Institut für Höhere Studien
Institute for Advanced Studies
Vienna

CONTINUING EDUCATION CENTER

MSc Economics

Learning Rare Disasters

A Master's Thesis submitted for the degree of
"Master of Science"

supervised by

Christian Haefke

Laszlo Tetenyi

1226494

Vienna, June 10, 2014.

MSc Economics

Affidavit

I, Laszlo Tetenyi,

hereby declare

that I am the sole author of the present Master's Thesis,

Learning Rare Disasters

49 pages, bound, and that I have not used any source or tool other than those referenced or any other illicit aid or tool, and that I have not prior to this date submitted this Master's Thesis as an examination paper in any form in Austria or abroad.

Vienna, June 10, 2014.

Tétényi László

Signature

Contents

List of Figures

Abstract

The purpose of this thesis is to evaluate the performance of a macro-finance model equipped with Epstein-Zin utility, learning and disasters. The observed high price of risk in the United States economy is successfully explained, but the variances of the financial variables generated by the model are unrealistically low.

1 Introduction

"I do not know which makes a man more conservative - to know nothing but the present, or nothing but the past."

-John Maynard Keynes (1926)

The purpose of my work is to merge the rare disaster and learning literature in order to evaluate their performance in explaining the financial movements of the United States economy. In one type of learning regime, the agent never forgets the disaster that had just happened, while the other type of learning allows the agent to almost completely erase it from her memory. The main focus then is on the extent the model can explain the premium or price of investing in risky financial assets. Traditionally, macro finance models predict a low premium while in the data a large price of risk is observed, therefore an extremely cautious behaviour must be explained. It is shown that the model presented here can partially do that without any unjustified deviation from the standard framework. The content is organised as follows : section 2 summarizes the learning and rare disaster literature while the theoretical framework and the computer implementation is discussed in section 3. The data description and the calibration is done in section 4, the performance of the model is shown in section 5 and in section 6 I conclude.

2 Literature Review

Explaining the large equity premium, that is, the difference between the risk free and the risky return has been a central issue in the macro finance literature, since Mehra and Prescott (1985). The standard Capital Asset Pricing Model (CAPM) fails to account for the premium for any reasonable parameter values. That is, assuming normal innovations, with zero mean and σ^2 variance to the dividend growth rate generates an equity premium of $\gamma\sigma^2$ where γ is the risk aversion parameter. Given the value for σ^2 is empirically around 0.00125 while the equity premium is 0.06 , the risk aversion γ should be around 48. There are at least three reasons why this value for γ is unreasonable : first, it would imply a very high risk free rate, second, micro studies suggest a maximum value of 3, see Barsky et al. (1997). Finally, such a value is grossly inconsistent with growth theory implying an extremely low growth rate, because if the utility function is of constant relative risk aversion (CRRA) type, γ also controls the inverse of the elasticity of intertemporal substitution[1].

There are various other asset pricing facts that are expected to be matched by a macro finance model. Although the equity premium is "large" it is also declining over time, while the price to dividend ratio is volatile and non-stationary as shown by Blanchard (1993),Jagannathan et al. (2000) and Fama and French (2002). In addition, the low volatility and return of the risk free asset and the high volatility of the risky return is also present in the data Barsky and De Long (1993). None of these facts can be explained by the standard CAPM with rational expectations, CRRA preferences, complete and frictionless asset markets as summarized by Kocherlakota (1996) and Mehra and Prescott (2003).

The first direction to modify the standard model is to use other utility functions then that of CRRA type. Epstein-Zin preferences, as shown by Epstein and Zin (1989) and Weil (1989), allow to separate risk aversion and intertemporal elasticity of substitution, therefore expanding the set of reasonable parameter values. Higher risk aversion is possible while the low inverse intertemporal elas-

[1] In the case of the Ramsey–Cass–Koopmans growth model, the balanced growth equals $\frac{1}{\gamma}(r(t) - \rho)$ where $r(t)$ is the return on capital and ρ is the depreciation

ticity of substitution allows for a lower risk free rate. The main usefulness of such preferences is in a setup where the consumption growth process is modelled as non stationary, as demonstrated by Bansal and Yaron (2004), justifying the choice in this paper of modeling consumption growth as a Markov process and taking the utility function as an EZW type.

The other direction is to increase the variance of the dividend growth rate. Since Rietz (1988) the rare disaster literature argues that there are catastrophic events that because of their infrequent realizations, do not affect significantly the mean of the growth rates,but are so dangerous to the agents in the economy that precautionary savings increase greatly, even for a moderately risk averse consumer. A critique of this approach is that such disasters - the mild scenario of Rietz (1988) consists of a one year drop of 30 % in consumption - are not observed in the data of the United States. The counter-argument of Barro and Ursúa (2008) is that the US has been lucky as these rare events did not realize while other countries were more unfortunate - nevertheless the US consumers fear that what happened to Argentina might happen to them too. In addition, as data might be missing or unreliable during disasters, such as wars and revolutions, the variance of the growth rate for the economies in the world are underestimated by this survivorship-bias. In this spirit Nakamura et al. (2013) estimates the size of the disasters using the new panel dataset of Barro and Ursúa and finds that $\gamma = 3$ can already explain a large portion of the equity premium.

Apart from the direct effect, an increase in the probability of a rare disaster alone could induce a recession. Shocks to the probability of the disaster indirectly explain various asset pricing facts as shown by Gourio (2012) and Gabaix (2012) where the standard framework is enriched with production and inflation, respectively. Barro and Ursúa (2012) is a survey of the rare disaster literature. In these models, to justify the exogenous shocks to the probability of the disaster an additional explanation is required — learning.

To understand why and how probabilities vary, one should think of subjective probabilities where the exact probability of an event is not known but is learnt about - along the transition when one learns about the probability, the estimate might fluctuate around the true value. As shown by Timmermann (1993) introducing Bayesian updating[2] to asset pricing models can help explaining a number of puzzles - but not the equity premium - given risk neutrality and normality of the stochastic dividend growth rate. Cecchetti et al. (2000) considers a Hidden Markov Model with two states as the generating process for consumption growth where the parameters of the transition matrix are not known. The learning is not optimal, the initial distorted prior persists. Bayesian learning in the same setup is implemented by Cogley and Sargent (2008) - with an exogenously given pessimistic prior the model can generate not only the high price of risk, but also it's decreasing rate. Here the same setup is extended to more states and a more general utility function, as Epstein-Zin is used instead of CRRA. Moreover it becomes possible to directly compare the implications of learning with a fixed-window -where the number of observations used to estimate the model parameters are constant in the spirit of Cecchetti et al. (2000) - with Bayesian updating. The main motivation of doing so is that switching between these learning regimes could possibly explain the oscillation between periods of non-stationary and stationary prices of risk. Learning about rare disasters is difficult therefore convergence to the rational expectations equilibrium is slow, justifying an "extreme" distortion of the prior transition matrix. Intuitively, suppose that the economies in the world are governed by the same Markov process but this is only known by the representative consumer, not by the econometricians of the US who use only US data to estimate consumption growth. Then having the data generating transition matrix as a prior for the consumer might seem as an extremely distorted prior from the viewpoint of the US econometricians. The interesting question is then to what extent such a model could possibly recreate the financial movements when the actual realizations of the states are "fed" in to it? It is shown in section 5 that the model is able to match a large fraction of the observed price of risk without assuming a large, exogenous deviation from rational expectations. The setup by Cogley et al. (2012) with a 3 state process where one type of consumer learns about the 3^{rd}, disaster state, is merging the disaster literature and learning to an extent, but the distortion of beliefs is still

[2]Least squares learning is used, but in the case that the parameters are constant and the process is Gaussian, these are equivalent

exogenous - although much less substantial than the distortion in Cogley and Sargent (2008) because uncertainty is present for the rare event only.

Although not modeled here — as the existence of a representative consumer is assumed — the wealth implications in an economy where the parameters of the model are unknown are discussed by Blume and Easley (2006). With complete markets the trader with "worse" learning process will eventually "die out" as their wealth will converge to zero. It is important to note that in a learning economy the subjective price of risk - the naive Sharpe Ratio - is not the one observed directly in the data, which implicitly assumes rational expectations when dealing with the future. Therefore Cogley and Sargent (2008), based on Hansen and Jagannathan (1991), construct a measure for the market price of risk in a learning economy that reconciles with rational expectations thereby referred to as the "price of risk".

3 Model Setup

Following Mehra and Prescott (1985) I study an endowment economy populated by an infinitely lived representative agent, endowed with preferences that are representable by a non-time-separable utility function U_t given in the following recursive form at time t:

$$U_t = ((1-\beta)C_t^{1-\theta} + \beta(\mathbb{E}_t^s(U_{t+1}^{1-\gamma}))^{\frac{1-\theta}{1-\gamma}})^{\frac{1}{1-\theta}}$$

as defined by Epstein and Zin (1989) and Weil (1989) (thereby referred to as EZW preferences). C_t is consumption in period t, β is the subjective discount factor, γ is the coefficient of relative risk aversion and θ is the inverse of the intertemporal elasticity of substitution[3]. As a special case, if $\theta = \gamma$ then the utility function becomes separable and of constant relative risk aversion (CRRA) type. The expectation operator $\mathbb{E}_t^s$ is the subjective expectation conditioned on the information set available to the consumer on date t. The dividend is assumed to follow an exogenous, stochastic process and is nonstorable - that is, there is no investment or government in the model so consumption equals dividend. Asset markets are complete and the only asset which is in non-zero supply is the stock of the Lucas (1978) tree. In this economy, the existence of the no-trade competitive equilibrium will be assumed which rules out both some stochastic processes for the dividend growth and learning algorithms as discussed by Weitzman (2007).

3.1 Subjective Euler Equation

To derive the subjective Euler equation characterizing the competitive equilibrium, the consumer has to solve the following maximization problem (given recursively as a Bellman problem):

$$\max_{C_t, h_{t+1}} U_t(h_t) = ((1-\beta)C_t^{1-\theta} + \beta(\mathbb{E}_t^s(U_{t+1}^{1-\gamma}))^{\frac{1-\theta}{1-\gamma}})^{\frac{1}{1-\theta}}$$

[3]for $\theta = 1$ the above formula is incorrect, but as in all specifications it is different than 1 , the correct formula for that case is skipped

st:

$$C_t + P_t h_{t+1} = D_t h_t + P_t h_t$$

where D_t is the dividend, h_t is the number of shares owned by the consumer (overall number of shares are normalized to 1) P_t is the price of shares at date t. Note that for simplicity all other assets - which in principle could exist as markets are complete - are suppressed from the budget constraint as once a stochastic discount factor is determined it can be used to price them. The derivations provided here closely follow Epstein and Zin (1991).

Theorem 1. *The stochastic discount factor defined as :* $m_{t,t+1} = \frac{\partial U_t/\partial C_{t+1}}{\partial U_t/\partial C_t}$ *is given by*

$$m_{t,t+1} = \beta(\frac{C_{t+1}}{C_t})^{-\theta}(\frac{U_{t+1}}{\mathbb{E}_t^s(U_{t+1}^{1-\gamma})^{\frac{1}{1-\gamma}}})^{\theta-\gamma} \tag{1}$$

Proof provided in Appendix A

Implementing Theorem 1 directly has a disadvantage: it increases the number of variables compared to the CRRA stochastic discount factor with the utility levels. Nevertheless if for example there is labor in the utility function, as in a Dynamic Stochastic General Equilibrium model then the above expression is the only useful formula (see Uhlig (2010)). Here however, as the model abstracts from labor, it becomes possible to derive a more intuitive expression for the stochastic discount factor.

Definition 1. The present discounted value of consumption - which equals the discounted overall income,referred as wealth, W_t in equilibrium - can be written recursively as:

$$W_t = C_t + \mathbb{E}_t^s(m_{t,t+1}W_{t+1})$$

Theorem 2.

$$W_t = \frac{U_t}{(1-\beta)C_t^{-\theta}(U_t)^{\theta}}$$

Proof provided in Appendix A

Definition 2. Return on stocks is given by:

$$R_{w,t+1} = \frac{W_{t+1}}{W_t - C_t}$$

Theorem 3.

$$R_{w,t+1} = \left\{\beta(\frac{C_{t+1}}{C_t})^{-\theta}(\frac{\mathbb{E}_t^s(U_{t+1}^{1-\gamma})^{\frac{1}{1-\gamma}}}{U_{t+1}})^{1-\theta}\right\}^{-1}$$

and therefore

$$m_{t,t+1} = \beta^{\frac{1-\gamma}{1-\theta}}(\frac{C_{t+1}}{C_t})^{-\theta\frac{1-\gamma}{1-\theta}}R_{w,t+1}^{\frac{\theta-\gamma}{1-\theta}}$$

Proof provided in Appendix A

Theorem 3 results in the subjective Euler equation:

$$\beta^{\frac{1-\gamma}{1-\theta}}\mathbb{E}_t^s[(\frac{C_{t+1}}{C_t})^{-\theta\frac{1-\gamma}{1-\theta}}R_{w,t+1}^{\frac{\theta-\gamma}{1-\theta}}R_{t+1}] = 1 \tag{2}$$

where R_{t+1} is the return of any one period asset.

