliefs coincide. This happens for both the 'EZ' calibration (upper plot) and the 'Power' calibration (lower plot) and is because the Young perceive the world to be more risky than the Old as their beliefs about the mean growth are less precise than that of the Old. Second, it is clear in the unbiased case that the 'EZ' case has the Old insuring the Young to a larger extent than for the 'Power' case. This is due to the fact that in the 'EZ' calibration, where agents have a preference for early resolution of uncertainty, model uncertainty is perceived as much more risky than in the 'Power' case (see Collin-Dufresne, Johannes, and Lochstoer (2013a)). As before, if the Young are sufficiently optimistic (here, about 2 standard deviations above the mean over a life time), the Young in fact are insuring the Old who are more pessimistic, and vice versa for the case where the Young are pessimistic.

The right-hand plots of Figure 2 show the portfolio weight of the Old versus the Young over the span of a generation (80 quarters). The wealth is held equal across the two agents and the beliefs of both agents are assumed to be unbiased over time. For the 'EZ'-case, the Old start with a portfolio allocation of 1.4 (140%) to the total wealth portfolio, while the Young starts with 0.6. Subsequently both are pulled towards 1 as the difference in the dispersion of beliefs decreases over time. This is an artifact of Bayesian learning in this case, as can be seen from Equation (3), where the variance of beliefs decreases more rapidly when prior uncertainty is high than low. Right before the generational shift, there is still a substantial difference, about 1.15 versus 0.85.

The 'Power'-case, however, is markedly different. First, portfolio weights barely budge over time and they are quite close, about 1.05 versus 0.95. Second, it is the *Young* who is more exposed to the total wealth fluctuations and thus has a higher portfolio weight. This somewhat counter-intuitive result is due to the fact that total wealth covaries *positively* with marginal utility in the 'Power' case due to the low level of the elasticity of substitution. For instance, an upward update in the mean belief of the growth rate, due to a positive consumption shock, *lowers* the price/consumption ratio as the wealth effect dominates. This effect is strong enough to make the return to

Figure 2 - Risk-sharing and portfolio allocations over time

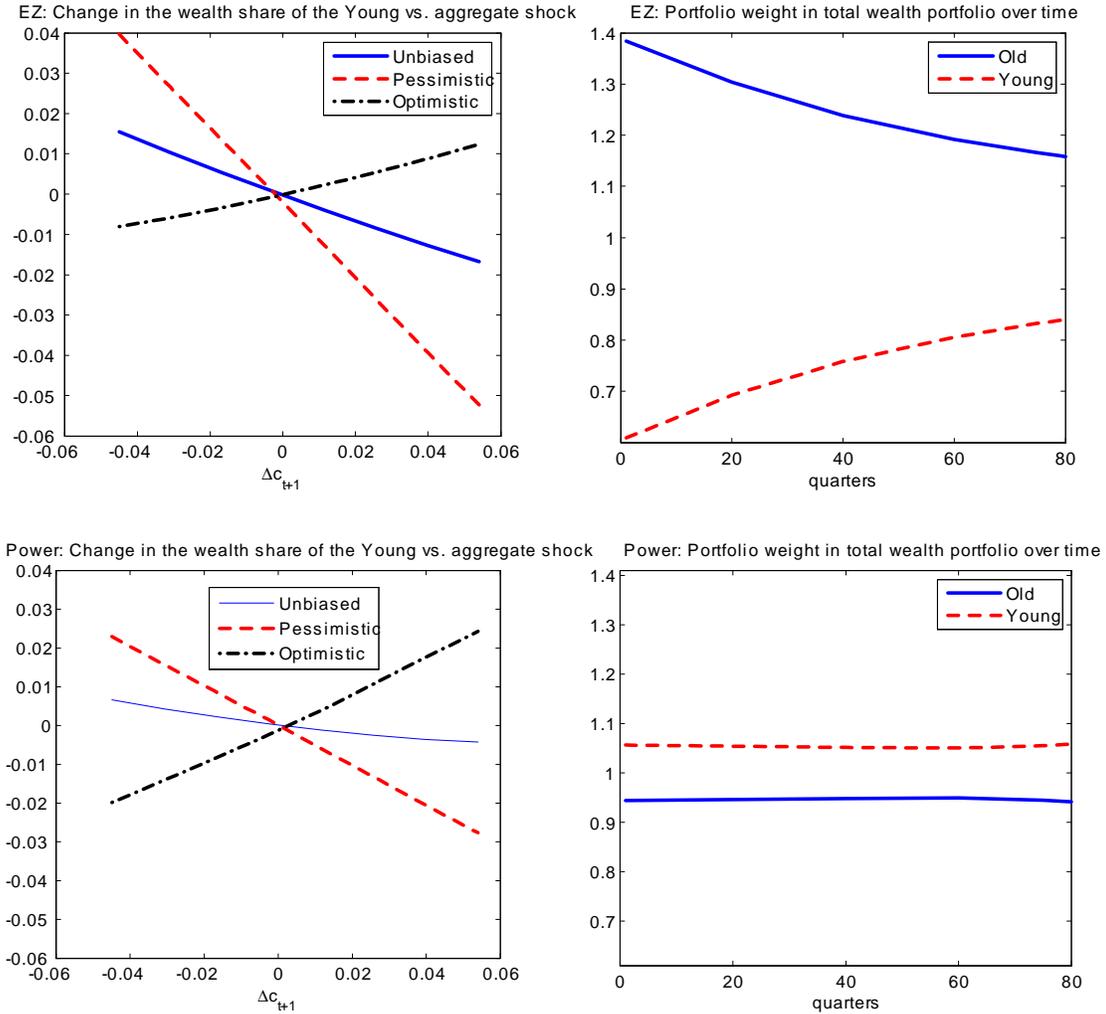


Figure 2: The left plots show the change in the wealth share of the Young for different realizations of the aggregate shock (consumption growth). The current wealth of the agents is set equal, the current age of the Young and the Old are in the middle of their generations (at 10 and 30 years, respectively), and the current beliefs of the Old are unbiased. The solid line shows the change in the relative wealth share when the current beliefs of the Young are also unbiased, whereas the red dashed line shows the case where the Young are pessimistic (the belief of the mean growth rate 2 standard deviations below the true mean), and the dash-dotted line shows the case where the Young are optimistic (2 standard deviations above the true mean)). The right plots show the portfolio allocation of the Young and the Old agent over time (from 1 to 80 quarters), where beliefs are held unbiased and the wealth-share is held equal across agents. The top plots show the result from the 'EZ' case whereas the bottom plots show the result from the 'Power' case.

total wealth positively related to marginal utility. The Young still perceive model risk as higher than the Old (remember, the subjective consumption dynamics are $\Delta c_{t+1} = m_{i,t} + \sqrt{1 + A_{i,t}}\sigma\varepsilon_{t+1}$, so this follows since $A_{Young,t} > A_{Old,t}$), but given the negative correlation between total wealth returns and aggregate consumption growth and the resulting negative risk premium on the total wealth portfolio, the Old hedges the Young by holding *less* of the total wealth claim.

