Likelihood Estimation of DSGE Models with Epstein-Zin Preferences*

Jules H. van Binsbergen

Duke University

Jesús Fernández-Villaverde

University of Pennsylvania, NBER, and CEPR

Ralph S.J. Koijen

New York University and Tilburg University

Juan F. Rubio-Ramírez

Duke University and Federal Reserve Bank of Atlanta

March 24, 2008

*Corresponding author: Juan F. Rubio-Ramírez, 213 Social Sciences, Duke University, Durham, NC 27708, USA. E-mail: Juan.Rubio-Ramirez@duke.edu. We thank Stephanie Schmitt-Grohé and Martín Uribe for comments. Beyond the usual disclaimer, we must note that any views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Finally, we also thank the NSF for financial support.

Abstract

This paper illustrates how to perform likelihood-based inference in dynamic stochastic general equilibrium (DSGE) models with Epstein-Zin preferences. This class of preferences has recently become a popular device to account for asset pricing observations and other phenomena that are challenging to address within the traditional state-separable utility framework. However, there has been little econometric work in the area, particularly from a likelihood perspective, because of the difficulty in computing an equilibrium solution to the model and in deriving the likelihood function. To fill this gap, we build a real business cycle model with Epstein-Zin preferences and long-run growth, solve it with perturbation techniques, and evaluate its likelihood with the particle filter. We estimate the model using U.S. macro and yield curve data. We discuss the ability of the model to explain the business cycle, asset prices, the comovements between these two, and the implications of our point estimates for the welfare cost of the business cycle.

Keywords: DSGE Models, Epstein-Zin Preferences, Likelihood Estimation.

JEL classification numbers: C32, C63, C68, G12

1. Introduction

This paper illustrates how to perform likelihood-based inference in dynamic stochastic general equilibrium (DSGE) models with Epstein-Zin preferences. Over the last years, this class of recursive utility functions (Kreps and Porteus, 1978, Epstein and Zin, 1989 and 1991, and Weil, 1990) has become popular.[1] The separation they allow between the intertemporal elasticity of substitution (IES) and risk aversion is intuitively appealing and provides an extra degree of freedom that can be put to good use. For instance, in the asset pricing literature, researchers have argued that Epstein-Zin preferences account for many patterns in the data, possibly in combination with other features like long-run risk. Bansal and Yaron (2004) is a prime representative of this line of work. From a policy perspective, Epstein-Zin preferences generate radically bigger welfare costs of the business cycle than the standard expected utility framework (Tallarini, 2000). Hence, they may change the trade-offs that policy-makers face, as in the example built by Levin, López-Salido, and Yun (2007). Finally, Epstein-Zin preferences can be reinterpreted, under certain conditions, as a particular case of robust control preferences (Hansen, Sargent, and Tallarini, 1999).

Unfortunately, much of the work with Epstein-Zin preferences has faced two limitations. First, except in a few papers like Backus, Routledge, and Zin (2007), Campaneli, Castro, and Clementi (2007), or Croce (2006), researchers have studied economies where consumption follows an exogenous process, sometimes estimated, often calibrated. This is a potentially important shortcoming, since production economies place tight restrictions on the comovements of endogenous variables that exogenous consumption models are not forced to satisfy. For example, in the Epstein-Zin setting, a first order approximation to the return of a risky asset is a weighted mean of the asset covariance with consumption growth and the asset covariance with the return on wealth. Since, in the data, the returns on wealth are more volatile than consumption, it has been argued that Epstein-Zin preferences may explain the equity premium. However, such reasoning suffers from the problem that consumption and wealth are linked through the budget constraint. For this mechanism to work in general equilibrium, the model has to generate a consumption that is smooth enough and wealth that is sufficiently volatile as to induce an equity premium of the right size. This is not an easy task. Furthermore, considering production economies with labor supply is quantitatively relevant. Uhlig (2007) has shown how, with Epstein-Zin preferences, leisure affects asset pricing in a

[1] Among dozens of papers, we can cite Backus, Routledge, and Zin (2007), Bansal and Yaron (2004), Campanale, Castro, and Clementi (2007), Campbell (1993) and (1996), Croce (2006), Dolmas (1998 and 2006), Gallmeyer *et al.* (2007), Gomes and Michealides (2005), Hansen, Heaton, and Li (2008), Kaltenbrunner and Lochstoer (2007), Lettau and Uhlig (2002), Piazzesi and Schneider (2006), Restoy and Weil (1998), Tallarini (2000), and Uhlig (2007).

significant way through the risk-adjusted expectation operator, even when it enters separately in the period utility function.

The second limitation of the literature is that there has been little econometric guidance in the selection of parameter values. Much of the intuition that we have about which parameter values are reasonable and about how to calibrate models efficiently come from years of experience with expected utility models. It is unclear how much of that accumulated learning can be translated into models with Epstein-Zin preferences (see, for instance, the cautionary arguments regarding identification in Kocherlakota, 1990b). As we will demonstrate with our analysis of the likelihood, the problem not only involves the obvious parameters controlling IES and risk aversion, but also the other parameters in the model. Learning about them from the data is important because some economists, like Cochrane (2008), have expressed skepticism about the empirical plausibility of having simultaneously high IES and high risk aversion, a condition that we need if we want Epstein-Zin preferences to have a quantitative impact.