3.2 Stochastic process

Equation 2 in general does not allow one to have a closed form solution of the model, even in the special case of time separable utility. However, by choosing the stochastic process for $\frac{C_{t+1}}{C_t}$ wisely, the solution can attain closed form or can be reasonably approximated by a closed form. The most common way is to assume lognormal innovations to consumption. There are two important reasons why I refrain from that path - both are connected to learning. First, in the case of even the slightest structural uncertainty as shown by Weitzman (2007), the posterior distribution of the discount factor can yield an infinite equity premium - that is, the competitive equilibrium does not exist. Therefore a model with such a stochastic process is not robust to deviations from rational expectations and can be seen as a knife-edge case. In addition, there are only two parameters to learn about and they are often conflicting when a positive shock occurs, resulting in an acyclic equity premium. In a more general setting, Barro and Ursúa (2012) show that a twisted lognormal distribution, with fatter tails approximately results in an equity premium:

$$r^e - r^f = \gamma\sigma^2 + pb(1-b)^{-\gamma}$$

where r^e is the log return on equity, r^f is the log risk free return, p is the probability of the disaster and b is the "growth" in the disaster state. If one does not know the true σ, the variance of the normal innovations, then the equity premium is increasing for both large positive and negative shocks as they both increase the perceived variance. In case the probability or the size of the disaster is unknown then the equity premium easily becomes $+\infty$ because of the previous, robustness problem.

That is why I choose a different path and specify the stochastic process as:

$$\frac{C_{t+1}}{C_t} = u_t \tag{3}$$

where the term u_t is assumed to follow a multi state, first order, ergodic, irreducible Markov process, the transition probabilities are given by the rows of Π and $u_t \in \{\lambda_1, \lambda_2, \ldots \lambda_N\} \ \forall \ t$.

3.3 Learning

Assumption 1. λ_1,λ_2,$\ldots \lambda_N$ *and all parameters apart from the elements of* Π *are known by the consumer.*

Assumption 1 seems reasonable if the agents in the economy are sure about to clustering of states — for example good or bad — and if there is uncertainty about the growth rate in a given state than it can be easily modeled by increasing the number of states in the perceived process by nesting the true model and avoiding the robustness problem of the normal distribution.

Before specifying the learning structure it is useful to derive the closed form solutions for the key asset prices and returns in case we also know the "true" Π.

Theorem 4. *Suppose that* Π *is known by the agent. In this case, the time independent, rational expectations equilibrium price to dividend ratio in state* i*, denoted by* w_i*, is given by*

$$w_i = \beta \cdot \Big(\sum_{j=1}^{n} \Pi_{ij} \lambda_j^{1-\gamma} \cdot (1 + w_j)^{\frac{1-\gamma}{1-\theta}} \Big)^{\frac{1-\theta}{1-\gamma}} \tag{4}$$

$\forall\ i \in \{1, 2 \ldots N\}$.

Proof provided in Appendix A

Notice that the implicit equation for the price to dividend ratio is linear in w if and only if $\gamma = \theta$, the CRRA time separable special case of the EZW preferences. In general, Equation 4 can only be solved for numerically as a fixed point problem transformed into root-finding. Once w_i has been solved $\forall\ i$, the asset returns can be determined. It is crucial to understand the assumption in Theorem 6 about the time independence of the price to dividend ratios. It does not say that the price to dividend ratio per se is independent of time - it only says that the price to dividend ratios for all possible states are independent of time so that the stacked w is time independent. It is shown in Theorem 5 that this type of time independence assumption is implied by the no-bubbles condition.

Corollary 1. *Suppose that $w \in \mathbb{R}^N$ is known. Then the following holds:*

$$\mathbb{E}_t^s R_{w,t+1} = \sum_{j=1}^{n} \Pi_{ij}\lambda_j \cdot \frac{1 + w_j}{w_i}$$

and the one period risk free rate $R_{f,t}$

$$R_{f,t} = \Big\{\beta^{\frac{1-\gamma}{1-\theta}} \cdot \Big(\sum_{j=1}^{n} \Pi_{i,j}\lambda_j^{-\gamma} \cdot \big(\frac{1 + w_j}{w_i}\big)^{\frac{\theta-\gamma}{1-\theta}}\Big)\Big\}^{-1}.$$

However the "true" Π is unknown. In the first place, assume Bayesian updating for the consumer. Let the prior distribution of the elements of Π given by:

$$(\Pi_{i,1}, \ldots \Pi_{i,N-1} | n_{i,1}^0 \ldots n_{i,N}^0) \sim Dir(N-1, n_{i,1}^0 \ldots n_{i,N}^0)$$

where $n_{i,j}^0$ is the prior knowledge of the transition from i to j. The notation is justified by the fact that the Dirichlet distribution has a probability density function defined on the $N-1$ dimensional open simplex, so that $\Pi_{i,N} = 1 - \sum_{j=1}^{N-1} \Pi_{i,j}$. The beta distribution is a special case for $N = 2$. The Dirichlet distributed random vectors are assumed to be independent $\forall i \in \{1, 2 \ldots N\}$. Then, as for given i the realization of states is:

$$(n_{i,1},\dots n_{i,N}) \sim Multinomial(N,\Pi_{i,1},\dots \Pi_{i,N})$$

where $n_{i,j}$ denotes the number of observed transitions from i to j. As the Dirichlet is the conjugate prior of the Multinomial distribution, the posterior distribution is also Dirichlet:

$$\Pi_{i,1},\dots \Pi_{i,N-1} \sim Dir(N-1, n^0_{i,1}+n_{i,1},\dots n^0_{i,N}+n_{i,N})$$

This relationship yields a simple updating rule and this is precisely why such a prior distribution is chosen. Suppose the consumer was at state i in period 0 and in the next period, the state indexed by j is observed, then her estimate $\hat{\Pi}$ becomes:

$$\hat{\Pi}_{i,j} = \frac{n^0_{i,j}+1}{1+\sum_{k=1}^{N}(n^0_{i,k})}$$

$\forall\, l \in \{1\dots j-1, j+1,\dots N\}$

$$\hat{\Pi}_{i,l} = \frac{n^0_{i,l}}{1+\sum_{k=1}^{N}(n^0_{i,k})}$$

and $\hat{\Pi}_{l,k} = \Pi_{0_{l,k}}$, $\forall\, l \in \{1\dots j-1, j+1,\dots N\}$ and $\forall\, k \in \{1\dots j-1, j, j+1,\dots N\}$ where Π_0 is the prior transition matrix given the histories $n^0_{i,j}$.

Using the posterior distribution of the transition probabilities, updated on the information available up to date t, lead to the evaluation of the Euler Equation 2 to obtain the price to dividend ratio w:

$$\beta^{\frac{1-\gamma}{1-\theta}}\mathbb{E}^s_t[(\frac{C_{t+1}}{C_t})^{1-\gamma}(1+w_{t+1})^{\frac{1-\gamma}{1-\theta}}] = w_t^{\frac{1-\gamma}{1-\theta}}. \tag{5}$$

3.4 Learning under CRRA utility and an upper bound for EZW

Assume first that $\theta = \gamma$. As shown before:

$$w_t = \mathbb{E}_t^s\Big[\hat{m}_{t,t+1}(1 + w_{t+1})\Big]$$

where $\hat{m}_{t,t+1} = \beta(\frac{C_{t+1}}{C_t})^{1-\gamma}$.

Theorem 5. *After iterating forward, using the law of iterated expectations and imposing the no-bubbles condition:*

$$w_t = \mathbb{E}_t^s \sum_{j=1}^{\infty} \prod_{s=1}^{j} \hat{m}_{t+s-1,t+s}.$$

Proof:

First, the law of iterated expectations with Bayesian updating holds, see for example appendix B by Cogley and Sargent (2008). Then iterating forward:

$$\begin{aligned}
w_t &= \mathbb{E}_t^s\Big[\hat{m}_{t,t+1}(1 + w_{t+1})\Big] \\
&= \mathbb{E}_t^s\Big[\hat{m}_{t,t+1}(1 + \mathbb{E}_{t+1}^s\Big[\hat{m}_{t+1,t+2}(1 + w_{t+2})\Big])\Big] \\
&= \mathbb{E}_t^s\Big[\hat{m}_{t,t+1} + \hat{m}_{t,t+1}\mathbb{E}_{t+1}^s\Big[\hat{m}_{t+1,t+2}(1 + w_{t+2})\Big]\Big] \\
&= \mathbb{E}_t^s\Big[\hat{m}_{t,t+1} + \hat{m}_{t,t+1}\hat{m}_{t+1,t+2} + \hat{m}_{t,t+1}\hat{m}_{t+1,t+2}w_{t+2}\Big] \\
&= \mathbb{E}_t^s \sum_{j=1}^{\infty} \prod_{s=1}^{j} \hat{m}_{t+s-1,t+s} + \mathbb{E}_t^s \prod_{j=1}^{\infty} \hat{m}_{t+j-1,t+j} w_{t+j}.
\end{aligned}$$

Imposing the no-bubbles condition that:

$$\prod_{j=1}^{\infty} \hat{m}_{t+j-1,t+j} w_{t+j} = 0$$

yields the result.

□

Notice that in case of rational expectations and CRRA preferences the time independence assumption of the price to dividend ratios is implied by the no-bubbles condition in Theorem 4.

Theorem 6. *Let us denote the rational expectations solution to Equation 4 for any* Π *matrix by* $w^*(\Pi)$*. Then:*

$$w_t = \int w^*(\Pi) f(\Pi | history) d\Pi \tag{6}$$

where $f()$ is the joint probability density function, a product of Dirichlet probability densities (the posterior of Π).

Proof: In appendix C by Cogley and Sargent (2008) in case of risk neutrality and two states this has already been proved and this proof is based on it. Invoking Theorem 5:

$$w_t = \mathbb{E}_t^s \sum_{j=1}^{\infty} \prod_{s=1}^{j} \hat{m}_{t+s-1,t+s}$$

and expressing the subjective expectation as a sum yields:

$$w_t = \sum_{j=1}^{\infty} \sum_{k=1}^{K_j} \hat{m}_{k,j} \mathbb{P}(g_t^j = k | history).$$

where g_t^j is a bijection, mapping from $\{1, 2 \dots N\}^j$ to $\{1, 2 \dots K_j\}$, ordering to each realization of states from period t until $t + j$ a natural number, $K_j = N^j$, and $\hat{m}_{k,j} = \prod_{s=1}^{j} \hat{m}_{t+s-1,t+s}$ given that $(g_t^j)^{-1}(k)$ realized.