The reason the portfolio share does not move much over time (holding beliefs and wealth constant) in the 'Power' case relative to the 'EZ' case is because with power utility, and the indifference to the timing of resolution of uncertainty, model uncertainty is simply not very important for total welfare. Thus, while the Young experience more model uncertainty relative to the Old, neither agent care very much about it. In the 'EZ' case, however, the agents experience large utility losses from being faced with model uncertainty. This is because of the subjective long-run risk induced by the model learning (see Equation (2)). The amount of long-run risk is proportional to the size of the update in mean beliefs, which from Equation (2) is $A_t\sigma$. In the middle of their respective generations, the relative difference in perception of short run risk ($\sqrt{1 + A_t}\sigma$) between the Young and the Old is 0.3%, while the relative difference in long-run risk ($A_t\sigma$) is 67%. Thus, the 'EZ' calibration leads to a much bigger difference in perceived risk between the two agents, which is also why both the dynamics of portfolio allocation and asset prices are much more pronounced in this case.

In sum, the Old hedge the larger amount of model uncertainty perceived by the Young. In the 'EZ'-case, model uncertainty is a first-order source of risk and therefore the difference in average portfolio allocation between the Young and Old is much larger in this case.

3.2 Asset pricing implications

Figure 3 shows the annualized price of risk (the conditional volatility of the pricing kernel divided by its conditional mean) for the 'EZ'- and the 'Power'-cases under the true probability measure—i.e., from the perspective of an agent that knows the true mean—versus the mean belief of the old. In the left-hand plots the wealth is equal across agents, and the solid line shows the case when the Young are pessimistic, while the dashed line shows the case where the Young are optimistic. The support of beliefs

Figure 3 - The Conditional Price of Risk

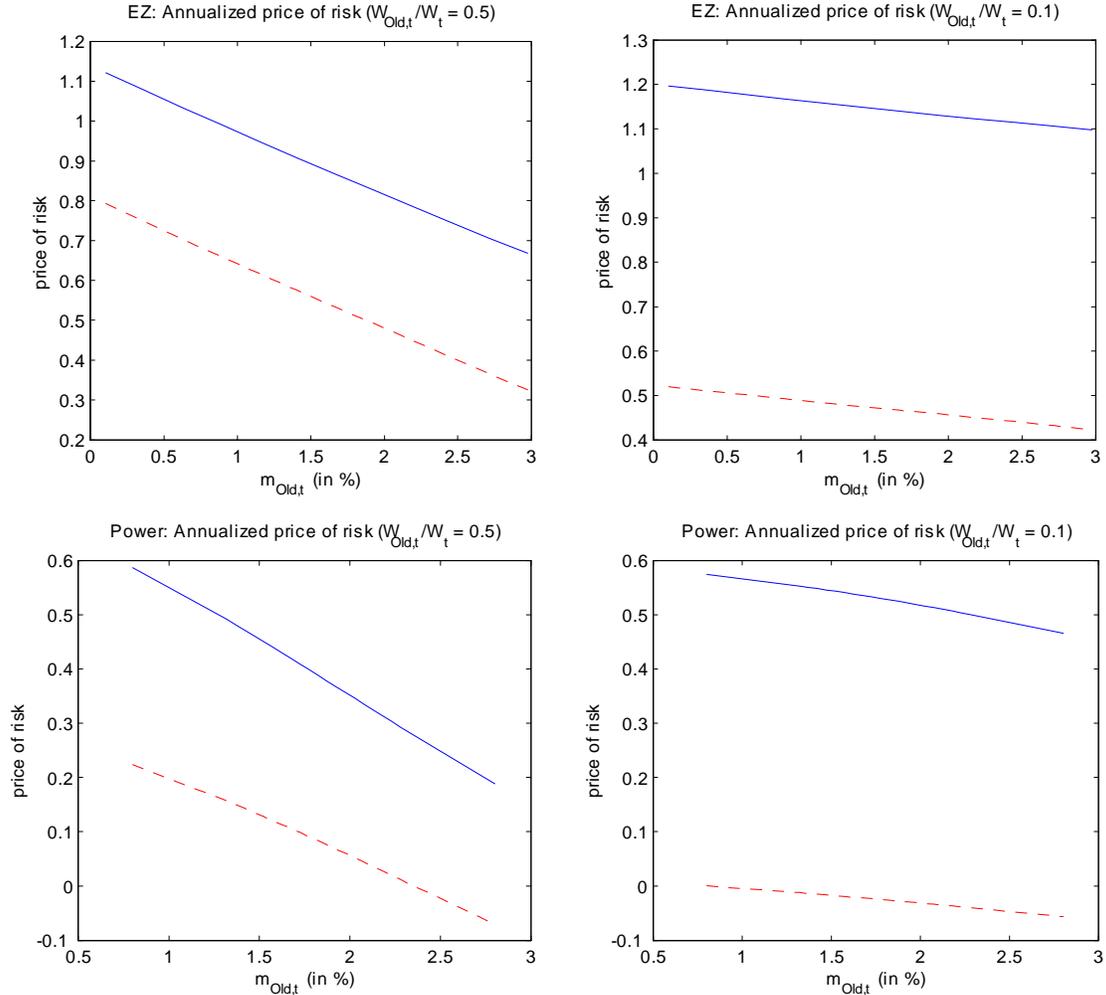


Figure 3: The figure shows the annualized conditional price of risk (the conditional volatility of the stochastic discount factor) for various configurations of the state variables in the economy. This is calculated under the true measure (i.e., using the conditional state probabilities of an agent that knows the true mean). In the plots, the x-axis has the current beliefs about the mean growth rate of the Old, while the solid line is for pessimistic beliefs of the Young and the dashed line is for the case of optimistic beliefs of the Young. The current age of the Young and the Old are in the middle of their generations (at 10 and 30 years, respectively). The left plots show cases where current wealth is equal across agents, while the right plots show cases where the current wealth of the Old is only 10% of total wealth. The top plots show results from the 'EZ' case whereas the bottom plots show results from the 'Power' case.

of both agents are plotted over approximately a ± 3 standard deviation range. First, for both agents, more optimistic (pessimistic) beliefs about the mean growth rate leads

to a lower (higher) price of risk under the true measure as the error in expectation will be reflected in the maximum Sharpe ratio available.

For both the 'EZ' and the 'Power'-cases, the variation in the price of risk in response to changes in beliefs is roughly the same. However, the level of the price of risk is on average about twice as high in the 'EZ' economy, despite the risk aversion parameter being the same in both calibrations. This is due to the endogenous subjective long-run risks generated from the parameter learning that are priced risks when agents prefer early resolution of uncertainty. Thus, the 'this time is different'-bias generates substantial additional macro risks in this case.