These two limitations, i.e., few production economies and few econometric exercises, share to a large extent a common origin: the difficulty in working with Epstein-Zin preferences. Instead of the simple optimality conditions of expected utility models, recursive preferences imply necessary conditions that include derivatives of the value function that, in general, we cannot eliminate.[2] Therefore, we cannot apply standard linearization techniques. Instead, we need to resort to either simplifying the problem by working only with an exogenous flow for consumption, or employing solution methods like value function iteration (Croce, 2006) or projection methods (Campaneli, Castro, and Clementi, 2007), which are compositionally costly and suffer from the curse of dimensionality. The former solution precludes numerous inference exercises of interest, since it does not solve for the equilibrium dynamics of the model. The latter solution makes likelihood estimation exceedingly challenging because of the time spent in the solution of the model for each set of parameter values.

We overcome the difficulty in working with Epstein-Zin preferences through the use of perturbation techniques and the particle filter. We perturb the value function formulation of the household problem to obtain a high-order approximation to the solution of the model given some parameter values in a trivial amount of time. In companion work (Caldara, Fernández-Villaverde, Rubio-Ramírez, and Yao, 2008), we document that this solution is highly accurate and we compare it with alternative computational approaches.

[2] Epstein and Zin (1989) avoid this problem by showing that if we have access to the total wealth portfolio, we can derive a first order condition in terms of observables that can be estimated using a method of moments. However, in general we do not observe the total wealth portfolio because of the difficulties in measuring human capital, forcing the researcher to proxy the return on wealth.

Our perturbation shows analytically that the first order approximation to the policy functions of our model with Epstein-Zin preferences is equivalent to that of the model with standard utility. The second order approximation changes the constant of the solution that captures precautionary behavior. The new value of the constant moves the ergodic distribution of states, affecting, through this channel, allocations, prices, and welfare up to a first order. More concretely, by changing the mean of capital in the ergodic distribution, risk aversion influences the average level of the yield curve. The third order approximation changes the slope of the response of the solution to variations in the states of the model.

An important advantage of our solution technique is that we do not limit ourselves to the case with unitary IES, as Tallerini (2000) and others are forced to do.[3] There are three reasons why this flexibility is welcomed. First, because restricting the IES to be equal to one seems an unreasonably tight restriction that is hard to reconcile with other econometric estimates (Hall, 1988). Second, because a value of the IES equal to one implies that the consumption-wealth ratio is constant over time. Checking for this implication of the model is hard because wealth is not directly observable as it includes human wealth. However, different attempts at measurement, like Lettau and Ludvigson (2001), who control for variation in human capital by using observable labor income, or Lustig, van Nieuwerburgh, and Verdehaln (2007) reject the hypothesis that the ratio of consumption to wealth is constant. Third, because the debate between Campbell (1996) and Bansal and Yaron (2004) about the usefulness of the Epstein-Zin approach pertains to the right value of the IES. By directly estimating this parameter, we contribute to this conversation.

The second step in our procedure is to use the particle filter to find the likelihood function of the model (Fernández-Villaverde and Juan Rubio-Ramírez, 2005 and 2007). Evaluating the likelihood function of a dynamic equilibrium model is equivalent to keeping track of the conditional distribution of unobserved states of the model with respect to the data. Our perturbation solution is inherently non-linear (otherwise, the parameter controlling risk-aversion would not appear). These non-linearities make the conditional distribution of states intractable and prevent the use of more conventional methods, such as the Kalman filter. The particle filter is a sequential Monte Carlo method that replaces the conditional distribution of states by an empirical distribution of states drawn by simulation. The insight of the particle filter is that the simulation is performed sequentially as we get new observations, a procedure known as sequential importance resampling (SIR). The resampling step guarantees that the simulation delivers sufficient accuracy in a small amount of time.

[3] There is also another literature, based on Campbell (1993), that approximates the solution of the model around a value of the IES equal to one. Since our perturbation is with respect to the volatility of the productivity shock, we can deal with arbitrary values of the IES.

To illustrate our procedure, we work with a prototype real business cycle model. We take the stochastic neoclassical growth model and introduce Epstein-Zin preferences and long-run growth through a unit root process in the law of motion for technological progress. We solve the model using a third order perturbation, and we estimate it with U.S. macro and yield curve data. Having Epstein-Zin preferences allows the model to have a fighting chance at accounting for asset pricing observations and brings to the table the information that financial data may have about the aggregate economy. In a finding reminiscent of Campbell (1993) and (1996), we document how macro data help us learn about the IES while saying little about risk aversion. With respect to financial data, the level of the yield curve teaches us much about the IES, while the slope is informative about risk aversion. This result is also present in Campbell (1993). We explore the different modes of the likelihood, describe how the change when we vary the data employed in the estimation, and document how those modes explain part of the variation in point estimates in the literature regarding the IES. We complete our presentation by discussing the ability of the model to account for the business cycle, asset prices, the comovements between these two, and the implications of our point estimates for the welfare cost of the business cycle.

We purposely select as our application a prototype business cycle model for three reasons. First, because we develop several new techniques in this paper whose performance was unknown to us ex-ante. It seems prudent to apply them first to a simple model that we understand extremely well in the case where we have standard preferences. Second, because even in such a simple model, we have faced computational challenges that have put us close to the limit of what is currently feasible. As we get more familiar with the techniques involved and as hardware and software progresses, we are confident that we will be able to handle much more realistic models. Finally, because it is a necessary first step to map which observations an estimated prototype business cycle model can and cannot account for when we have Epstein-Zin preferences. After we know the strengths and weaknesses of the model, we can venture into incorporating richer mechanisms such as the ones that have been popular in the business cycle and asset pricing literatures (habit formation, long-run risks, real and nominal rigidities, etc.).