The predictive probabilities can be expressed as:

$$\mathbb{P}(g_t^j = k | history) = \int \mathbb{P}(g_t^j = k | history, \Pi) f(\Pi | history) d\Pi.$$

Therefore:

$$w_t = \int \{ \sum_{j=1}^{\infty} \sum_{k=1}^{K_j} \hat{m}_{k,j} \mathbb{P}(g_t^j = k | history, \Pi) \} f(\Pi | history) d\Pi.$$

Notice the term in brackets is equivalent to the rational expectations solution for a known Π matrix due to the Markov property and the updating rule:

$$w_t = \int w^*(\Pi) f(\Pi | history) d\Pi.$$

$\square$

For EZW preferences, by transforming Equation 5 the following holds:

$$\mathbb{E}_t^s[\hat{m}_{t,t+1}(1 + w_{t+1})^{\frac{1-\gamma}{1-\theta}}] = w_t^{\frac{1-\gamma}{1-\theta}}. \tag{7}$$

Assumption 2. *: γ and θ are both greater than one or both smaller. In addition, $\theta \leq \gamma$.*

Theorem 7. *With Assumption 2 an upper bound for the price to dividend ratio exists for EZW preferences:*

$$w_t^{EZW} \leq w_t^{CRRA},$$

where the CRRA parameter equals γ

Proof: First, consider rational expectations. Using the inequality that $(1+x)^\alpha \leq 1 + x^\alpha$ for $\alpha \in (0, 1)$ and $0 \leq x$, the no-bubbles condition as in Theorem 5 and first order stochastic dominance:

$$\begin{aligned}
w_t^{EZW} &= \mathbb{E}_t^s[\hat{m}_{t,t+1}(1 + w_{t+1})^{\frac{1-\gamma}{1-\theta}}] \\
&\leq \mathbb{E}_t^s[\hat{m}_{t,t+1}(1 + w_{t+1}^{\frac{1-\gamma}{1-\theta}})] \\
&= \mathbb{E}_t^s[\hat{m}_{t,t+1} + \hat{m}_{t,t+1}\mathbb{E}_{t+1}^s[\hat{m}_{t+1,t+2}(1 + w_{t+2})^{\frac{1-\gamma}{1-\theta}}]] \\
&\leq \mathbb{E}_t^s \sum_{j=1}^{\infty} \prod_{s=1}^{j} \hat{m}_{t+s-1,t+s} + \mathbb{E}_t^s \prod_{j=1}^{\infty} \hat{m}_{t+j-1,t+j}(1 + w_{t+j})^{\frac{1-\gamma}{1-\theta}} \\
&\leq \mathbb{E}_t^s \sum_{j=1}^{\infty} \prod_{s=1}^{j} \hat{m}_{t+s-1,t+s} + \mathbb{E}_t^s \prod_{j=1}^{\infty} \hat{m}_{t+j-1,t+j}(1 + w_{t+j}) \\
&= w_t^{CRRA}.
\end{aligned}$$

By invoking the last step in the proof of Theorem 6 it is also apparent, by first order stochastic dominance, that the inequality holds under Bayesian updating too.

□

Theorem 7 implies that the price dividend ratio for EZW preferences is bounded. Therefore if there is a unique fix-point to Equation 4 then it will be a solution only if it is smaller than the corresponding price to dividend ratio for CRRA. Intuitively it is also apparent from the proof that if θ does not differ a lot from γ and the variation in the stochastic discount factor is not substantial - because the variance of the stochastic process is small - then the solution is close to the CRRA solution. From the proof of Theorem 6 it can also be seen that there is no obvious way to derive a closed form for w_t^{EZW} under Bayesian updating, hence an additional assumption is required.

Assumption 3. *The representative consumer is unaware of the fact that she is reestimating the transition probabilities each period. That is, she thinks that her current estimate of the transition matrix will be true from now on forever.*

Assumption 3 is equivalent to the assumption that the probability density function in Equation 6 is no longer a density but a single mass point. It can be thought that this approximates what a true Bayesian would do relatively well if the pdf is "centered" around the current estimate of the transition matrix. Assumption 3 is necessary to decrease the computational burden as otherwise at each time point a numerical integration of the product Dirichlet densities would be required. Moreover, for general EZW preferences each evaluation of the price to dividend ratio requires a non-linear root-finding which is sensitive to starting values. It is shown in section 5 that for CRRA preferences, Assumption 3 seems not to be a bad approximation compared to what a true Bayesian would do.

3.5 Implementation in MATLAB

The formula for the price of a one period ahead risk-free discount coupon given that the current state is characterized by λ_i, using Matlab notation can be written as:

$$p^f = \beta^{\frac{1-\gamma}{1-\theta}} \cdot \Pi(i,:) \cdot Diag(\lambda.^{-\gamma}) \cdot (\mathbf{1}_N + w).^{\frac{\theta-\gamma}{1-\theta}} / (w(i))^{\frac{\theta-\gamma}{1-\theta}}$$

where $Diag()$ denotes the linear transformation $\mathbb{R}^N \rightarrow \mathbb{R}^{N\times N}$, such that the diagonal elements of the matrix in the range are the elements of the vector from the domain and all off-diagonal elements are 0. $\mathbf{1}_N$ denotes the N dimensional vector where all components are equal to 1. $w \in \mathbb{R}^N$ is the price to dividend ratio. The operation "." denotes element by element operation. Therefore the return on the one year risk free asset is given by $r^f = \frac{1}{p^f} - 1$. Return on equity is given by:

$$r^e = \Pi(i,:)Diag(\lambda) \cdot (w + \mathbf{1}_N) \cdot \frac{1}{w(i)} - 1,$$

and w solves the nonlinear system:

$$\mathbf{0}_N = \beta^{\frac{1-\gamma}{1-\theta}} \cdot \Pi \cdot Diag(\lambda.^{1-\gamma}) \cdot (\mathbf{1}_N + w).^{\frac{1-\gamma}{1-\theta}} - w^{\frac{1-\gamma}{1-\theta}}. \quad (8)$$

The solution is obtained by applying the Trust-region Dogleg algorithm developed by Powell (1968). The timing is the following:

For period t:

- Realization of the dividend growth rate
- Based on the realization, the consumer updates the histories and her estimate $\hat{\Pi}$
- Expectations and consumption decisions are formed and the assets are priced until the no-trade equilibrium is reached (rootfinding)

4 Calibration and Data Description

In this section the parameters of the model are determined. The transition probabilities and growth rates of the Markov process are estimated while the utility parameters β, γ and θ are calibrated to match certain features of financial markets in the US economy.

4.1 Consumption Data Description

The dataset used for the estimation of the consumption process is assembled by Barro and Ursúa (2008). It includes yearly data of GDP and consumption levels for 42 countries[4] including the biggest economies, starting from 1790 to 2009. Viewed as an unbalanced panel, it has an additional property that if GDP is missing then consumption is missing too. The dataset contains the major disasters in consumption in the 19^{th} and the 20^{th} century which were missing from previous data leading to a survivorship-bias in consumption growth. As data is likely to be missing during crises, omitting missing observations from estimating a consumption process will lead to a bias both about the mean and about the variance. The long panel structure is essential to counter the argument that the US had not experienced disasters, which would render the disaster literature unfounded. First, the US did experience disastrous events prior to 1871 - the starting date of reliable financial data - for example the civil war with a yearly drop of 4% in consumption for five years, which most certainly affected the agents beliefs about the US economy. Second the US might have been just lucky avoiding disaster realizations - nevertheless agents could have considered the US similar to countries which were not so lucky. Argentina, which was similar to the US in many economic features in the beginning of the 20^{th} century has experienced various economic disasters. Of course, if one could observe the realizations of consumption growth for the US for thousands of years it would be possible to estimate the consumption generating process relatively well. As this is not possible, the assumption that the countries experienced different realizations of the same stochastic process increases the number of relevant observations sufficiently.

[4]The list of countries is provided in Appendix B

It is possible to fill in the missing observations for consumption by estimating the following model:

$$\Delta \log C_{it} = \Delta \log Y_{it}\beta + \xi_t + \eta_i + u_{it}$$

Table 1: Regressing consumption on GDP

Parameter	Point estimate	Standard Error	t-value
β	0.595284	0.013408	44.399
Summary	R^2	Adjusted R^2	Number of observations
Values	0.29647	0.28137	4929

Note: Only the coefficient for the GDP is shown.

Table 2: Consumption Summary Statistics

Statistics	Consumption Growth Rates
Number of observations	6079
Mean	0.0204
Median	0.0197
Standard deviation	0.0661
Min	-0.4272
Max	0.6276

Note: The consumption series obtained after forecasting whenever it was missing but GDP was not.

where C and Y is the level of consumption and income, respectively, ξ_t is a constant time effect, η_i is a constant individual —country specific — effect and u_{it} is an iid normal process uncorrelated with the exogenous variables. Where GDP is available, but consumption is not, this model is used to conditionally forecast it. Estimation is performed by the within estimator, results are shown in Table 1. In order not to decrease the relative standard deviation of the consumption growth rates, the estimated error terms are added randomly to the conditionally forecasted values of consumption. After dealing with missing observation the consumption growth rates with 6079 observations are obtained. In Table 2 the properties of this series are summarized. The obtained series are then used to estimate a first order, multi state Markov process.

4.2 Estimation of the Consumption Process

Estimating a Markov process is usually done in a Hidden Markov Model (HMM) framework. Assume that the DGP for consumption is:

$$\frac{C_{it}}{C_{i,t-1}} - 1 = \lambda_{S_t} + u_{it}$$

where λ_{S_t} is the growth rate[5] depending on the current state $S_t = \{1, 2, ...N\}$. S_t is unobservable and follows an ergodic, irreducible Markov process with transition matrix Π. u_{it} is assumed to be iid normal. It is important to note that such a model is justified by the observed non stationarity in the consumption data. The standard maximum likelihood approach developed by Hamilton (1989) produces statistically insignificant coefficients as the number of states increase[6]. Also, the unobserved state assumption is in conflict with the model setup where agents are assumed to be able to determine the state they are in. Therefore an alternative estimator is constructed.

Let $b_1, b_2, \ldots, b_{N-1}$ be an initial guess. Then a growth rate is assumed to be a realization of state $k \in \{2, ...N-1\}$ if

$$b_{k-1} < \frac{C_{it}}{C_{i,t-1}} - 1 \leq b_k$$

and if $\frac{C_{it}}{C_{i,t-1}} - 1 \leq b_1$ then a realization of state 1 occurred, else it is a realization of state N. The coefficients λ_{S_t} are calculated as the averages of the realizations in their respective states. The next step is to estimate by maximum likelihood the transition matrix on these realization of states as if they were truly observed. Then calculate the stationary distribution of this estimated Markov process and the moments of the stationary distribution. As there are $N-1$ free variables (b_1, $b_2, \ldots, b_{N-1}$), the procedure requires the first $N-1$ moments of the sample to be equal to the moments of the stationary distribution estimated - implemented as a simple root-finding problem where the dimensionality of the problem increases only linearly with the number of states, not exponentially as in the HMM frame-

[5]Note the slight change in the definition of growth which is used with Markov process as in Rietz (1988) - it only makes a difference for extreme rates

[6]The estimation is not included in the appendix due to length issues.

work. Therefore it is much faster (the maximum likelihood step is relatively fast for observable states), especially suited for estimating several state processes and it is exactly how agents would estimate a Markov process given that they observe the state. The estimated transition matrix, growth rates, stationary distribution and duration times for the 5 state process are reported in Table 3. Notice that the worst state is a 16 % drop in consumption per year, happening on average in 3 % of the observations - quite often, but it is in line with the findings of (Nakamura et al., 2013). Moreover, observe that from the worst state there is almost 10 % chance of a transition to the state with the highest growth rate and it is also probable that the disaster will continue with 16 % chance. There is an intuitive reason for letting $N = 5$. As the number of states increases it is easier to observe variation in the states in the 20^{th} century US history. However, there are $N(N-1)$ parameters to learn about and for a large N the estimating agent would have a difficult job to do. Results would not differ a lot if a process with $N = 4$ or $N = 6$ has been estimated.

Table 3: Estimated Markov chain with five states

Π_0 matrix	To state 1	To state 2	To state 3	To state 4	To state 5
From state 1	0.065	0.272	0.261	0.294	0.109
From state 2	0.017	0.298	0.39	0.253	0.043
From state 3	0.006	0.099	0.656	0.221	0.018
From state 4	0.02	0.155	0.468	0.316	0.042
From state 5	0.097	0.175	0.248	0.32	0.160
Growth rates	27 %	9.78 %	2.68 %	-3.26 %	-16.1 %
Duration times	1.07	1.42	2.9	1.46	1.19
Stationary distr.	0.015	0.148	0.549	0.254	0.034

Note: The transition probability matrix, growth rates, duration times (in years) and the stationary distribution is shown, respectively.