Figure 4 - The Risk Premium on the Dividend Claim over a Generation

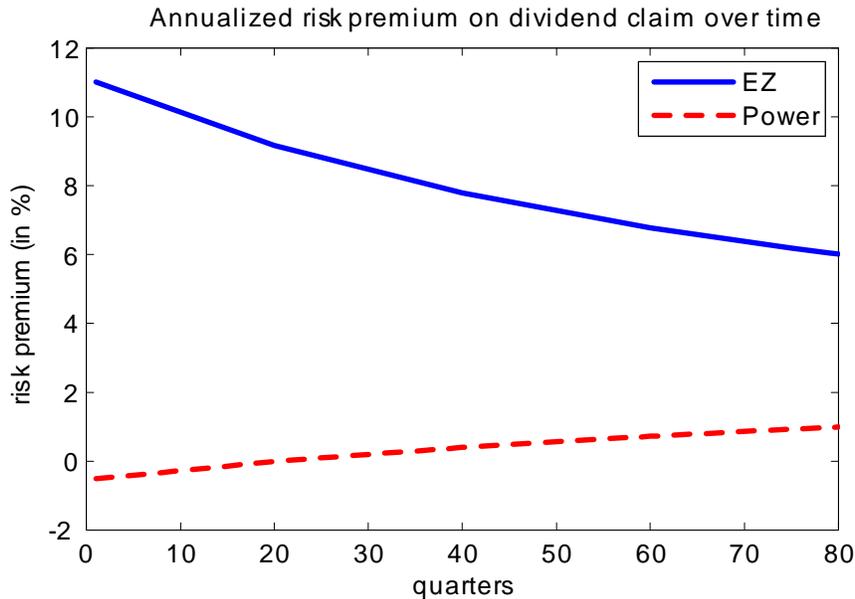


Figure 4: The figure shows the annualized conditional risk premium on the dividend claim over time. Quarter 1 is when a new generation of Young are born and the previously Young transition to the new Old, while Quarter 80 is the last date the Old are alive. The wealth of the agents is set equal at all dates, and beliefs are held unbiased (i.e., $m_{i,t} = \mu$). The solid line corresponds to the 'EZ' case, while the dashed line corresponds to the 'Power' case. The risk premium is calculated using the true probability measure (i.e., using the conditional state probabilities of an agent that knows the true mean).

From the price of risk dynamics, it may appear that one could simply increase γ for the 'Power'-case to obtain the same level of risk premiums and Sharpe ratios as in the 'EZ'-case. This is not true, however. Figure 4 shows the annualized risk premium

for the dividend claim over the course of a generation (80 quarters) where the beliefs of both agents are assumed unbiased and where their wealth is held equal. For the 'EZ'-case, the risk premium is high, as in the data, and slowly declining as learning occurs over the life of the two agents currently alive.

In the 'Power'-case, however, the risk premium is small, in fact negative to begin with, and slowly increasing over the sample. The negative risk premium is due to the low elasticity of intertemporal substitution in this calibration. In particular, since the wealth effect dominates, an upward (downward) revision in the mean belief about μ leads to a lower (higher) price-dividend ratio. When there is enough uncertainty about the mean, as in the start of a generation when the update in beliefs is sufficiently large, this effect dominates the effect of the dividend payout and returns on the dividend claim are negatively correlated with aggregate consumption growth. Overall, the absolute value of the risk premium is low in this calibration as agents that are indifferent to the timing of uncertainty do not suffer much of a utility loss from the model uncertainty we consider here.

In sum, the power utility preferences in combination with a 'this time is different'-bias leads to undesirable asset pricing implications. On the contrary, the 'EZ'-calibration, which has both pro-cyclical valuation ratios due to the high elasticity of intertemporal substitution and a high price of risk for model uncertainty (effectively $\gamma - 1/\psi$), yields asset pricing implications in line with observed aggregate asset price data. Figure 5 shows the annualized risk premium and Sharpe ratio of the dividend claim as a function of agents' mean beliefs. Both the risk premium and Sharpe ratio are high, of reasons explained above, and move counter-cyclically as agents tend to be pessimistic after a sequence of negative shocks to consumption growth.

In the following , we present asset pricing results from feeding the actual post-WW2 real, per capita U.S. quarterly consumption growth series (nondurables + services) from 1947Q2 to 2009Q4 into the model. We normalize the shocks to have zero mean and unit variance so that the mean and volatility of consumption within the model corresponds to that in the Bansal and Yaron (2004) calibration. The reason for using these shocks are thus mainly to have a time series that can be related to historical data in an intuitive way. We start both agents with mean beliefs set equal to the true mean. We let agent A start as the Young and agent B as the Old. We set the sample starting age (conditional on being Young or Old) of the agents to all possible values

Figure 5 - The Conditional Risk Premium and Sharpe ratio of Dividend Claim

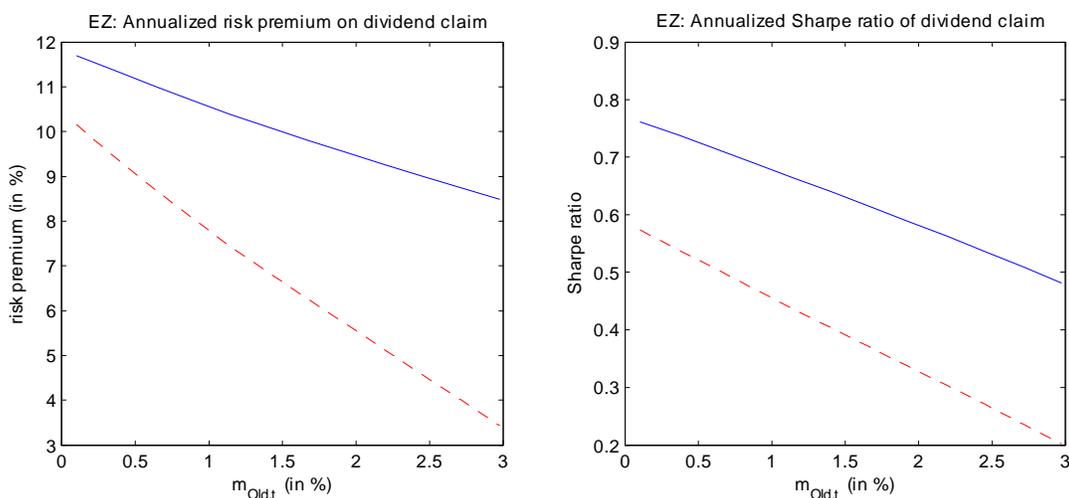


Figure 5: The left plot shows the annualized conditional risk premium on the dividend claim versus the mean belief of the Old. The solid line corresponds to the case of pessimistic beliefs of the Young and the dashed line is for the case of optimistic beliefs of the Young. The current age of the Young and the Old are in the middle of their generations (at 10 and 30 years, respectively), and the wealth is held equal across agents. The right plot shows the conditional annualized Sharpe ratio of the dividend claim. Both quantities are calculated using the true probability measure (i.e., using the conditional state probabilities as perceived by an agent who knows the true mean).

and then average across these when presenting statistics from the model. This is to avoid focusing the analysis on the discrete events associated with the death and birth of a new generations. The assumption of only two sets of agents alive at the same time was made purely for tractability and not as an important part of the model.