Of course, constraining ourselves to a prototype business cycle model exacts a steep tribute: many aspects of interest will be left out. Most notably, we will take inflation as an exogenous process about which we will have little to say. This is an important weakness of our paper and one about which we are totally up-front. For example, Piazzesi and Schneider (2006) suggest that the interaction between inflation risk and consumption is key to understanding the dynamics of the yield curve. We think, however, that the rewards of our exploratory exercise are sufficiently high as to justify our heroic simplifications.

The rest of the paper is organized as follows. In section 1, we present a simple DSGE model, rewrite it in a stationary recursive representation, and show how to price bonds with it. Section 2 explains our computational algorithm and writes the model in a state space form. Section 3 discusses the behavior of the model and illustrates how the likelihood function will identify the parameters. Section 4 builds the likelihood of the model. Section 5 describes the data and reports the estimation results. Section 6 concludes. An appendix provides further details about some aspects of the paper.

2. A Prototype DSGE Model

In this section, we present our prototype DSGE model, we rewrite it in a recursive, stationary form, and we derive expressions for the nominal and real price of uncontingent bonds.

2.1. Basic Set-Up

As we explained in the introduction, we work with a conventional stochastic neoclassical growth model where we incorporate only one peculiarity: the presence of Epstein-Zin preferences. There is a representative agent whose utility function over streams of consumption c_t and leisure l_t is:

$$U_t = \left[\left(c_t^{\upsilon} \left(1 - l_t\right)^{1-\upsilon} \right)^{\frac{1-\gamma}{\theta}} + \beta \left(\mathbb{E}_t U_{t+1}^{1-\gamma} \right)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}}$$

where $\gamma \geq 0$ is the parameter that controls risk aversion, $\psi \geq 0$ is the IES, and:

$$\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}.$$

The term $\left(\mathbb{E}_t U_{t+1}^{1-\gamma} \right)^{\frac{1}{1-\gamma}}$ is often called the risk-adjusted expectation operator. When $\gamma = \frac{1}{\psi}$, $\theta = 1$, the recursive preferences collapse to the standard case of state-separable, CRRA utility. The discount factor is β and our period is one quarter.

The budget constraint of the households is given by:

$$p_t c_t + p_t k_{t+1} + q_{t,t+1} b_t = p_t w_t l_t + p_t r_t k_t + b_t \tag{1}$$

where p_t is the price level at period t, k_t is capital, b_t is an uncontingent bond that pays one nominal unit in period t, $q_{t,t+1}$ is the nominal price at t of the uncontingent bond that pays at period $t+1$, w_t is the real wage, and r_t is the real rental price of capital. Weexplicitly include in the budget constraint only a one-period uncontingent bond because below we will derive its price explicitly. With that pricing, we derive the prices of all the other possible securities. We omit them from the budget constraint in the interest of clarity. In any case, their price in equilibrium will be such that the representative agent will hold a zero amount of them.

The technology is described by a neoclassical production function $y_t = k_t^{\zeta} (z_t l_t)^{1-\zeta}$, where output y_t is produced with capital, labor and technology z_t. Technology evolves as a random walk in logs with drift:

$$\log z_{t+1} = \lambda + \log z_t + \chi \sigma_\varepsilon \varepsilon_{zt+1} \text{ where } \varepsilon_{zt} \sim \mathcal{N}(0,1)$$

The drift induces long-run technological progress and the random walk introduces a unit root in the model. We pick this specification over trend stationarity convinced by Tallarini's (2000) demonstration that a different stationary representation for technological progress facilitates the work of a model similar to ours at matching the observed market price of risk.[4] Part of the reason, as emphasized by Rouwenhorst (1995), is that the presence of a unit root shifts, period by period, the long-run growth path of the economy, increasing the variance of future paths of the variables and hence the utility cost of risk. The parameter χ scales the standard deviation of the productivity shock. This parameter will facilitate the presentation of our perturbation solution method later on.

Capital depreciates at rate δ. This observation, together with the presence of competitive markets for inputs, implies that the resource constraint of the economy is:

$$c_t + k_{t+1} = k_t^{\zeta} (z_t l_t)^{1-\zeta} + (1-\delta) k_t$$

We will use the resource constraint to derive allocations, and we will return to the budget constraint in (1) to derive asset prices.

In our data, we will observe nominal bond prices. Hence, we need to take a stand on the evolution of the nominal price level p_t, how it evolves over time, and the expectations that the representative household form about it. Since at this stage we want to keep the model as stylized as possible, we take the price level, p_t, as an exogenous process that does not affect the allocations and real prices. Therefore, our economy presents a strikingly strong form of neutrality of nominal prices. But if we take inflation as exogenous, we must, at least, choose a reasonable process for it. Stock and Watson (2007) suggest that the following IMA(1,1)

[4] Similarly, Álvarez and Jerman (2005) argue that the observed market price of risk strongly points out to the presence of non-stationarities in the shocks that hit the economy.