4.3 Financial data description

The financial data is obtained from the webpage[8] of Aswath Damodaran and is summarized in Table 4. First, it is important to notice that the standard

[8] *http : /pages.stern.nyu.edu/adamodar/New_Home_Page/datafile/histretSP.html*

Table 4: Financial data summary in the US, 1928-2009

Variable	Mean	Standard deviation	Corr with consumption
Sharpe ratio (from 1928)[7]	0.56	-	-
Sharpe ratio (from 1955)	0.245	0.0222	43 %
Equity premium	7.53 %	0.21	10.9 %
Risk-free rate	1.46 %	0.043	-44.7 %
Price to dividend ratio	31.07	16.15	14.4 %

Note: The risk-free rate is measured as the yearly return on the 3-month Treasury Bill, the risky return and the price to dividend ratio are the market return and the inverse dividend to price ratio of S&P500. The Sharpe ratio is calculated from 1955 to allow for a conditioning period for the variance and the mean of the equity premium.

deviation of the equity premium is high, rendering the estimate for the average equity premium unreliable. Therefore the Sharpe ratio should be considered as the reliable variable for measuring the price of risk as it is discussed by Hansen and Jagannathan (1991). Second, the correlation with consumption growth with any of the variables is not high and is usually the opposite of what is reported in the literature - price to dividend ratio,the risk premium and the Sharpe Ratio are usually reported to be strongly countercyclical, see Jagannathan et al. (2000). It can also be that consumption growth is not the correct measure of cycles. Therefore evaluating the correlation with the consumption growth aspect of the model should be taken with a grain of salt.

4.4 Baseline Calibration of utility parameters and the priors

Table 5: Baseline calibration

β	θ	γ	fixed-window	Simulation period	T_0	Prior
0.985	2	3	30 years	1928-2009	30 years	Π_0

Note: Apart from the utility parameters, in case of the fixed-window learning, the length of the memory is specified. The agent must observe a number of years prior to the starting of the simulation (T_0), but the prior distribution is nonetheless assumed to be the DGP one.

The baseline calibration of the utility parameters are presented in Table 5. This calibration targets are presented in Table 6. With a higher discount factor β, the risk free rate would be lower. There are two reasons not to increase it.

Table 6: Baseline calibration targets

Variable	Steady State Value
risk-free rate	4.62%
risk-free rate - normalized consumption	1.52%
Price to dividend ratio	32.67

Note: The growth rate of consumption equals the average consumption growth observed in the US data from 1928, a value of 1.51%.

First, the price to dividend ratio w is increasing with β. Moreover, as positive shocks occur, the estimated growth rate can become infinite - the likelihood of that event depends on β and the highest possible growth rate λ_N. θ is picked such that the price to dividend ratio in steady state approximately equals the one observed. The risk aversion parameter γ is high, but not exorbitant. It is set higher than θ to demonstrate the effects of Epstein-Zin utility, while keeping in mind Assumption 2. The higher the value of γ the higher the equity premium and thus the Sharpe Ratio will be. Note however that the Sharpe Ratio generated by a learning economy is not directly comparable to the Sharpe ratio in the data. Cogley and Sargent (2008) based on, Hansen and Jagannathan (1991) construct the correct measure of the price of risk. First, the per period Sharpe ratio depends on the state while the one observed in the data is "smoothed" out — therefore a transformation that eliminates this type of non-stationarity is required. More importantly, the probabilities should be corrected with the objective probabilities because the data is assembled that way. The formula is given in Matlab notation:

$$PR_{RE}(t) = \frac{\pi_0 \cdot (\Pi_0. \cdot (\Pi./\Pi_0).^2) \cdot m.^2 - (\pi_0 \cdot ((\Pi_0. \cdot (\Pi./\Pi_0)) \cdot m).^2)^{1/2}}{\pi_0 \cdot (\Pi_0. \cdot (\Pi./(\Pi_0))) \cdot m}$$

where π denotes the stationary distribution of the Markov process characterized by Π transition matrix, the index 0 denotes the objective, true parameters and m is the per-period stochastic discount factor. The prior distribution in the baseline calibration equals to the objective transition matrix Π_0 as if coming from 30 years of observations. More formally, the prior $\forall i$ is distributed as:

$$\Pi(i,:) \sim Dir(4, 30\Pi_0(i,:)).$$

The fixed-window parameter $fixedw$ refers to the scenario when the representative agent at time t forgets the transition from $t - fixedw - 1$ to $t - fixedw$. If $fixedw$ is smaller than the number of conditioning periods than the initial prior persists - creating a similar model to that of Cecchetti et al. (2000).

4.5 Alternative Calibration

Table 7: Alternative calibration

β	θ	γ	fixed-window	Simulation period	T_0	Prior
0.985	0.25	0.25	8 if $T_0 = 10$,10 otherwise	1928-1999	10, 30 50 or 70 years	Distorted

Note: In this specification T_0 varies across simulations and the distorted prior distribution and the fixed window with it.

Table 8: Alternative calibration targets

Variable	Steady State Value
risk-free rate	1.91%
risk-free rate - normalized consumption	1.52%
Price to dividend ratio	260.1

Note: The growth rate of consumption equals the average consumption growth observed in the US data from 1928, a value of 1.51%.

In order to compare results with Cogley and Sargent (2008)[9] , their calibration of the utility function and stochastic process is also used - see Table 7 and Table 9 and the steady state targets in Table 8. The reason for not using this calibration only, is that the 2 state specification does not generate enough "volatility" - that is, if one wants to evaluate the model's performance on the "true" history of the US. In addition, their model performs well only with an exogenous distortion of the prior, which is hard to justify. Moreover, in steady state, apart from generating an implausibly high price to dividend ratio - after all, the baseline

[9]more precisely, with the working paper version of it, as in the published one the agent is assumed to be risk neutral

Table 9: Markov chain with two states

Π_0 matrix	To state 1	To state 2
From state 1	0.978	0.022
From state 2	0.485	0.515
Growth rates	2.251 %	-6.785 %
Duration time	45.45	2.06
Stationary distr.	0.9566	0.0434

Note: The transition matrix, growth rates, duration time (in years) and stationary distribution of the 2 state process reproduced from Cecchetti et al. (2000) on the US data only with HMM, without standard errors.

calibration generates too high risk free rate - this undermines numerical precision as it is a key variable determining all the others in case of EZW preferences. The prior distributions are constructed in such a way that they are statistically indistinguishable from the true process but are pessimistic. Indistinguishable in terms of the Bayesian factor, as discussed by Kass and Raftery (1995):

$$B = \frac{L(history|\Pi_0)}{L(history|\Pi_{worst-case})}.$$

Table 10: Distorted priors

	$\Pi_{T_0=10}$		$\Pi_{T_0=30}$		$\Pi_{T_0=50}$		$\Pi_{T_0=70}$	
	To 1	To 2	To 1	To 2	To 1	To 2	To 1	To 2
From 1	0,775	0.225	0.886	0.114	0.914	0.086	0.926	0.074
From 2	0.068	0.932	0.142	0.858	0.183	0.817	0.21	0.79

Note: The transition matrix for the distorted priors — depending on T_0.

If $2 \log B < 10$ then the decision maker cannot differentiate significantly between the worst case and the "true" DGP. By distorting the priors to the maximum amount — which depends on the length of the observed history T_0 — the transition matrices are obtained (see Table 10) and then with these transition matrices the prior distributions are constructed as:

$$\Pi(i, i) \sim beta(T_0 \Pi_{T_0}(i, :)),$$

where $\forall\, T_0 \in \{10, 30, 50, 70\}$ and $i = 1, 2$. Although it is true, that they are "statistically" indistinguishable, it is important to note that these priors correspond to extreme disasters - if one calculates how these priors could arise from the data, then, for example in the $T_0 = 10$ case, 8 realizations of disasters are needed - no countries ever experienced such an implausible event. Therefore distorting the priors in some sense is equivalent to a rare disaster event, ultimately making the alternative and the baseline calibration comparable, because although the baseline calibration assumes the correct prior, it can generate rare disaster realizations along the simulations, a starting point for the alternative calibration.

5 Results

5.1 The effects of disasters

Table 11: No-disaster summary

Variable	Mean	Std deviation	Cyclicality
Sharpe ratio	0.0140	0.0003	-8.62 %
Equity premium	0.25 %	0.0333	-0.2 %
Risk-free rate	5.07 %	0.0146	100 %
Price to dividend ratio	30.1662	0.2577	-100 %

Note: The summary statistics of the simulated financial series when the stochastic process is taken as in Table 17. Cyclicality is measured as correlation with consumption.

Table 12: Disaster summary

Variable	Mean	Std deviation	Cyclicality
Sharpe ratio	0.0442	0.0003	-1.61 %
Equity premium	0.68 %	0.0576	-2.94 %
Risk-free rate	4.38 %	0.0109	79.15 %
Price to dividend ratio	32.2678	0.1838	-80.3 %

Note: The summary statistics of the simulated financial series when the stochastic process is taken as in Table 3 .

First the difference between the disaster and the standard, rational expectations solution with CRRA ($\theta = \gamma = 2$) utility is demonstrated. The parameters are as in the baseline calibration (see Table 5). The summary statistics of 2000 simulations are presented in Table 11 and in Table 12. For the no disaster simulation a two state Markov consumption process is estimated, where both of the growth rates are positive[10]. The no disaster model performs extremely poorly. The equity premium and therefore the price of risk is a small fraction of what is observed in the data in Table 4 - a known failure of the standard model. The mean of the risk free rate and the price to dividend ratio was calibrated. The standard deviation of the price to dividend ratio is low - with iid shocks it would be zero (recall Equation 8) - and as the risky return and the risk premium is depending on the price to dividend ratio, their variation is also low. The correlation of the risk-free rate with consumption is the opposite of what is observed

[10]see Table 17 in Appendix C

- this weakness of the Lucas tree model comes from the lack of production. If there was a storage technology - capital - and consumption would come from production, then the interest rate would determine consumption, not the other way around. That is, an increase in the interest rate would increase the growth rate of consumption which, if the substitution effect dominates, would lead to a decrease in current consumption.

With disasters, that is, when the five state process in Table 3 is used[11], the model performs better in all variables. The key mechanism of how disasters generate higher risk premium is that a possible consumption drop increases the (expected) marginal utility and so the stochastic discount factor in Equation 1. This implies that the demand for the risky asset would go down unless it's expected return goes up. Intuitively, the agents demand compensation for the event that the dividends are poor and they are left without anything to consume - as they are equipped with risk averse utility, these events are considered way worse than a comparable increase in the dividends - a bonanza - is considered to be good. Correlation with consumption became less substantial as the probability of reaching the disaster state is non-monotonic - in the best state there is very high probability that it will end badly for the consumer (see Table 3). The disaster calibration clearly improved the performance of the model and this was done only by allowing for more states - the estimation was performed on the same dataset and the first moment of the stationary distribution is also the same in both Markov specifications due to the estimation method. However, even the model with the disasters generates statistics that are not much closer to the reality.

5.2 Epstein-Zin preferences and Bayesian or fixed-window learning

The alternative calibration summarized in Table 7, is used to demonstrate the effects of introducing Epstein-Zin preferences and learning - it is chosen because then the results become directly comparable with Cogley and Sargent (2008). First consider the rational expectation solution summarized in Table 13, where

[11] It does not make much difference if lets say the 3 state process is used as it does have a bad state - it just makes the point clearer to stick to the baseline calibration

Table 13: Rational Expectation summary

	CRRA			Epstein-Zin		
	Mean	Std. dev.	Cyclicality	Mean	Std. dev.	Cyclicality
Sharpe ratio	0.008	0.0014	-2.2 %	0.0084	0.0013	-3.4 %
Equity premium	0.01 %	0.0281	0.37 %	0.05 %	0.0281	-2.4 %
Risk-free rate	1.98 %	0.0026	100 %	1.94 %	0.0029	100 %
P-D ratio	764.72	30.78	37.7 %	770.69	33.1867	35.8 %

Note: The summary statistics of the simulated financial series when the stochastic process is taken as in Table 9 and rational expectations is assumed.