The top Panel Figure 6 shows the annualized mean beliefs of both Dynasties (A and B) over the sample. The graph shows how on average beliefs differ. At the beginning of the sample beliefs are similar, by assumption given the same starting value, but over the sample beliefs diverge somewhat. Beliefs of the two agents do move together as a consequence of updating from the consumption shocks. In particular, the Great Recession around the financial crisis at the end of the sample sees a notable drop in the beliefs about long-term growth for both sets of agents.

The lower panel of Figure 6 shows the annualized price of risk for the both the 'EZ'- and the 'Power'-cases. With power utility the price of risk is about 0.25, whereas

Figure 6 - Mean beliefs and the price of risk

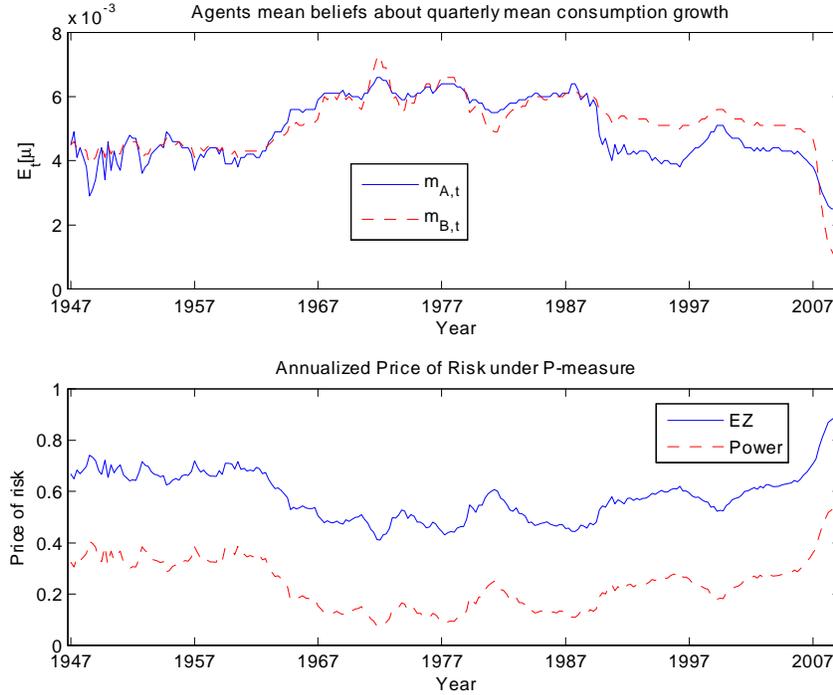


Figure 6: The top plot shows the mean beliefs of the agents over the post-war sample. In this illustration, both agent start the sample with unbiased beliefs and are then fed the actual, post-war consumption growth series (1947Q2 to 2009Q4). Agent A is the first Young generation, while B is the first Old (this changes during the sample as, e.g., A becomes the new Old after 20 years, and so on). The bottom plot shows the annualized conditional price of risk for both the 'EZ' calibration (solid line) and the 'Power' calibration (dashed line) over the sample. The price of risk is calculated using the true probability measure (i.e., using the conditional state probabilities as perceived by an agent who knows the true mean).

for the case of Epstein-Zin preferences it is about 0.7. This difference is in large part due to the endogenous long-run risks introduced by parameter learning, as in Collin-Dufresne, Johannes, and Lochstoer (2013a). The time-variation in the price of risk, which is viewed under the objective measure of the true model with known mean, corresponds closely to the time-variation in beliefs as seen in the upper panel. In particular, when agents are pessimistic the effective price of risk is high, whereas when they are optimistic the effective price of risk is low as future cash flows disappoint. The time-variation in the price of risk can be quite large and reaches its maximum during the financial crisis for both models.

It is notable that the price of risk in the power utility model is on average less than in the benchmark model with no parameter uncertainty, where it is $\gamma\sigma(\Delta c) = 0.33 > 0.25$. This is due to the this time is different feature of the model which enables better risk sharing among the agents. Thus, there is a trade-off in the model between the extra risks associated with parameter uncertainty and learning relative to the decrease in 'risk' due to the agree-to-disagree feature of the belief formation.

Figure 7 - The equity risk premium and return volatility

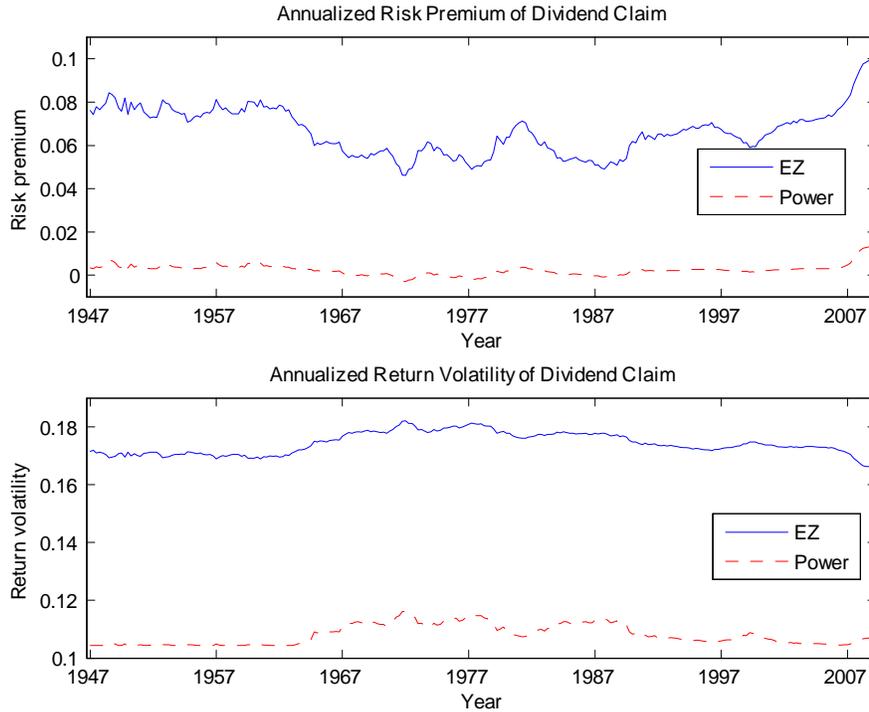


Figure 7: The top plot shows annualized conditional expected excess return on the dividend claim for the 'EZ' case (solid line) and the 'Power' case (dashed line). The lower plot shows the annualized, conditional return volatility of the dividend claim. Both quantities are calculated using the true probability measure (i.e., using the conditional state probabilities as perceived by an agent who knows the true mean). In this illustration, both agent start the sample with unbiased beliefs and equal wealth and are then fed the actual, post-war consumption growth series (1947Q2 to 2009Q4). The plotted quantities are the average time-series that obtains after averaging across all possible ages of the agents in the economy at the start of the sample.