process for inflation:

$$\log \frac{p_{t+1}}{p_t} - \log \frac{p_t}{p_{t-1}} = \omega_{t+1} - \tau\omega_t \text{ where } \omega_t \sim \mathcal{N}(0, \sigma_\omega) \tag{2}$$

is a good representation of the evolution of prices in the U.S. and that it is difficult to beat as a forecasting model. Following Stock and Watson, we take this IMA representation as the law of motion for prices in our economy. Also, households will have rational expectations: they know process (2) and they observe the sequence of shocks ω_t (i.e., the process is invertible). Under these assumptions, we show in the appendix that the expectation of the ratio of prices between p_t and p_{t+j}, which we will use later when we price bonds, is given by:

$$\mathbb{E}_t \frac{p_t}{p_{t+j}} = \left(\frac{p_{t-1}}{p_t}\right)^j \exp\left(j\tau\omega_t + \frac{1}{2}\sigma_\omega^2 \sum_{i=0}^{j-1} (i\tau - (i+1))^2\right)$$

The definition of a competitive equilibrium in this economy is standard and we omit it for the sake of brevity.

2.2. Rewriting the Problem in a Stationary Recursive Form

The problem of the household is indexed by the states k_t and z_t. Thus, we can rewrite its problem in the form of a value function:

$$V(k_t, z_t; \chi) = \max_{c_t, l_t} \left[\left(c_t^\upsilon (1 - l_t)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}} + \beta \left(\mathbb{E}_t V^{1-\gamma}(k_{t+1}, z_{t+1}; \chi)\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}}$$

subject to the budget constraint:

$$c_t + k_{t+1} = k_t^\zeta (z_t l_t)^{1-\zeta} + (1 - \delta) k_t$$

We index the value function by the parameter χ because it will help us later when we perturb the problem. For now, the reader can think about χ as just one more parameter that we chose to write explicitly in the value function.

A first task is to make this problem stationary. Define $x_t = z_{t-1}$ and $\widetilde{var}_t = \frac{var_t}{x_t}$ for a variable var_t. Thus, given the law of motion of productivity, we have:

$$\widetilde{z}_{t+1} = \frac{z_{t+1}}{z_t} = \frac{x_{t+2}}{x_{t+1}} = \exp(\lambda + \chi\sigma\varepsilon_{zt+1})$$

First, we transform the budget constraint:

$$\widetilde{c}_t x_t + \widetilde{k}_{t+1} x_{t+1} = \widetilde{k}_t^{\zeta} x_t^{\zeta} \left(z_t l_t\right)^{1-\zeta} + (1-\delta)\, x_t \widetilde{k}_t$$

Dividing both sides by x_t:

$$\widetilde{c}_t + \widetilde{k}_{t+1} \widetilde{z}_t = \widetilde{z}_t^{1-\zeta} \widetilde{k}_t^{\zeta} l_t^{1-\zeta} + (1-\delta)\, \widetilde{k}_t$$

or:

$$\widetilde{k}_{t+1} = \widetilde{z}_t^{-\zeta} \widetilde{k}_t^{\zeta} l_t^{1-\zeta} + (1-\delta)\, \widetilde{z}_t^{-1} \widetilde{k}_t - \widetilde{z}_t^{-1} \widetilde{c}_t \quad (3)$$

Also:

$$\widetilde{y}_t = \widetilde{k}_t^{\zeta} \left(\widetilde{z}_t l_t\right)^{1-\zeta}$$

Given the homotheticity of the utility function (see Epstein and Zin's derivations in their 1989 paper, and Dolmas, 1996), the value function is homogeneous of degree υ in k_t and z_t:

$$V\left(k_t, z_t; \chi\right) = V\left(\widetilde{k}_t x_t, \widetilde{z}_t x_t; \chi\right) = V\left(\widetilde{k}_t, \widetilde{z}_t; \chi\right) x_t^{\upsilon}$$

Then:

$$V\left(\widetilde{k}_t, \widetilde{z}_t; \chi\right) x_t^{\upsilon} = \max_{c_t, l_t} \left[x_t^{\frac{\upsilon(1-\gamma)}{\theta}} \left(\widetilde{c}_t^{\upsilon} \left(1-l_t\right)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}} + \beta x_{t+1}^{\frac{\upsilon(1-\gamma)}{\theta}} \left(\mathbb{E}_t V^{1-\gamma}\left(\widetilde{k}_{t+1}, \widetilde{z}_{t+1}; \chi\right)\right)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}}$$

or

$$V\left(\widetilde{k}_t, \widetilde{z}_t; \chi\right) = \max_{c_t, l_t} \left[\left(\widetilde{c}_t^{\upsilon} \left(1-l_t\right)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}} + \beta \widetilde{z}_t^{\frac{\upsilon(1-\gamma)}{\theta}} \left(\mathbb{E}_t V^{1-\gamma}\left(\widetilde{k}_{t+1}, \widetilde{z}_{t+1}; \chi\right)\right)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}} \quad (4)$$

Now, to save on notation below, define:

$$B_t = B\left(\widetilde{k}_t, \widetilde{z}_t; \chi\right) = \left(\widetilde{c}_t^{\upsilon} \left(1-l_t\right)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}} + \beta \widetilde{z}_t^{\frac{\upsilon(1-\gamma)}{\theta}} \left(\mathbb{E}_t V^{1-\gamma}\left(\widetilde{k}_{t+1}, \widetilde{z}_{t+1}; \chi\right)\right)^{\frac{1}{\theta}}$$

and we have:

$$V\left(\widetilde{k}_t, \widetilde{z}_t; \chi\right) = \max_{c_t, l_t} B\left(\widetilde{k}_t, \widetilde{z}_t; \chi\right)^{\frac{\theta}{1-\gamma}}$$

Given our assumptions, $B\left(\widetilde{k}_t, \widetilde{z}_t; \chi\right) > 0$.