Table 14: $T_0 = 30$ summary

	CRRA			Epstein-Zin		
	Mean	Std. dev.	Cyclicality	Mean	Std. dev.	Cyclicality
Sharpe ratio	0.2940	0.1255	-27.6 %	0.3414	0.1351	-28.6 %
Equity premium	1.37 %	0.0549	-44 %	1.66 %	0.0565	-45.1 %
Risk-free rate	1.82 %	0.0042	98.95 %	1.67 %	0.0047	98.1 %
P-D ratio	99.41	32.5159	33.5 %	88.19	27.8974	34.2 %
	CRRA,fixed-window			EZW,fixed-window		
	Mean	Std. dev.	Cyclicality	Mean	Std. dev.	Cyclicality
Sharpe ratio	0.4138	0.1253	-41.3 %	0.4283	0.1337	-42.9 %
Equity premium	0.9 %	0.0337	-26.3 %	1.03 %	0.0343	-27.7 %
Risk-free rate	1.81 %	0.0030	94.36 %	1.76 %	0.0034	92.74 %
P-D ratio	158.38	47.546	41.3 %	151.22	45.6902	41.1 %

Note: The summary statistics of the simulated financial series when the stochastic process is taken as in Table 9 and Bayesian updating (first row) or fixed-window learning (second row) is assumed with priors as in Table 10 for $T_0 = 30$.

the average statistics of 1000 simulations are shown. Introducing Epstein-Zin improves the performance of the model, but it is still not getting much closer to the data. While increasing risk aversion parameter to $\gamma = 0.75$ increases the risk premium, because θ remains the same, the risk-free return and the price to dividend ratio also remains the same. The intuition is that while a higher price of risk is demanded, the risk free rate and how much consumption an asset should provide is ultimately pinned down by the consumer's willingness to trade consumption between periods - the intertemporal elasticity of substitution. This argument is only true if the perceived difference between the risk free asset and the risky asset is not substantial in terms of possible consumption loss, because risk aversion does enter Equation 8.

To examine the effects of Bayesian learning consider the $T_0 = 30$ case in Table 14 where the average statistics of a 1000 simulations for 70 periods are

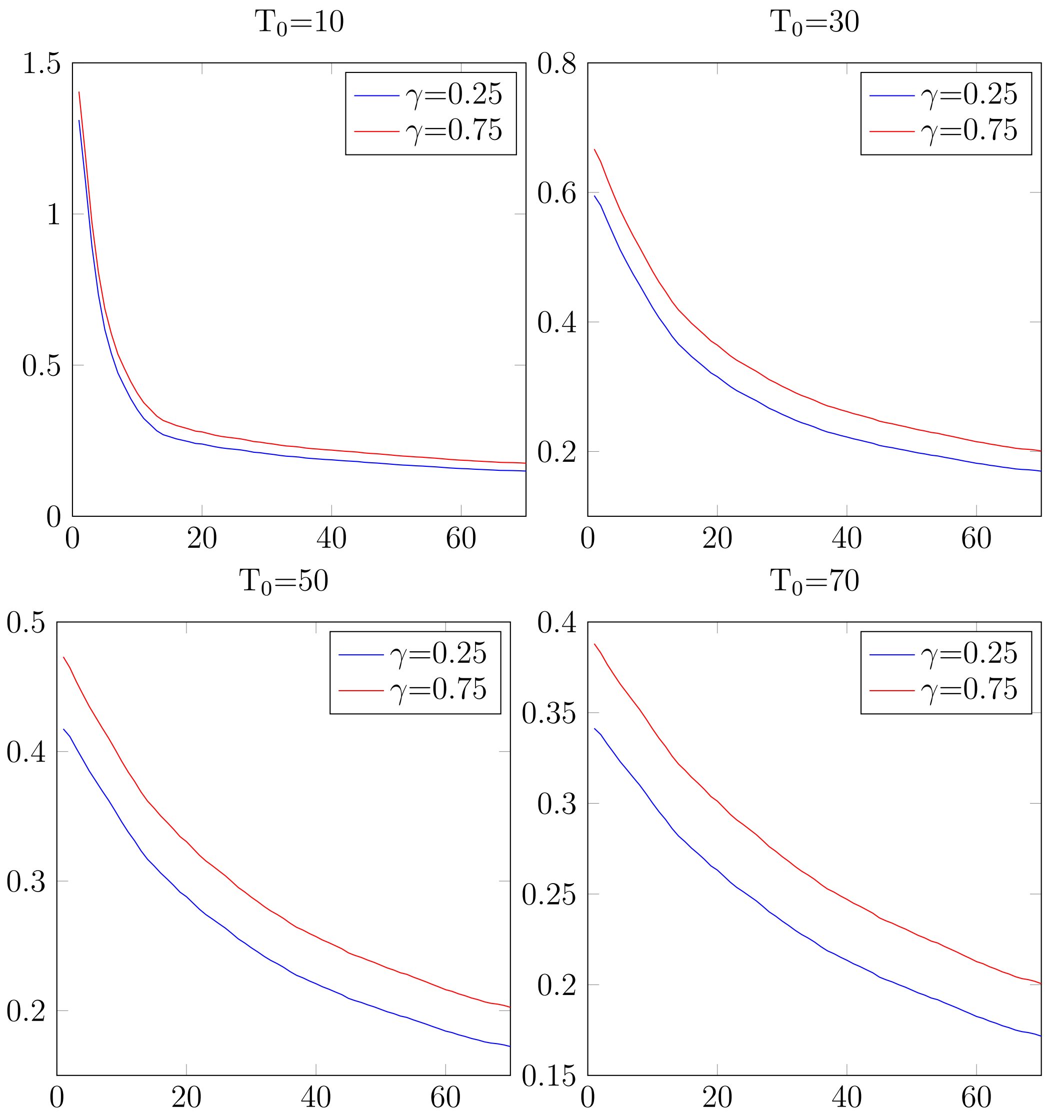

Figure 1: Unconditional Sharpe Ratio with $\theta = 0.25$

shown with the corresponding prior transition matrix in Table 10[12]. Bayesian learning is optimal, after a time the learning economy would converge to a rational expectations economy, but with this prior there is a sizeable difference between the two economies in the beginning of the simulation period. Overall, the pessimistic prior and learning generate a huge improvement in the model. The mean Sharpe Ratio is now comparable to what is actually observed. As the agent overestimates the probability and the duration of the bad state, they are demanding very high expected return from the risky asset. The key point here is that the variance of the equity premium is not increasing as much as the mean of the equity premium

[12] All other cases are summarized in Table 18, Table 19 and in Table 20 in Appendix C.

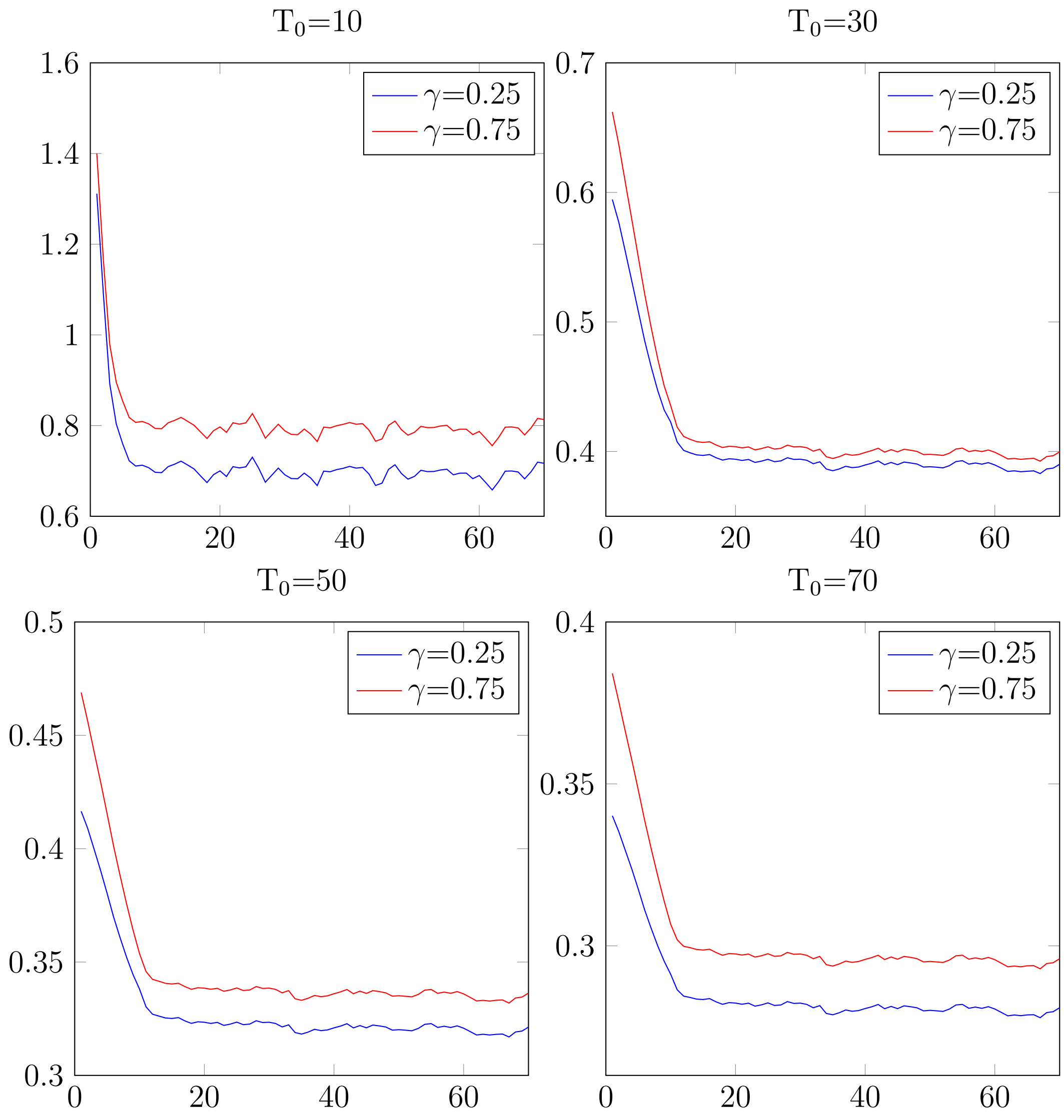

Figure 2: Unconditional Sharpe Ratio with fixed-window and $\theta = 0.25$

as both economies (the learning and the rational expectations one) are ultimately driven by the same shocks and therefore the increase in variance is due only to the persistence of a pessimistic bias about the economy. The price to dividend ratio is lower in the learning economy as the agents dislike the large perceived riskiness of the risky asset. Also, as the perceived risk is much larger for the consumer in the learning economy, the difference in the price to dividend ratio between the CRRA and the Epstein-Zin utility is greater than in the rational expectations economy -after all, risk aversion is only interesting if there is risk involved.

Fixed-window learning - when transitions that happened a certain time ago are forgotten - affects the model mechanics in two ways. Forgetfulness is not