Figure 7 shows the time path of the conditional, annualized equity risk premium along with the conditional annualized return volatility for both models. A striking

result is that the power utility model often yields a *negative* equity risk premium. This is again due to the low elasticity of intertemporal substitution. For the Epstein-Zin case, where the substitution effect dominates, the valuation ratio is strongly pro-cyclical, equity is risky, and the risk premium is therefore high, as in the data. Also, in this case there is substantial excess return volatility (17.4% p.a. versus dividend volatility of 11.5%). The equity premium increases and reaches its highest point during the financial crisis in excess of 8% p.a. In contrast, the power utility model has a negative and decreasing risk premium during the financial crisis. Again, the power utility model with model learning has a number of undesirable predictions that make it perform even worse than usual when faced with the standard asset pricing moments.

Table 2 shows the sample average moments of the two models versus those in the data and shows that the Epstein-Zin model with learning and this time is difference in fact does very well in terms of these standard moments. Notably, the 'EZ'-model has in the benchmark case with known mean growth rate a price of risk of 0.33 versus 0.58 achieved by this model with 'this time is different' learning. This contrast with the power utility specification which sees a decrease also in the price of risk when introducing the 'this time is different'-bias.

3.3 Mispricing: over- and undervaluation

In the benchmark economy with a known mean growth rate, the price-dividend ratio is constant. Thus, the time-variation in the price-dividend ratio can be characterized as mispricing from the perspective of a fully rational, long-lived agent that knows the true mean. This leads to excess return predictability as documented in predictive regression in the literature. Further, the optimal risk-sharing in the economy tends to exacerbate the pricing effects of biased beliefs in the Epstein-Zin economy. This is due to the endogenous wealth dynamics. In particular, with the Epstein-Zin preference parameters, the more optimistic holds more stock and a positive shock therefore has two effects – it increases the average belief about the mean growth rate in the economy and it shifts wealth to the more optimist agent, thus further increasing prices.

While these dynamics are also present in the power utility model, the amplification channel is much stronger when model risk is priced through the endogenous long-run risk channel. This can be seen in Figure 2, where the left plots show that show that

Table 2 - Unconditional Moments

Table 2: This table gives average sample moments from 247 quarters. The shocks to consumption are constructed from the real, per capita post WW2 nondurable and services data series from 1947Q2 to 2009Q4 and fed through each model. The 'data' column sample average moments for the U.S. from 1929 to 2011.

	Data 1929 – 2011	EZ : $\gamma = 10$ $\psi = 1.5, \beta = 0.994$	Power : $\gamma = 10$ $\psi = 1/10, \beta = 0.994$
$E_T [r_m - r_f]$	5.1	6.6	0.2
$\sigma_T [r_m - r_f]$	20.2	17.4	10.8
$SR_T [R_M - R_f]$	0.36	0.38	0.02
$E_T [r_f]$	0.86	1.9	17.9
$\sigma_T [r_f]$	0.97	0.4	3.32
$\sigma_T [M_{t+1}]$	> 0.36	0.58	0.24
$\gamma \times \sigma_T (\Delta c_{t+1})$		0.33	0.33
$E_T [\Delta c_{t+1}]$	1.8%	1.8%	1.8%
$\sigma_T [\Delta c_{t+1}]$	2.7%	2.7%	2.7%
$ExcessKurtosis_T [r_m - r_f]$	6.50	2.69	0.70

the relative wealth of the agents is more volatile in the 'EZ' case relative to the 'Power' case. In particular, Table 2 shows that excess stock returns in the former case display significant excess kurtosis as in the data, despite the fact that the aggregate shock is Normally distributed.

3.4 Return Predictability and Excess Volatility

As in Lewellen and Shanken (2002), this model also generates excess return predictability due to the parameter learning channel. However, the behavioral aspect adds to this by inducing over- and under-pricing. With the 'EZ'-calibration the dynamics of predictability are such that the price-dividend ratio predicts excess stock returns with a negative sign, as in the data. This is in contrast to the 'Power'-calibration where the sign in these benchmark regressions is positive.⁸

⁸These results are available upon request from the authors.

4 Asset Pricing in a 'This Time is Different' Production Economy

In the introduction, we noted that while there is micro evidence of an age dependent 'this time is different'-bias with respect to beliefs about macro dynamics, the average (or median) belief is an excellent predictor of future actual outcomes, indicating that the bias cannot be too severe. Survey data is not available for aggregate consumption dynamics, however, but it is available for both expected inflation and nominal GDP, which means we can construct the implied survey forecasts of *real* GDP. The latter can be related to output in a standard general equilibrium production economy.

In this Section, we show that agents' median GDP growth forecast is as good a predictor of future GDP as standard time-series GDP growth forecasting models, both in the data and in the model. Further, the production economy allows us to relate aggregate investment behavior to asset price over- and under-valuation. Before we turn to the production economy model, we first consider the accuracy of quarterly median real GDP growth forecasts from the Survey of Professional Forecasters versus a standard time-series forecasting model using GDP data.

The top left plot in Figure 8 shows realized, quarterly real U.S. GDP growth from 1972Q2 to 2009Q2 plotted versus the median forecast from the Survey of Professional Forecasters. The R^2 of a regression of the median forecast on next quarter's GDP growth is 31%, with a regression coefficient of 0.8 with 0.1 standard error. The top right plot shows the expected GDP growth obtained from an in-sample estimation of an AR(4) model on the GDP growth data plotted against the realized GDP growth data. In this case, the R^2 of the AR(4) is 37%. We can think of the AR(4) estimated expected GDP growth as the 'objective' time-variation in the expected growth rate versus the subjective growth rate given by the survey data. The R^2 's being quite similar indicates that the survey data in fact is a pretty good forecasting variable, especially considering that the AR(4) expected growth rate is determined in-sample. The bottom plot shows the two expected growth rates plotted together. They clearly co-move and the time-series standard deviation of the 'objective' measure from the AR(4) is only slightly lower than the standard deviation of the survey-based expectation (0.7% versus 0.8%).

These results indicate that also for GDP data, as for the inflation data as discussed in Ang, Bekaert, and Wei (2007), there is not a lot of evidence of extrapolative beliefs.