Finally, note that the model has a steady state in the transformed variables.

2.3. Pricing Bonds

To find the price of a bond in our model, we compute the derivative of the value function with respect to consumption today:

$$\begin{aligned}
\frac{\partial V_t}{\partial \widetilde{c}_t} &= \frac{\theta}{1-\gamma} B_t^{\frac{\theta}{1-\gamma}-1} \frac{1-\gamma}{\theta} \upsilon \frac{\left(\widetilde{c}_t^{\upsilon} (1-l_t)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}}}{\widetilde{c}_t} \\
&= \frac{\theta}{1-\gamma} V_t^{\frac{1-\gamma}{\theta}\left(\frac{\theta+\gamma-1}{1-\gamma}\right)} \frac{1-\gamma}{\theta} \upsilon \frac{\left(\widetilde{c}_t^{\upsilon} (1-l_t)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}}}{\widetilde{c}_t} \\
&= V_t^{1-\frac{1-\gamma}{\theta}} \upsilon \frac{\left(\widetilde{c}_t^{\upsilon} (1-l_t)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}}}{\widetilde{c}_t}
\end{aligned}$$

and the derivative of the value function with respect to consumption tomorrow:

$$\begin{aligned}
\frac{\partial V_t}{\partial \widetilde{c}_{t+1}} &= \frac{\theta}{1-\gamma} B_t^{\frac{\theta}{1-\gamma}-1} \beta \frac{1-\gamma}{\theta} \widetilde{z}_t^{\frac{\upsilon(1-\gamma)}{\theta}} \left(\mathbb{E}_t V_{t+1}^{1-\gamma}\right)^{\frac{1}{\theta}-1} V_{t+1}^{-\gamma} \frac{\partial V_{t+1}}{\partial \widetilde{c}_{t+1}} \\
&= \beta V_t^{1-\frac{1-\gamma}{\theta}} \widetilde{z}_t^{\frac{\upsilon(1-\gamma)}{\theta}} \left(\mathbb{E}_t V_{t+1}^{1-\gamma}\right)^{\frac{1}{\theta}-1} V_{t+1}^{-\gamma} V_{t+1}^{1-\frac{1-\gamma}{\theta}} \upsilon \frac{\left(\widetilde{c}_{t+1}^{\upsilon} (1-l_{t+1})^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}}}{\widetilde{c}_{t+1}}
\end{aligned}$$

where in the last step we have used the result regarding the derivative of the value function with respect to current consumption forwarded by one period.

Now, we define the variable $\widetilde{m}_{t+1}$:

$$\begin{aligned}
\widetilde{m}_{t+1} &= \widetilde{z}_t^{-1} \frac{\partial V_t / \partial \widetilde{c}_{t+1}}{\partial V_t / \partial \widetilde{c}_t} \\
&= \widetilde{z}_t^{-1} \frac{\beta V_t^{1-\frac{1-\gamma}{\theta}} \left(\mathbb{E}_t V_{t+1}^{1-\gamma}\right)^{\frac{1}{\theta}-1} V_{t+1}^{-\gamma} V_{t+1}^{1-\frac{1-\gamma}{\theta}} \upsilon \frac{\left(\widetilde{c}_{t+1}^{\upsilon}(1-l_{t+1})^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}}}{\widetilde{c}_{t+1}}}{V_t^{1-\frac{1-\gamma}{\theta}} \upsilon \frac{\left(\widetilde{c}_t^{\upsilon}(1-l_t)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}}}{\widetilde{c}_t}} \\
&= \beta \widetilde{z}_t^{\frac{\upsilon(1-\gamma)}{\theta}-1} \left(\frac{\widetilde{c}_{t+1}^{\upsilon} (1-l_{t+1})^{1-\upsilon}}{\widetilde{c}_t^{\upsilon} (1-l_t)^{1-\upsilon}}\right)^{\frac{1-\gamma}{\theta}} \frac{\widetilde{c}_t}{\widetilde{c}_{t+1}} \left(\frac{V_{t+1}^{1-\gamma}}{\mathbb{E}_t V_{t+1}^{1-\gamma}}\right)^{1-\frac{1}{\theta}}
\end{aligned}$$

To see that $\widetilde{m}_{t+1}$ is the pricing kernel in our economy, we rewrite the budget constraint of the household:

$$p_t c_t + p_t k_{t+1} + q_{t,t+1} b_t = p_t w_t l_t + p_t r_t k_t + b_t$$

as:

$$\widetilde{c}_t x_t + \widetilde{k}_{t+1} x_t + \frac{1}{p_t} q_{t,t+1} b_t = \widetilde{w}_t x_t l_t + r_t \widetilde{k}_t x_t + \frac{1}{p_t} b_t$$

Then, if we take first order conditions with respect to the uncontingent bond, we get:

$$\frac{1}{x_t p_t} q_{t,t+1} \frac{\partial V_t}{\partial \widetilde{c}_t} = \mathbb{E}_t \left\{ \frac{\partial V_t}{\partial \widetilde{c}_{t+1}} \frac{1}{x_{t+1} p_{t+1}} \right\}$$