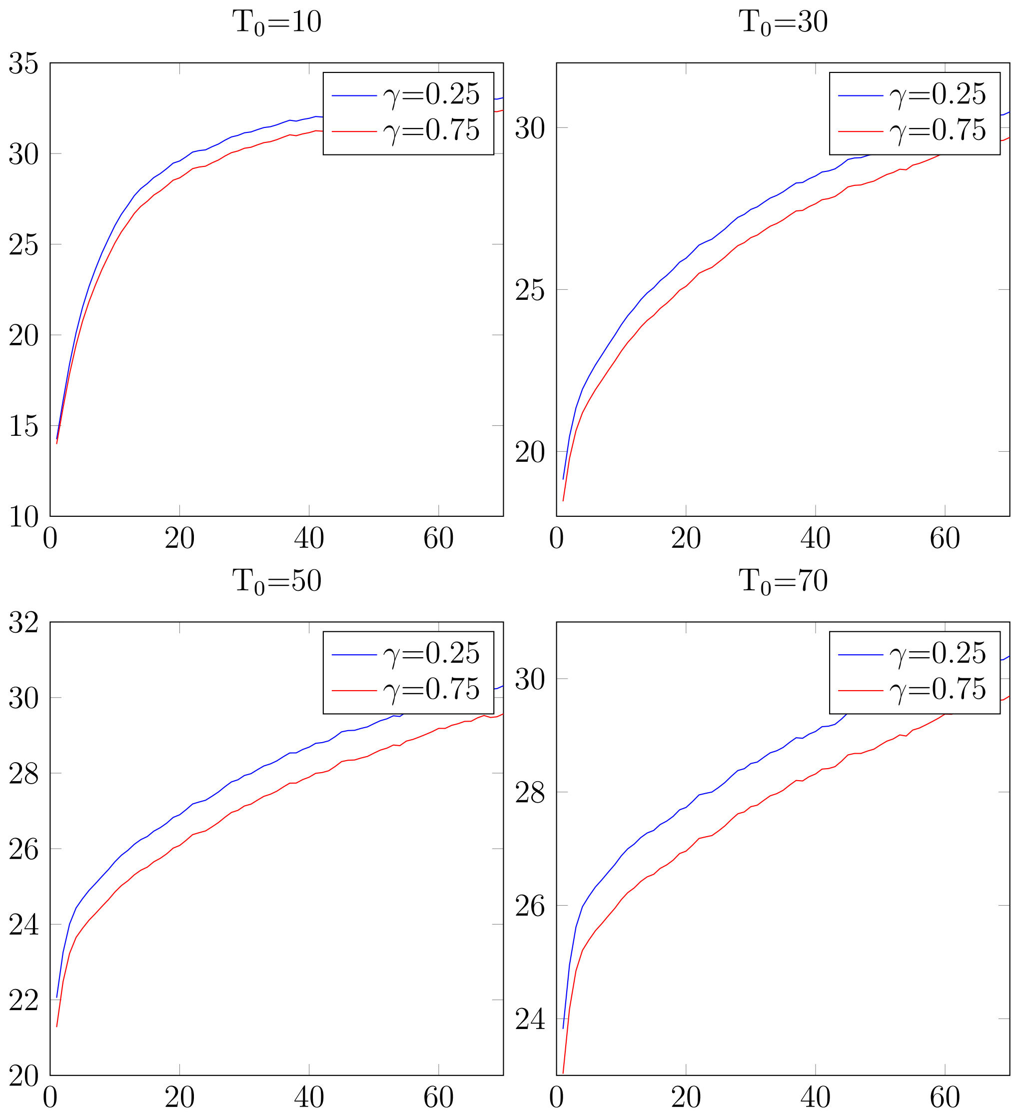

Figure 3: Price to dividend ratio with $\theta = 0.25$

complete - some initial realizations of the bad state are never forgotten, but the extreme pessimism that was present in the beginning of the simulation declines to an extent. Also, the pessimism is lost faster than in the Bayesian learning because not only the consumer obtains "good" signs as new realizations occur but in addition forgets the bad realizations. Therefore on the one hand, the systematic bias of the economy is not that substantial - hence the decrease in the relative deviation of the variables[13]. With complete forgetfulness, the equity premium on average would quickly decline to the rational expectation value. On the other hand though, as the initial prior persists to an extent even in the second half of

[13]That is, the standard deviation divided by the mean.

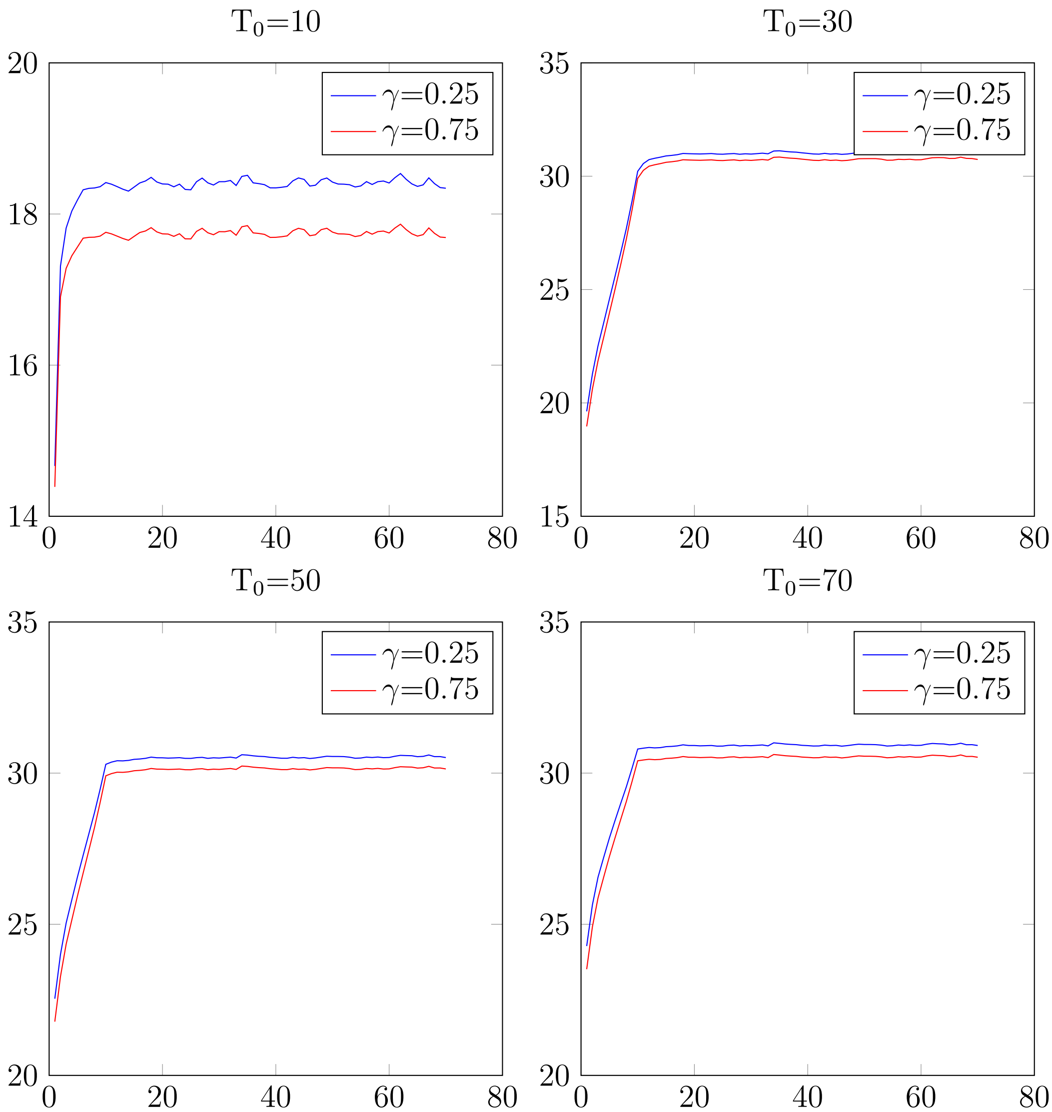

Figure 4: Price to dividend ratio with fixed-window and $\theta = 0.25$

the simulation period as opposed to the Bayesian learning, the equity premium on average decreases less than it's variance, increasing the Sharpe Ratio greatly. As the agents demand a lower premium, the price to dividend ratio increases while the difference between the CRRA and EZW utility decreases.

To directly compare results with Cogley and Sargent (2008) the figures 1-6 show the average of 1000 simulations for each point in time for the different priors. First it is important to see in Figure 1 that Assumption 3 is not restrictive at least in the case of CRRA utility as the graphs look exactly the same as in Cogley and Sargent (2008). As discussed before, increasing the risk aversion parameter γ while holding θ constant, increases the price of risk (the Sharpe Ratio) as

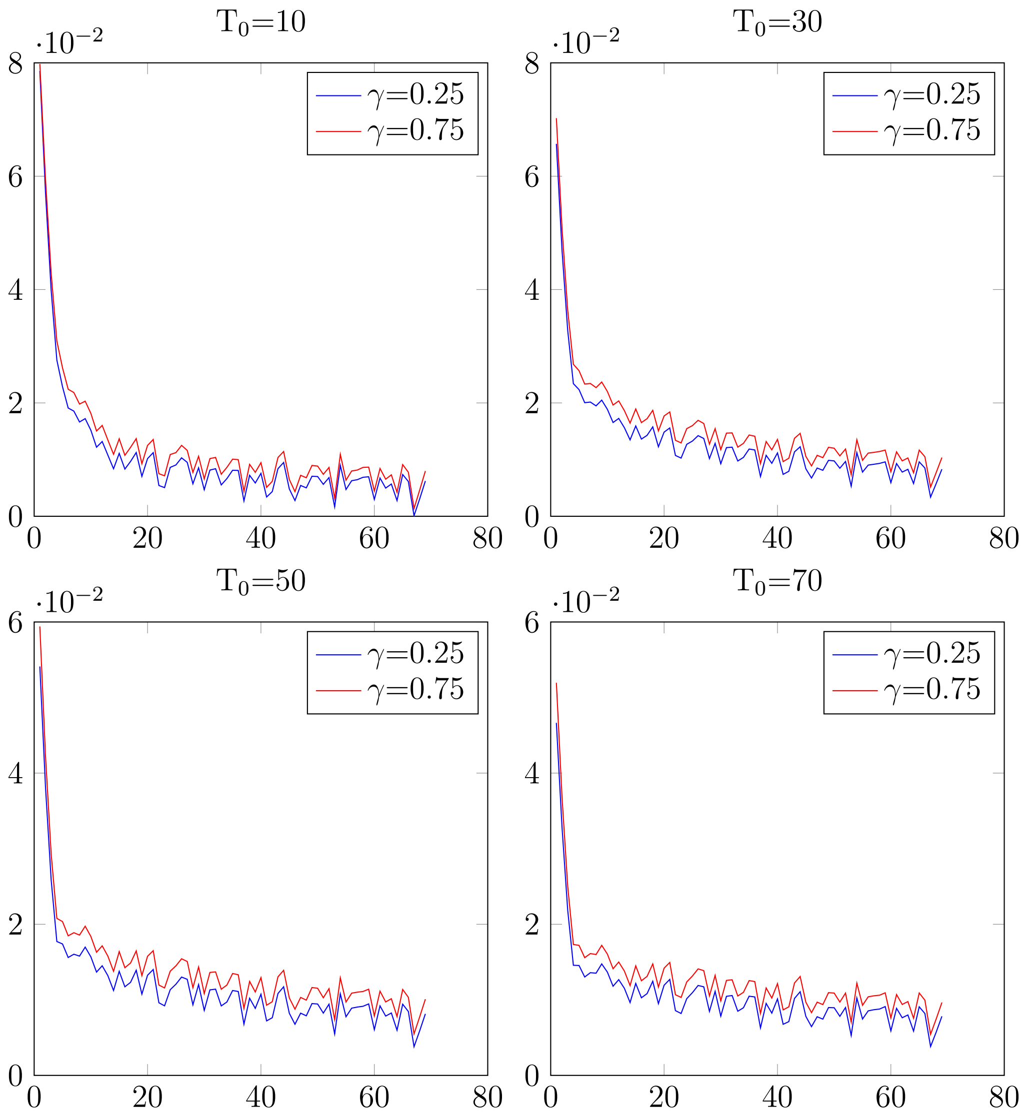

Figure 5: Realized equity risk premium with $\theta = 0.25$

more premium is demanded by the agents for a unit of additional risk. The price to dividend ratio decreases as shown on Figure 3 - the consumer demands more dividend for the same price because of increased risk aversion. Higher T_0 - the initial number of periods conditioning on - decreases the maximum possible distortion in the prior[14] making the agent less and less pessimistic, so that the initial price of risk decreases. It also affects the shift of the variables with higher γ - increasing T_0 decreases the shift in the Sharpe Ratio and in the price to dividend