Figure 8 - Forecasts of quarterly real GDP growth

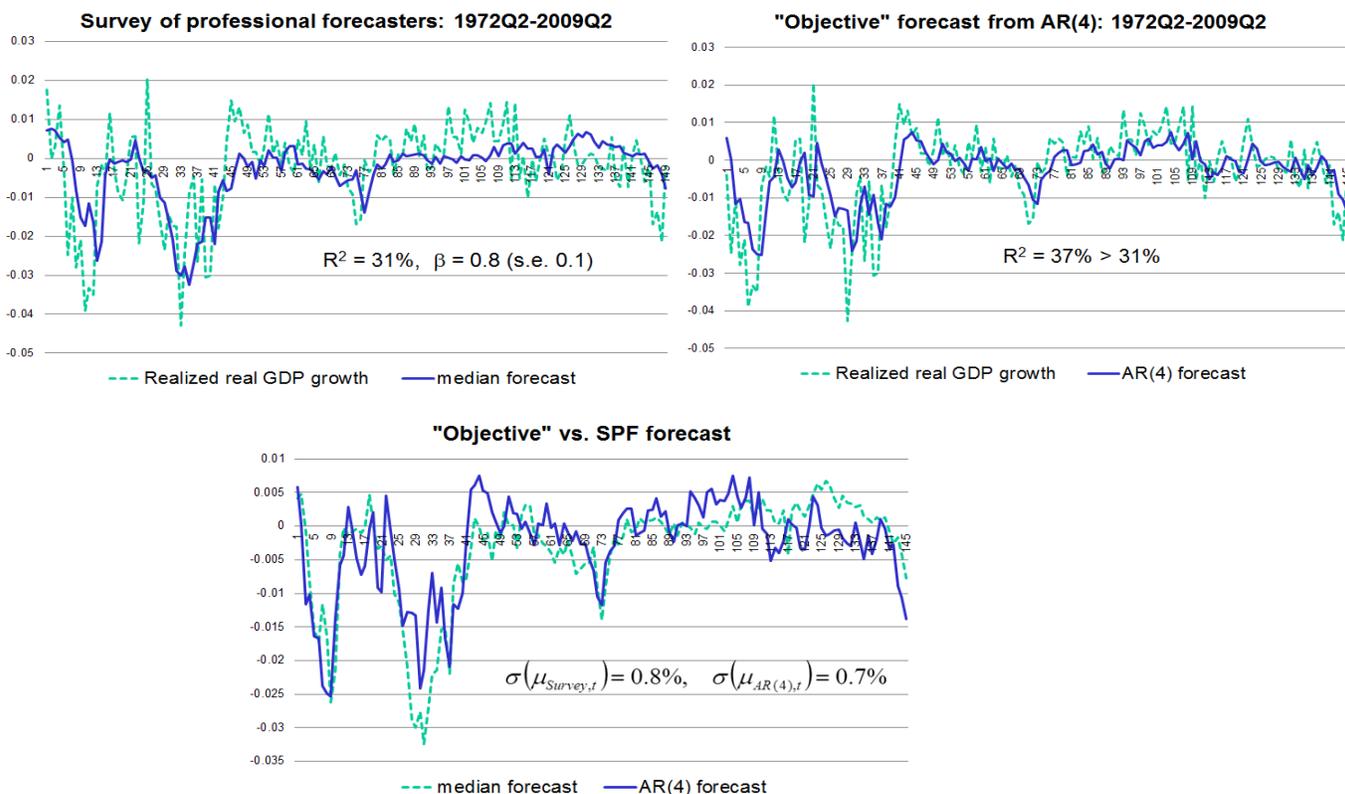


Figure 8: The top left plot shows realized quarterly real U.S. GDP growth from 1972Q2 to 2009Q2 (dashed line) versus the median quarterly forecast of real GDP growth (calculated using forecasts of inflation and nominal GDP) from the Survey of Professional Forecasters. The R^2 refers to a regression of realized on predicted growth, β is the regression coefficient, and s.e. refers to its standard error. The top right plot shows the same realized GDP growth plotted versus expected GDP growth obtained from running an AR(4) in sample on the realized GDP data. This is the 'objective', data-based forecast. The bottom plot shows the predicted growth obtained from the Survey and the AR(4) in-sample forecast. Here $\sigma(\mu_{Survey,t})$ and $\sigma(\mu_{AR(4),t})$ refer to the time-series standard deviation of the two forecasts.

Extrapolation implies that beliefs are substantially more volatile than what a Rational Expectations case would imply, but this is clearly not the case in this sample with respect to the GDP data. Next, we show that the production economy model with the 'this time is different;-bias as modeled in this paper can replicate this fact.

4.1 Production Function and Technology

Preferences, the overlapping generations setup, and belief formation are as before, with the only difference being that agents now learn about the mean growth rate of labor-enhancing technology, Z_t , which is assumed to be i.i.d.:

$$\Delta z_t = \mu + \sigma_z \varepsilon_{t+1}, \quad (13)$$

where $z_t \equiv \ln Z_t$. The production function of the representative firm is Cobb-Douglas:

$$Y_t = (Z_t N_t)^{1-\alpha} K_t^\alpha, \quad (14)$$

where labor supply N_t is assumed fixed at 1 and K_t is capital, and the capital accumulation equation is:

$$K_{t+1} = (1 - \delta) K_t + \phi(I_t/K_t) K_t, \quad (15)$$

where δ is the depreciation rate, $\phi(\cdot)$ is a concave function that governs capital adjustment costs, and I_t is gross investment. Thus, $Y_t = C_t + I_t$. If agents know μ , the economy reduces to the that in Kaltenbrunner and Lochstoer (2010). We let adjustment costs be quadratic with higher costs for downward adjustment:

$$\begin{aligned} \phi(x_t) &= x_t - \frac{a_t}{2} (x_t - \delta)^2, \\ a_t &= \begin{cases} a_- & \text{if } x_t < \delta \\ a_+ & \text{if } x_t \geq \delta \end{cases}, \end{aligned} \quad (16)$$

as in Zhang (2005).

The return to investment in this model is the same as the unlevered equity return (e.g., Cochrane (1991), Restoy and Rockinger (1994)):

$$R_{I,t+1} = \phi'(I_t/K_t) \left(\alpha (K_t/Z_t)^{\alpha-1} + \frac{1 - \delta + \phi(I_{t+1}/K_{t+1})}{\phi'(I_{t+1}/K_{t+1})} - \frac{I_{t+1}}{K_{t+1}} \right), \quad (17)$$

where the firm's first order condition over investment yields:

$$E_t^A [M_{t+1}^A R_{I,t+1}] = E_t^B [M_{t+1}^B R_{I,t+1}] = 1. \quad (18)$$

We report levered equity returns, following Hirshleifer and Yu (2013) as

$$R_{E,t+1} = (1 + B/E) R_{I,t+1}, \quad (19)$$

where B/E is the average debt to equity ratio in the data, set to $2/3$.

4.2 Calibration and Moments

Table 3 - Parameter values for Production Economy

Table 3: The table gives the parameters used in the calibrations of the production economy model. The numbers correspond to the quarterly frequency of the model calibration.