Using the fact that $\widetilde{z}_t = \frac{x_{t+1}}{x_t}$

$$q_{t,t+1} = \mathbb{E}_t \left\{ \widetilde{z}_t^{-1} \frac{\partial V_t / \partial \widetilde{c}_{t+1}}{\partial V_t / \partial \widetilde{c}_t} \frac{p_t}{p_{t+1}} \right\}$$

By iterating in the price of a bond one period ahead and using the law of iterated expectations, the price of a $j-$period bond is given by:

$$\begin{aligned} q_{t,t+j} &= \mathbb{E}_t \left\{ \prod_{i=1}^{j} \widetilde{m}_{t+j} \frac{p_t}{p_{t+j}} \right\} \\ &= \mathbb{E}_t \left(\prod_{i=1}^{j} \widetilde{m}_{t+j} \right) \mathbb{E}_t \left(\frac{p_t}{p_{t+j}} \right) \end{aligned}$$

where the last line exploits the fact that the evolution of the price level is independent of the rest of the economy.

Before, we argued that, given our assumption about the law of motion for prices in this economy, the expectation of the ratio of prices evolves as:

$$\mathbb{E}_t \frac{p_t}{p_{t+j}} = \left(\frac{p_{t-1}}{p_t} \right)^j \exp \left(j\tau\omega_t + \frac{1}{2}\sigma_\omega^2 \sum_{i=0}^{j-1} (i\tau - (i+1))^2 \right)$$

Then:

$$q_{t,t+j} = \left(\frac{p_{t-1}}{p_t} \right)^j \exp \left(j\tau\omega_t + \frac{1}{2}\sigma_\omega^2 \sum_{i=0}^{j-1} (i\tau - (i+1))^2 \right) \mathbb{E}_t \left(\prod_{i=1}^{j} \widetilde{m}_{t+j} \right)$$

Also, note that the conditional expectation is a function of the states of the economy:

$$\mathbb{E}_t \left(\prod_{i=1}^{j} \widetilde{m}_{t+j} \right) = f^j \left(\widetilde{k}_t, \widetilde{z}_t; \chi \right)$$

Our perturbation method will approximate this function f^j using a high-order expansion.

We close by remembering that the real price of a bond is also given by f^j:

$$qr_{t,t+j} = \frac{q_{t,t+j}}{\mathbb{E}_t \left(\frac{p_t}{p_{t+j}} \right)} = \mathbb{E}_t \left(\prod_{i=1}^{j} \widetilde{m}_{t+j} \right) = f^j \left(\widetilde{k}_t, \widetilde{z}_t; \chi \right)$$

3. Solution of the Model

This section explains how to use a perturbation approach to solve the model presented in the previous section by working with the value function formulation of the problem in equation (4). The steps that we take are the same as in the standard perturbation. The reader familiar with that approach should have an easy time following our argument. We begin by finding a set of equations that includes the value function and the first order conditions derived from the value function. Then, we solve those equations in the case where there is no uncertainty. We use this solution to build a first order approximation to the value function. Second, we take derivatives of the previous equations and solve for the case with no uncertainty. With this solution, we can build a second order approximation to the value function and a first order approximation to the policy function. The procedure can be iterated by taking higher derivatives and solving for the deterministic case to get arbitrarily high approximations to the value function and policy function.

Perturbation methods have become a popular strategy for solving DSGE models (Judd and Guu, 1992 and 1987; see also Judd, 1998, and references cited therein, for a textbook introduction to the topic). Usually, the perturbation is undertaken on the set of equilibrium conditions defined by the first order conditions of the agents, resource constraints, and exogenous laws of motion (for a lucid exposition, see Schmitt-Grohé and Uribe, 2004). However, when we deal with Epstein-Zin preferences, it is more transparent to perform the perturbation starting from the value function of the problem. Moreover, if we want to perform welfare computations, the object of interest is the value function itself. Hence, obtaining an approximation to the value function is a natural goal.

We are not the first, of course, to explore the perturbation of value function problems. Judd (1998) already presents the idea of perturbing the value function instead of the equilibrium conditions. Unfortunately, he does not elaborate much on the topic. Schmitt-Grohé and Uribe (2005) employ a perturbation approach to find a second order approximation to the value function to be able to rank different policies in terms of welfare. We just emphasize the generality of the approach and discuss some of its theoretical and numerical advantages.

Our approach is also linked with Benigno and Woodford (2006) and Hansen and Sargent (1995). Benigno and Woodford present a new linear-quadratic approximation to solve optimal policy problems that avoids some problems of the traditional linear-quadratic approximation when the constraints of the problem are non-linear.[5] Thanks to this alternative approximation, Benigno and Woodford find the correct local welfare ranking of different policies. Our method, as theirs, can deal with non-linear constraints and obtain the correct local approxi-

[5] See also Levine, Pearlman, and Pierse (2006) for a similar treatment of the problem.

mation to welfare and policies. One advantage of our method is that it is easily generalizable to higher order approximations without adding further complications. Hansen and Sargent (1995) modify the linear-quadratic regulator problem to include an adjustment for risk. In that way, they can handle some versions of recursive utilities like the ones that motivate our investigation. Hansen and Sargent's method, however, requires imposing a tight functional form for future utility. Moreover, as implemented in Tallarini (2000), it requires solving a fixed point problem to recenter the approximation to control for precautionary behavior that can be costly in terms of time. Our method does not suffer from those limitations.