[14] see Table 10

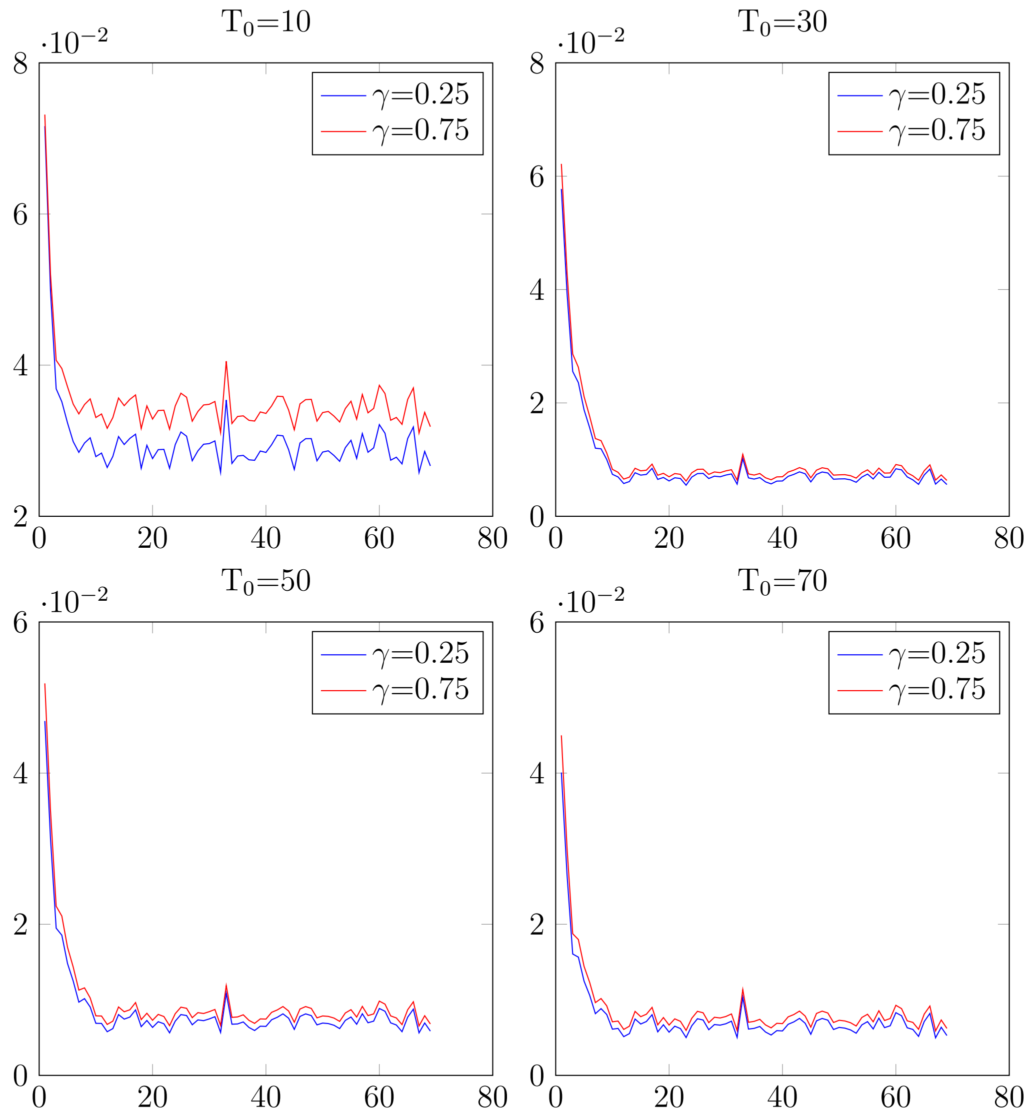

Figure 6: Realized equity risk premium with fixed-window and $\theta = 0.25$

ratio as the perceived risk decreases[15]. The realized (ex post) risk premium is positive on average, making ex post arbitrage possible[16] as can be seen in Figure 5

The effects of introducing fixed-window learning can be seen in Figure 2, Figure 4, and Figure 6. Note that the size of the fixed-window is depending on T_0. As mentioned in section 2, regime changes of fixed-window with Bayesian learning could be an explanation to the stationary, non-stationary switches of the variables, so that the speed of convergence to rational expectations can vary. That is, when there is a disaster one tends to remember it forever (Bayesian

[15]Note the different scaling of the axis though
[16]as discussed in Timmermann (1993)

learning), while when business is as usual then not importance is assigned to an individual observation so it might as well be forgotten (fixed window).

5.3 Testing the preferred calibration

Table 15: Historical simulation summary

Variable	Mean	Std deviation	Cyclicality	Corr with data
Sharpe ratio	0.2638	0.0293	7.2 %	-14.3%
Equity premium	0.3 %	0.0232	-9.5 %	52.1%
Risk-free rate	4.99 %	0.0127	81.6 %	-31.8%
Price to dividend ratio	31.5	3.0434	-22.4 %	-18.6%

Note: The summary statistics of the financial series when the stochastic process is taken as in Table 3, the calibration as in Table 5 and the historical realizations are fed to the model. The last column is the correlation between the observed and the simulated.

After a clear understanding of each component in the model separately, the baseline calibration is implemented in order to evaluate the overall performance of learning and disaster models. If the learning economy experiences the same shocks as the US economy did, starting from 1928, it can be seen on Figure 7 and in Table 15 that overall, the learning economy performs poorly - there is not nearly enough variation in the variables to explain the financial properties of the US. Nevertheless, it performs better than any previous model lacking either Epstein-Zin utility, fixed-window or disaster estimation. The main success of the model is that the mean and the standard deviation of the Sharpe ratio is now comparable to the one observed - even though the prior transition matrix was the data generating one, not exogenously distorted as before.

There are two main reasons for the failure in reproducing the other observed financial statistics. The lack of variation in consumption data might not reflect the true variance in the marginal utility - it might be just a problem that consumption is badly measured or is a bad proxy of marginal utility as discussed by Campbell (1993). Connected to this, it is also possible that the return on wealth is not the same as the return on the market portfolio, altering the risk premium that is needed to be "matched". Second, as discussed above, production[17] would

[17] with a non AK production technology, as there is equivalence between an AK and a Lucas tree economy

be really essential to change the correlation structure of the financial variables (importantly, the risk free rate) with consumption.

6 Conclusion

In this paper, a Lucas (1978) tree model with learning, disasters and Epstein-Zin utility has been applied in order to match the high observed price of risk. The representative consumer learns about the exogenous consumption growth process either by Bayesian or fixed-window updating. The method of updating ultimately determines whether the equilibrium converges to the rational expectations equilibrium or not. The exogenous process is assumed to be a Markov chain with at least one disaster state, in which consumption declines significantly inducing the agent to demand a high return from the risky return. Epstein-Zin utility allows a departure from the CRRA case so that the parameter controlling risk aversion is no longer need to also control the intertemporal elasticity of substitution. These extensions of the standard framework usually appear separately in the literature as discussed in section 2. The behavior of a model equipped with learning and Epstein-Zin utility is analyzed in section 3 while a calibration of the utility parameters and the estimation of the Markov process on a disaster dataset is presented in section 4.

The results in section 5 indicate that merging all these components greatly improve the performance of a Lucas tree model. Epstein-Zin utility enables one to match the steady state value of the price to dividend ratio which also improves the numerical precision during the simulations. In addition, compared to the CRRA case, it is possible to analyze an increase in risk aversion only, keeping the elasticity of intertemporal substitution constant. Disasters justify the deviation from the rational expectations by shattering beliefs while learning preserves this distance from the objective truth — overall accounting for the large perceived risk.

Although the observed averages of the financial variables are met when the historical realizations of consumption growth are fed into the model, the volatilities and the correlation with consumption are not. A future research direction could be to solve these problems by finding a better proxy than consumption for marginal utility and by introducing production.

References

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

Barro, R. J. and J. F. Ursúa (2008): "Consumption Disasters in the Twentieth Century," *American Economic Review*, 98, 58–63.

——— (2012): "Rare Macroeconomic Disasters," *Annual Review of Economics*, 4, 83–109.

Barsky, R. B. and J. B. De Long (1993): "Why Does the Stock Market Fluctuate?" *The Quarterly Journal of Economics*, 108, 291–311.

Barsky, R. B., M. S. Kimball, F. T. Juster, and M. D. Shapiro (1997): "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Survey," NBER Working Papers 5213, National Bureau of Economic Research, Inc.

Blanchard, O. J. (1993): "Movements in the Equity Premium," *Brookings Papers on Economic Activity*, 24, 75–138.

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.

Campbell, J. Y. (1993): "Intertemporal Asset Pricing without Consumption Data," *American Economic Review*, 83, 487–512.

Cecchetti, S. G., P. sang Lam, and N. C. Mark (2000): "Asset Pricing with Distorted Beliefs: Are Equity Returns Too Good to Be True?" *American Economic Review*, 90, 787–805.

Cogley, T. and T. J. Sargent (2008): "The Market Price of Risk and the Equity Premium: A Legacy of the Great Depression?" *Journal of Monetary Economics*, 55, 454–478.

Cogley, T., T. J. Sargent, and V. Tsyrennikov (2012): “Market Prices of Risk with Diverse Beliefs, Learning, and Catastrophes,” *American Economic Review*, 102, 141–46.

Epstein, L. G. and S. E. Zin (1989): “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework,” *Econometrica*, 57, 937–69.

——— (1991): “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis,” *Journal of Political Economy*, 99, 263–86.

Fama, E. F. and K. R. French (2002): “The Equity Premium,” *Journal of Finance*, 57, 637–659.

Gabaix, X. (2012): “Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance,” *The Quarterly Journal of Economics*, 127, 645–700.

Gourio, F. (2012): “Disaster Risk and Business Cycles,” *American Economic Review*, 102, 2734–66.

Hamilton, J. (1989): “A new approach to the economic analysis of nonstationary time series and the business cycle.” *Econometrica*, 57(2), 357–384.

Hansen, L. P. and R. Jagannathan (1991): “Implications of Security Market Data for Models of Dynamic Economies,” *Journal of Political Economy*, 99, 225–62.

Jagannathan, R., E. R. McGrattan, and A. Scherbina. (2000): “The declining U.S. equity premium,” *Quarterly Review*, 3–19.

Kass, R. E. and A. E. Raftery (1995): “Bayes factors,” *Journal of the american statistical association*, 90, 773–795.

Keynes, J. (1926): *The end of laissez-faire*, Reprints in History Series, L. & Virginia Woolf.

Kocherlakota, N. R. (1996): "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature*, 34, 42–71.

Lucas, Robert E, J. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1429–45.

Mehra, R. and E. C. Prescott (1985): "The equity premium: A puzzle," *Journal of Monetary Economics*, 15, 145–161.

——— (2003): "The equity premium in retrospect," in *Handbook of the Economics of Finance*, ed. by G. Constantinides, M. Harris, and R. M. Stulz, Elsevier, vol. 1 of *Handbook of the Economics of Finance*, chap. 14, 889–938.

Nakamura, E., J. Steinsson, R. Barro, and J. F. Ursúa (2013): "Crises and Recoveries in an Empirical Model of Consumption Disasters," *American Economic Journal: Macroeconomics*, 5, 35–74.

Powell, M. (1968): *A FORTRAN subroutine for solving systems of nonlinear algebraic equations.*

Rietz, T. A. (1988): "The equity risk premium: a solution," *Journal of Monetary Economics*, 22, 117–131.

Timmermann, A. G. (1993): "How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices," *The Quarterly Journal of Economics*, 108, 1135–45.

Uhlig, H. (2010): "Easy EZ in DSGE," *Society for Economic Dynamics,2010 Meeting Papers.*

Weil, P. (1989): "The Equity Premium Puzzle and the Riskfree Rate Puzzle," NBER Working Papers 2829, National Bureau of Economic Research, Inc.

Weitzman, M. L. (2007): "Subjective Expectations and Asset-Return Puzzles," *American Economic Review*, 97, 1102–1130.

A Appendix

Proof of Theorem 1:

Taking derivatives of the unconstrained objective function:

$$\frac{\partial U_t}{\partial C_t} = (1-\beta)C_t^{-\theta}(U_t)^{\theta} \tag{9}$$

$$\frac{\partial U_t}{\partial C_{t+1}} = (1-\beta)C_{t+1}^{-\theta}(U_{t+1})^{\theta-\gamma}\mathbb{E}_t^s(U_{t+1}^{1-\gamma})^{\frac{1}{1-\gamma}-1}\beta\mathbb{E}_t^s(U_{t+1}^{1-\gamma})^{\frac{-\theta}{1-\gamma}}U_t^{\theta} \tag{10}$$

Dividing Equation 10 with Equation 9 yields the result. □

Proof of Theorem 2:

Guess and verify

$$\begin{aligned}
W_t &= C_t + \mathbb{E}_t^s(m_{t,t+1}W_{t+1}) \\
&= C_t + \mathbb{E}_t^s\Big(\beta\big(\frac{C_{t+1}}{C_t}\big)^{-\theta}\big(\frac{U_{t+1}}{\mathbb{E}_t^s(U_{t+1}^{1-\gamma})^{\frac{1}{1-\gamma}}}\big)^{\theta-\gamma}\frac{U_{t+1}}{(1-\beta)C_{t+1}^{-\theta}(U_{t+1})^{\theta}}\Big) \\
&= C_t + \mathbb{E}_t^s\Big(\beta\big(\frac{C_{t+1}}{C_t}\big)^{-\theta}U_{t+1}^{1-\gamma}(\mathbb{E}_t^s(U_{t+1})^{1-\gamma})^{\frac{\gamma-\theta}{1-\gamma}}\frac{1}{(1-\beta)C_{t+1}^{-\theta}}\Big)
\end{aligned}$$

$$\begin{aligned}
\frac{U_t}{(1-\beta)C_t^{-\theta}(U_t)^{\theta}}(1-\beta)C_t^{-\theta}(\mathbb{E}_t^s(U_{t+1})^{1-\gamma})^{\frac{\theta-\gamma}{1-\gamma}} &\stackrel{?}{=} (1-\beta)C_t^{1-\theta}(\mathbb{E}_t^s(U_{t+1})^{1-\gamma})^{\frac{\theta-\gamma}{1-\gamma}} + \beta\mathbb{E}_t^s(U_{t+1}^{1-\gamma}) \\
U_t^{1-\theta}(\mathbb{E}_t^s(U_{t+1})^{1-\gamma})^{\frac{\theta-\gamma}{1-\gamma}} &\stackrel{?}{=} (1-\beta)C_t^{1-\theta}(\mathbb{E}_t^s(U_{t+1})^{1-\gamma})^{\frac{\theta-\gamma}{1-\gamma}} + \beta\mathbb{E}_t^s(U_{t+1}^{1-\gamma}) \\
U_t^{1-\theta} &\stackrel{?}{=} (1-\beta)C_t^{1-\theta} + (\mathbb{E}_t^s(U_{t+1})^{1-\gamma})^{\frac{1-\theta}{1-\gamma}}
\end{aligned}$$

which holds. This confirms the guess. □

Proof of Theorem 3:

$$\begin{aligned}
R_{w,t+1} &= \frac{U_{t+1}}{(1-\beta)C_{t+1}^{-\theta}(U_{t+1})^{\theta}} \frac{1}{\frac{U_t}{(1-\beta)C_t^{-\theta}U_t^{\theta}} - C_t} \\
&= \frac{U_{t+1}}{(1-\beta)C_{t+1}^{-\theta}(U_{t+1})^{\theta}} \frac{(1-\beta)C_t^{-\theta}U_t^{\theta}}{U_t - (1-\beta)C_t^{1-\theta}U_t^{\theta}} \\
&= \frac{U_{t+1}}{(1-\beta)C_{t+1}^{-\theta}(U_{t+1})^{\theta}} \frac{(1-\beta)C_t^{-\theta}U_t^{\theta}}{\beta U_t^{\theta}(\mathbb{E}_t^s(U_{t+1})^{1-\gamma})^{\frac{1-\theta}{1-\gamma}}} \\
&= \Big\{\beta\big(\frac{C_{t+1}}{C_t}\big)^{-\theta}\big(\frac{\mathbb{E}_t^s(U_{t+1}^{1-\gamma})^{\frac{1}{1-\gamma}}}{U_{t+1}}\big)^{1-\theta}\Big\}^{-1}
\end{aligned}$$

Substituting back to Equation 1 yields

$$\begin{aligned}
m_{t,t+1} &= \beta\big(\frac{C_{t+1}}{C_t}\big)^{-\theta}\Big(\frac{U_{t+1}}{\mathbb{E}_t^s(U_{t+1}^{1-\gamma})^{\frac{1}{1-\gamma}}}\Big)^{\theta-\gamma} \\
&= \beta\big(\frac{C_{t+1}}{C_t}\big)^{-\theta}\Big(\frac{R_{w,t+1}^{-1}}{\beta(\frac{C_{t+1}}{C_t})^{-\theta}}\Big)^{\frac{\theta-\gamma}{\theta-1}} \\
&= \beta^{\frac{1-\gamma}{1-\theta}}\big(\frac{C_{t+1}}{C_t}\big)^{-\theta\frac{1-\gamma}{1-\theta}} R_{w,t+1}^{\frac{\theta-\gamma}{1-\theta}}
\end{aligned}$$

□

Proof of Theorem 4: Using Equation 2 for the stock return gives:

$$\beta^{\frac{1-\gamma}{1-\theta}}\mathbb{E}_t^s\Big[\big(\frac{C_{t+1}}{C_t}\big)^{-\theta\frac{1-\gamma}{1-\theta}} R_{w,t+1}^{\frac{1-\gamma}{1-\theta}}\Big] = 1. \tag{11}$$

Asset returns for stocks can be written as:

$$R_{w,t+1} = \frac{P_{t+1}+C_{t+1}}{P_t} = \frac{C_{t+1}}{C_t}\frac{1+w_{t+1}}{w_t}. \tag{12}$$

Plugging Equation 12 back into Equation 11 yields:

$$\beta^{\frac{1-\gamma}{1-\theta}}\mathbb{E}_t^s\Big[\big(\frac{C_{t+1}}{C_t}\big)^{1-\gamma}\big(\frac{1+w_{t+1}}{w_t}\big)^{\frac{1-\gamma}{1-\theta}}\Big] = 1. \tag{13}$$

Suppose that from state i a transition to state j will occur in period $t+1$. Then the expectation operator in Equation 13 can be written using Equation 3 as:

$$1 = \beta^{\frac{1-\gamma}{1-\theta}} \cdot \Big(\sum_{j=1}^{n} \Pi_{ij} \lambda_j^{1-\gamma} \cdot \big(\frac{1 + w_j}{w_i} \big)^{\frac{1-\gamma}{1-\theta}} \Big).$$

Reordering this expression gives the result.

□

B Appendix

Argentina	Australia	Austria	Belgium	Brazil	Canada	Chile
China	Colombia	Denmark	Egypt	Finland	France	Germany
Greece	Iceland	India	Indonesia	Italy	Japan	Korea
Mexico	Malaysia	Netherlands	N.Z.	Norway	Peru	Philippines
Portugal	Russia	S. Africa	Singapore	Spain	Sri Lanka	Sweden
Switzerland	Taiwan	Turkey	U.K.	Uruguay	U.S.A.	Venezuela

Table 16: List of countries included for the estimation of the consumption process

C Appendix

Table 17: No-disaster Markov chain

Π_0 matrix	To state 1	To state 2
From state 1	0.317	0.683
From state 2	0.131	0.869
Growth rates	11.57 %	0.113 %
Duration time	1.46	7.62
Stationary distr.	0.161	0.839

Note: The transition matrix, growth rates, duration time (in years) and stationary distribution of the 2 state process estimated on the disaster dataset — with two states only, negative growth is not possible thus it is referred to as the no-disaster calibration.

Table 18: $T_0 = 10$ summary

	CRRA			Epstein-Zin		
	Mean	Std. dev.	Cyclicality	Mean	Std. dev.	Cyclicality
Sharpe ratio	0.2820	0.2576	-35.5 %	0.3262	0.2710	-36.3 %
Equity premium	1.12 %	0.0619	-46.5 %	1.37 %	0.0643	-46.9 %
Risk-free rate	1.85 %	0.0045	97.3 %	1.7 %	0.0050	95.7 %
P-D ratio	227.91	158.6209	30.3 %	199.08	136.8826	30.6 %
	CRRA,fixed-window			EZW,fixed-window		
	Mean	Std. dev.	Cyclicality	Mean	Std. dev.	Cyclicality
Sharpe ratio	0.7290	0.4254	-71.1 %	0.8332	0.4152	-72.2 %
Equity premium	3.11 %	0.0519	-36.8 %	3.67 %	0.0511	-34.7 %
Risk-free rate	1.68 %	0.0044	96.9 %	1.32 %	0.0037	97.9 %
P-D ratio	33.73	4.3164	86.4 %	31.50	3.8376	85.4 %

Note: The summary statistics of the simulated financial series when the stochastic process is taken as in Table 9 and Bayesian updating (first row) or fixed-window learning (second row) is assumed with priors as in Table 10 for $T_0 = 10$.

Table 19: $T_0 = 50$ summary

	CRRA			Epstein-Zin		
	Mean	Std. dev.	Cyclicality	Mean	Std. dev.	Cyclicality
Sharpe ratio	0.2642	0.0781	-26.5 %	0.3061	0.0841	-27.8 %
Equity premium	1.21 %	0.0508	-41.7 %	1.47 %	0.0525	-43.22 %
Risk-free rate	1.83 %	0.0040	99.5 %	1.7 %	0.0045	99.1 %
P-D ratio	104.29	25.5976	37.3 %	93.12	22.1637	38.25 %
	CRRA,fixed-window			EZW,fixed-window		
	Mean	Std. dev.	Cyclicality	Mean	Std. dev.	Cyclicality
Sharpe ratio	0.3335	0.0714	-41.3 %	0.3518	0.0765	-43.7 %
Equity premium	0.9 %	0.0343	-28.8 %	1 %	0.0351	-30.6 %
Risk-free rate	1.83 %	0.0031	98.3 %	1.8 %	0.0036	97.7 %
P-D ratio	144.67	30.8452	46.8 %	136.12	29.3167	46.6 %

Note: The summary statistics of the simulated financial series when the stochastic process is taken as in Table 9 and Bayesian updating (first row) or fixed-window learning (second row) is assumed with priors as in Table 10 for $T_0 = 50$.

Table 20: $T_0 = 70$ summary

	CRRA			Epstein-Zin		
	Mean	Std. dev.	Cyclicality	Mean	Std. dev.	Cyclicality
Sharpe ratio	0.2434	0.0562	-26.2 %	0.2811	0.0604	-27.7 %
Equity premium	1.08 %	0.0480	-39.8 %	1.32 %	0.0497	-41.6 %
Risk-free rate	1.84 %	0.0039	99.7 %	1.72 %	0.0045	99.5 %
P-D ratio	111.74	22.0567	40.7 %	100.14	19.2585	41.76 %
	CRRA,fixed-window			EZW,fixed-window		
	Mean	Std. dev.	Cyclicality	Mean	Std. dev.	Cyclicality
Sharpe ratio	0.2887	0.0489	-41.68 %	0.3065	0.0527	-44.33 %
Equity premium	0.8 %	0.0337	-28.1 %	0.9 %	0.0346	-30.1 %
Risk-free rate	1.85 %	0.0031	98.84 %	1.78 %	0.0036	98.5 %
P-D ratio	154.68	27.9229	48.8 %	144.94	26.5856	48.55 %

Note: The summary statistics of the simulated financial series when the stochastic process is taken as in Table 9 and Bayesian updating (first row) or fixed-window learning (second row) is assumed with priors as in Table 10 for $T_0 = 70$.

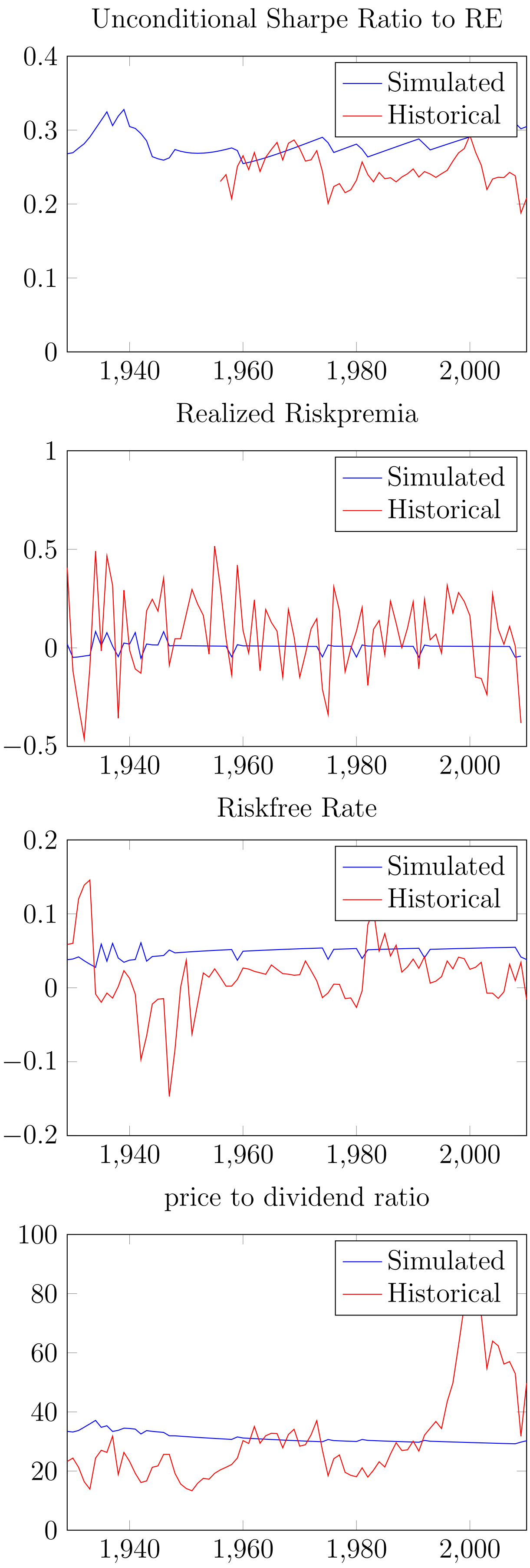

Figure 7: US baseline calibration with the true history