Preference parameters:	EZ – Production
γ (risk aversion parameter)	5
ψ (elasticity of intertemporal substitution)	2
β (quarterly time discounting)	0.994
Priors:	
A_0 (dispersion parameter for prior at birth)	0.025
\bar{m} (upper truncation point of prior)	1.5%
\underline{m} (lower truncation point of prior)	-0.5%
Technology and Production Parameters:	
α (the capital share)	0.36
μ (true mean growth rate)	0.5%
σ_z (volatility of technology shocks)	3.8%
δ (capital depreciation rate)	2.5%
a_- (downside adjustment costs parameter)	20
a_+ (upside adjustment costs parameter)	6

Table 3 shows the parameters chosen for the production economy. The preference parameter and learning parameters are the same as for the exchange economy considered in the previous section, with the caveat of risk aversion which set to half, at 5, as the production economy generates some long-run risks through the endogenous capital

accumulation and therefore requires less risk aversion to match the equity Sharpe ratio (see Kaltenbrunner and Lochstoer (2010)).

The parameters governing the production technology are standard and the volatility of technology shocks is set such as to match U.S. real per capita GDP growth volatility over the available 1929-2013 sample. The adjustment cost parameters are set such as to both be in line with empirical estimates (which are provide a quite wide range) and such that consumption growth volatility is the same as in the data over the same sample period.

Table 4 - Unconditional Moments: Production Economy

Table 4: This table gives average sample moments from 247 quarters. The shocks to consumption are constructed from the real, per capita post WW2 nondurable and services data series from 1947Q2 to 2009Q4 and fed through each model. The 'data' column sample average moments for the U.S. from 1929 to 2011.

	Data 1929 – 2011	Model
$\sigma_T [\Delta y]$ (%)	4.9	4.9
$\sigma_T [\Delta c] / \sigma_T [\Delta y]$	0.52	0.47
$\sigma_T [\Delta i] / \sigma_T [\Delta y]$	2.3	2.2
$E_T [r_m - r_f]$ (%)	5.1	3.1
$\sigma_T [r_m - r_f]$ (%)	20.2	6.0
$SR_T [R_M - R_f]$	0.36	0.51
$E_T [r_f]$ (%)	0.86	1.3
$\sigma_T [r_f]$ (%)	0.97	0.6

Table 4 shows the unconditional moments from the production economy model with the 'this time is different'-bias. The model matches the volatilities of output, consumption and investment very well. The model also matches the equity Sharpe ratio and the real risk-free rate mean and volatility well. The equity return volatility is, however, only 6%, which leads to a somewhat low equity premium of 3.1%. While this is short of the equity volatility in the data, it is an order of magnitude higher than the equity volatility in the model without the 'this time is different'-bias when agents known μ . The low volatility of the endogenous dividend claim is due to dividends in

this model being counter-cyclical as investment is strongly pro-cyclical, as explained in Kaltenbrunner and Lochstoer (2010). This is counter-factual. Danthine and Donaldson (2002) show that sticky wages is a feature of the data that if incorporated in the model creates operating leverage that helps match the dividend dynamics better to the data (see also Favilukis and Lin (2013)), which in turn generates higher return volatility. To keep the model simple, we do not add these features, but note that the 'this time is different'-bias helps tremendously in matching the 'difficult' return volatility moment in the production economy model.

This relative success is due to increase incentive to invest and disinvest in response to shocks as both agents update beliefs about the mean growth rate in the direction of the shock. This leads to higher investment volatility, which allows us to increase capital adjustment costs to make investment smoother to again match the data (see also Jerman (1998)).

Figure 9 shows that the time-variation in agent beliefs again lead to counter-cyclical risk premium and price of risk. Further, a low (high) risk premium under the true probability measure, due to optimistic beliefs, is associated with high (low) investment rates, which implies that investment forecasts future excess returns with a negative sign.

4.3 Output Predictability

In this final section, we turn to the predictability of output growth within the model, using both average agent beliefs and standard time-series models estimated on a sample of the same length as in the data. Figure 10 shows the true conditional mean of output growth for a simulated sample based on the same shocks as those used in the exchange economy of the previous section, as well as agents' mean belief about the conditional mean of output growth, average across agents.

As the graph shows, there is considerable comovement between these series, indicating the agents beliefs about output growth is quite accurate. This is due to the endogenous capital accumulation, which creates high output if agents believe technology growth is high, as they in these cases invest more aggressively.

In a more formal evaluation, using annual output growth from the simulated sample, we verify that the agents' average subjective expectation is a good predictor of future

Figure 9 - Conditional dynamics in production economy

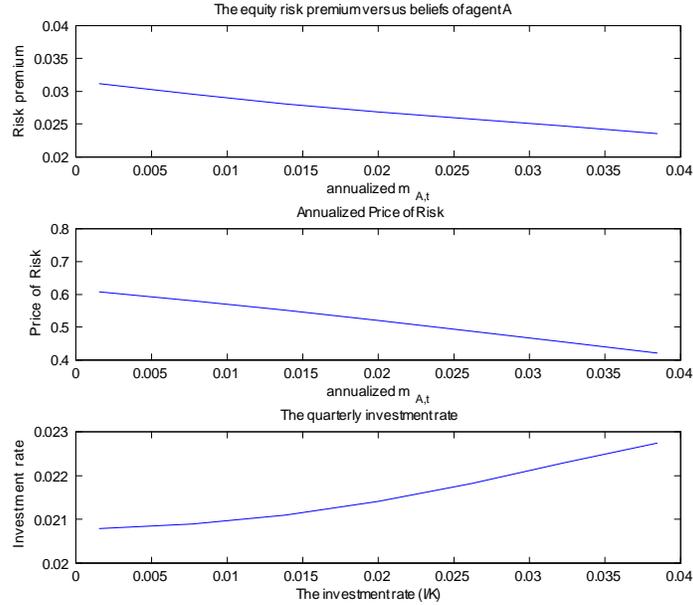


Figure 9: The top plot shows annualized conditional expected excess return on the dividend claim in the production economy versus the beliefs of agent A. All other state variables are at their average value. The middle plot shows the annualized conditional price of risk – the maximum Sharpe ratio in the economy. The price of risk and risk premium are calculated using the true probability measure. The bottom plot shows the quarterly investment rate, again as a function of the mean beliefs of agent A.

annual output growth relative to estimates of true output growth from an AR(4) model. In particular, the R^2 's and volatility of beliefs are very similar.

In sum, the 'this time is different'-bias is not excessive in this model along the following dimensions: (a) it corresponds to micro estimates, (b) average beliefs are not excessively volatile and as good a predictor of output growth as standard in-sample time-series estimates (from an AR(4)), and (c) it helps account jointly for quantities and asset prices in a production-based general equilibrium economy.