3.1. Perturbing the Value Function

To illustrate our procedure, we limit our exposition to derive the second order approximation to the value function, which generates a first order approximation to the policy function. Higher order terms are derived in similar ways, but the algebra becomes too cumbersome to be developed explicitly in the paper (in our application, the symbolic algebra is undertaken by the computer employing `Mathematica` which automatically generates Fortran 95 code that we can evaluate numerically). Hopefully, our steps will be enough to understand the main thrust of the procedure and to let the reader obtain higher order approximations by herself.

Under differentiability conditions, the second order Taylor approximation of the value function around the deterministic steady state $\left(\widetilde{k}_{ss}, \widetilde{z}_{ss}; 0\right)$ is:

$$\begin{aligned} V\left(\widetilde{k}_t, \widetilde{z}_t; \chi\right) \simeq\ & V_{ss} + V_{1,ss}\left(\widetilde{k}_t - \widetilde{k}_{ss}\right) + V_{2,ss}\left(\widetilde{z}_t - \widetilde{z}_{ss}\right) - +V_{3,ss}\chi \\ & + \frac{1}{2}V_{11,ss}\left(\widetilde{k}_t - \widetilde{k}_{ss}\right)^2 + \frac{1}{2}V_{12,ss}\left(\widetilde{k}_t - \widetilde{k}_{ss}\right)\widetilde{z}_t + \frac{1}{2}V_{13,ss}\left(\widetilde{k}_t - \widetilde{k}_{ss}\right)\chi \\ & + \frac{1}{2}V_{21,ss}z_t\left(\widetilde{k}_t - \widetilde{k}_{ss}\right) + \frac{1}{2}V_{22,ss}\left(\widetilde{z}_t - \widetilde{z}_{ss}\right)^2 + \frac{1}{2}V_{23,ss}\left(\widetilde{z}_t - \widetilde{z}_{ss}\right)\chi \\ & + \frac{1}{2}V_{31,ss}\chi\left(\widetilde{k}_t - \widetilde{k}_{ss}\right) + \frac{1}{2}V_{32,ss}\chi\left(\widetilde{z}_t - \widetilde{z}_{ss}\right) + \frac{1}{2}V_{33,ss}\chi^2 \end{aligned}$$

where:

$$\begin{aligned} V_{ss} &= V\left(\widetilde{k}_{ss}, \widetilde{z}_{ss}; 0\right) \\ V_{i,ss} &= V_i\left(\widetilde{k}_{ss}, \widetilde{z}_{ss}; 0\right) \text{ for } i = \{1,2,3\} \\ V_{ij,ss} &= V_{ij}\left(\widetilde{k}_{ss}, \widetilde{z}_{ss}; 0\right) \text{ for } i,j = \{1,2,3\} \end{aligned}$$

By certainty equivalence, we will have that:

$$V_{3,ss} = V_{13,ss} = V_{23,ss} = 0$$

Moreover, taking advantage of the equality of cross-derivatives, and setting $\chi = 1$, which is just a normalization of the perturbation parameter implied by the standard deviation of the shock σ, the approximation we search has the simpler form:

$$\begin{aligned} V\left(\widetilde{k}_t, \widetilde{z}_t; 1\right) \simeq\; & V_{ss} + V_{1,ss}\left(\widetilde{k}_t - \widetilde{k}_{ss}\right) + V_{2,ss}\left(\widetilde{z}_t - \widetilde{z}_{ss}\right) \\ & + \frac{1}{2}V_{11,ss}\left(\widetilde{k}_t - \widetilde{k}_{ss}\right)^2 + \frac{1}{2}V_{22,ss}\left(\widetilde{z}_t - \widetilde{z}_{ss}\right)^2 + V_{12,ss}\left(k_t - k_{ss}\right)\left(\widetilde{z}_t - \widetilde{z}_{ss}\right) \\ & + \frac{1}{2}V_{33,ss} \end{aligned}$$

Note that $V_{33,ss} \neq 0$, a difference from the standard linear-quadratic approximation to the utility functions, where the constants are dropped. However, this result is intuitive, since the value function is, in general, translated by uncertainty. Since the term is a constant, it affects welfare (that will be an important quantity to measure in many situations) but not the policy functions of the social planner.[6] Furthermore, the term has a straightforward interpretation. At the deterministic steady state, we have:

$$V\left(\widetilde{k}_{ss}, \widetilde{z}_{ss}; \chi\right) \simeq V_{ss} + \frac{1}{2}V_{33,ss}$$

Hence:

$$\frac{1}{2}V_{33,ss}$$

is a measure of the welfare cost of the business cycle, i.e., of how much utility changes when the variance of the productivity shocks is σ^2 instead of zero.[7] This term is an approximation to the welfare cost of third order in the value function because in the third order approximation of the value function, all terms will drop when evaluated at the deterministic steady state, except potentially the term $V_{333,ss}$ which happens to be equal to zero (as all the terms in odd powers of χ).

Moreover, this cost of the business cycle can easily transformed into consumption equivalent units. We can compute the decrease in consumption τ that will make the household indifferent between consuming $(1-\tau)\,c_{ss}$ units per period with certainty or c_t units with uncertainty. To do so, note that our definition implies that:

$$\left(c_{ss}^{\upsilon}\left(1 - l_{ss}\right)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}} + \frac{1}{2}V_{33,ss} = \left(\left(c_{ss}\left(1-\tau\right)\right)^{\upsilon}\left(1 - l_{ss}\right)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}}$$

[6] Note that the converse is not true: changes in the policy functions, for example, those induced by changes in the persistence of the shock, modify the value of $V_{33,ss}$.