5 Conclusion

We have proposed a relatively simple, but quantitatively realistic model that incorporates the 'this time is different'-bias across generations documented by Malmendier and

Figure 10 - Expected output growth

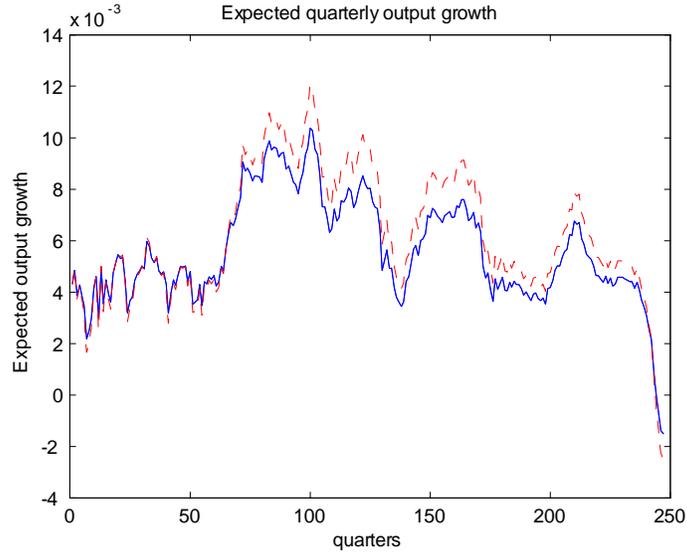


Figure 10: The figure shows a sample of expected quarterly output growth from the production economy. The solid line is the true (P-measure) expected growth rate, while the dashed line is the average agents' subjective expected growth rate.

Nagel (2011, 2013). In this model, model uncertainty persists indefinitely, and model uncertainty is an added risk factor due to the assumed preference for early resolution of uncertainty (Epstein-Zin preferences as in Bansal and Yaron (2004)). Importantly, consistent with Ang, Bekaert, and Wei (2007), as well as original empirical evidence presented in this paper, agents' beliefs are very good predictive variables for real output growth. Thus, the calibration of the bias is not excessive.

The pricing implications of the endogenous long-run risk introduced by such learning, as shown in Collin-Dufresne, Johannes, and Lochstoer (2013a), remain with the added feature of more time-variation in Sharpe ratios and expected excess returns. In particular, optimal risk-sharing between the agents induce fat tails in returns as over- and underpricing is exacerbated more as a consequence of optimal risk sharing.

When adding the calibrated 'this time is different'-bias in a production economy, the endogenous subjective long-run risk helps increase equity return volatility by an order of magnitude relative to the case where the mean growth rate is known. Thus, even though the departure from Rational Expectations is 'small,' the pricing implications are profound. In the model, over (under) valuation is associated with over (under) investment, as in the data.

References

- Ang, A., G. Bekaert, and M. Wei (2007), "Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?," *Journal of Monetary Economics* 54, 1163-1212.
- Baker, S., B. Hollifield, and E. Osambela (2014), "Disagreement, Speculation, and Aggregate Investment," Carnegie Mellon University working paper.
- Bansal, R. and A. Yaron (2004), "Risks for the Long-Run: A Potential Resolution of Asset Pricing Puzzles", *Journal of Finance* 59(4), 1481 - 1509
- Borovicka, J. (2013), "Survival and long-run dynamics with heterogeneous beliefs under recursive preferences", NYU working paper.
- Choi, J. and T. Mertens (2013), "Extrapolative Expectations and the Equity Premium," NYU and Yale working paper.
- Collin-Dufresne, P., M. Johannes, and L. A. Lochstoer (2013a), "Parameter Learning in General Equilibrium: The Asset Pricing Implications", Columbia University working paper.
- Collin-Dufresne, P., M. Johannes, and L. A. Lochstoer (2013b), "A Robust Numerical Procedure for Solving Risk-Sharing Problems with Recursive Preferences", Columbia University working paper.
- Detemple, J. (1986), "Asset pricing in a production economy with incomplete information." *Journal of Finance*, 41, 383–390.
- Doob, J. L. (1949), "Application of the theory of martingales." Coll. Int. du C. N. R. S. Paris, 23 - 27.
- Dumas, B., A. Kurshev, and R. Uppal (2009), "Equilibrium Portfolio Strategies in the Presence of Sentiment Risk and Excess Volatility," *Journal of Finance*, 64, 579 - 629.
- Ehling, P., A. Graniero, and C. Heyerdahl-Larsen (2013), "Asset Prices and Portfolio Choice with Learning from Experience," London Business School working paper.
- 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

- Garleanu, N. and S. Panageas (2012), "Young, Old, Conservative and Bold. The implication of finite lives and heterogeneity for asset pricing." UC Berkeley and Chicago Booth working paper.
- Gennotte, G. (1986), "Optimal Portfolio Choice under Incomplete Information," *Journal of Finance*, 61, 733-749.
- Geweke, J. (2001), "A note on some limitations of CRRA utility," *Economics Letters* 71, 341 - 345.
- Goyal A. and I. Welch, "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction," July 2008, *Review of Financial Studies* 21(4), 1455-1508.
- Hall, R. E (1988), "Intertemporal Substitution in Consumption," *Journal of Political Economy* 96, 339 – 357.
- Harrison J. and D. Kreps (1978), "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations," *Quarterly Journal of Economics*, 92, 323 - 336.
- Hansen, L. (2007), "Beliefs, Doubts and Learning: Valuing Macroeconomic Risk," Richard T. Ely Lecture, *The American Economic Review* 97, No. 2., 1 - 30
- Hansen, L. and T. Sargent (2001), "Robust control and model uncertainty," *American Economic Review* 91, 60-66.
- Hirshleifer, D. and J. Yu (2013), "Asset Pricing in Production Economics with Extrapolative Expectations", UC Irvine and University of Minnesota working paper.
- Jermann, U. (1998), "Asset Pricing in Production Economies," *Journal of Monetary Economics* 41, 257 - 275.
- Johannes, M., L. A. Lochstoer, and Y. Mou (2013), "Learning about Consumption Dynamics," forthcoming at the *Journal of Finance*.
- Johnson, T. (2007), "Optimal Learning and New Technology Bubbles," *Journal of Monetary Economics*, 87, 2486 – 2511.
- Kaltenbrunner, G. and L. A. Lochstoer (2010), "Long-Run Risk through Consumption Smoothing," *Review of Financial Studies* 23, 3141 - 3189.

- Kreps, D. and E. Porteus (1978), "Temporal Resolution of Uncertainty and Dynamic Choice Theory", *Econometrica* 46, 185 – 200.
- Lewellen, J. and J. Shanken (2002), "Learning, Asset-Pricing Tests, and Market Efficiency," *Journal of Finance* (57(3)), 1113 - 1145.
- Lucas, R. (1978), "Asset Prices in an Exchange Economy," *Econometrica* 46, 1429-1446
- Pastor, L. (2000), "Portfolio Selection and Asset Pricing Models," *Journal of Finance* 55, 179 - 223
- Pastor, L. and P. Veronesi (2006), "Was there a NASDAQ bubble in the late 1990s?," *Journal of Financial Economics*, 81, 61 – 100.
- Pastor, L. and P. Veronesi (2009), "Learning in Financial Markets," *Annual Review of Financial Economics*.
- Timmermann, A. (1993), "How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices." *Quarterly Journal of Economics*, 1993, 108, 1135-1145.
- Veronesi, P. (2000), "How does information quality affect stock returns?" *Journal of Finance*, 55.
- Working, H. (1960), "Note on the Correlation of First Differences of Averages in a Random Chain," *Econometrica* 28(4), 916 - 918.