[7] Note that this quantity is not necessarily negative. Indeed, it may well be positive in many models, as in a real business cycle model with leisure choice. See Cho and Cooley (2000) for an example and further discussion.

or

$$\left((1-\tau)^{\upsilon\frac{1-\gamma}{\theta}}-1\right)\left(c_{ss}^{\upsilon}\left(1-l_{ss}\right)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}}=\frac{1}{2}V_{33,ss}$$

that implies:

$$\tau=1-\left[1+\left(c_{ss}^{\upsilon}\left(1-l_{ss}\right)^{1-\upsilon}\right)^{\frac{\gamma-1}{\theta}}\frac{1}{2}V_{33,ss}\right]^{\upsilon\frac{\gamma-1}{\theta}}$$

The policy functions for consumption, labor, and capital (our controls, the last one of which is also the endogenous state variable) can be expanded as:

$$\begin{aligned}\widetilde{c}_t &= c\left(\widetilde{k}_t,\widetilde{z}_t;\chi\right)\simeq\widetilde{c}_{ss}+c_{1,ss}\left(\widetilde{k}_t-\widetilde{k}_{ss}\right)+c_{2,ss}\left(\widetilde{z}_t-\widetilde{z}_{ss}\right)+c_{3,ss}\chi\\ l_t &= l\left(\widetilde{k}_t,\widetilde{z}_t;\chi\right)\simeq l_{ss}+l_{1,ss}\left(\widetilde{k}_t-\widetilde{k}_{ss}\right)+l_{2,ss}\left(\widetilde{z}_t-\widetilde{z}_{ss}\right)+l_{3,ss}\chi\\ \widetilde{k}_{t+1} &= k\left(\widetilde{k}_t,\widetilde{z}_t;\chi\right)\simeq\widetilde{k}_{ss}+k_{1,ss}\left(\widetilde{k}_t-\widetilde{k}_{ss}\right)+k_{2,ss}\left(\widetilde{z}_t-\widetilde{z}_{ss}\right)+k_{3,ss}\chi\end{aligned}$$

where:

$$\begin{aligned}c_{1,ss} &= c_1\left(\widetilde{k}_{ss},\widetilde{z}_{ss};0\right),c_{2,ss}=c_2\left(\widetilde{k}_{ss},\widetilde{z}_{ss};0\right),c_{3,ss}=c_3\left(\widetilde{k}_{ss},\widetilde{z}_{ss};0\right)\\ l_{1,ss} &= l_1\left(\widetilde{k}_{ss},\widetilde{z}_{ss};0\right),l_{2,ss}=l_2\left(\widetilde{k}_{ss},\widetilde{z}_{ss};0\right),l_{3,ss}=l_3\left(\widetilde{k}_{ss},\widetilde{z}_{ss};0\right)\\ k_{1,ss} &= k_1\left(\widetilde{k}_{ss},\widetilde{z}_{ss};0\right),k_{2,ss}=k_2\left(\widetilde{k}_{ss},\widetilde{z}_{ss};0\right),k_{3,ss}=k_3\left(\widetilde{k}_{ss},\widetilde{z}_{ss};0\right)\end{aligned}$$

In addition we have functions that give us the evolution of other variables of interest, like the real price of a one-period bond:[8]

$$qr_{t,t+1}=f^1\left(\widetilde{k}_t,\widetilde{z}_t;\chi\right)\simeq f_{ss}^1+f_{1,ss}^1\left(\widetilde{k}_t-\widetilde{k}_{ss}\right)+f_{2,ss}^1\widetilde{z}_t+f_{3,ss}^1\chi$$

where

$$f_{1,ss}^1=f_1^1\left(\widetilde{k}_{ss},\widetilde{z}_{ss};0\right),f_{2,ss}=f_2^1\left(\widetilde{k}_{ss},\widetilde{z}_{ss};0\right),f_{3,ss}^1=f_3^1\left(\widetilde{k}_{ss},\widetilde{z}_{ss};0\right)$$

To find the linear approximation to the value function, we take derivatives of the value function with respect to controls $\left(\widetilde{c}_t,l_t,\widetilde{k}_{t+1}\right)$, the endogenous variable $qr_{t,t+1}$, states $\left(\widetilde{k}_t,\widetilde{z}_t\right)$, and the perturbation parameter χ. These derivatives plus the value function in equation (4) give us a system of 8 equations. By setting our perturbation parameter $\chi=0$, we can solve for the 8 unknowns V_{ss}, $V_{1,ss}$, $V_{2,ss}$, $V_{3,ss}$, $\widetilde{c}_{ss}$, $\widetilde{k}_{ss}$, l_{ss}, and f_{ss}^1. As mentioned before, the system solution implies that $V_{3,ss}=0$.

To find the quadratic approximation to the value function, we come back to the 7 first

[8]For clarity we present the price of only one bond. The approximating functions of all the other bonds prices and endogenous variables can be found in an analogous manner.

derivatives found in the previous step, and we take second derivatives with respect to the states $\left(\widetilde{k}_t, \widetilde{z}_t\right)$, and the perturbation parameter χ. This step gives us 18 different second derivatives. By setting our perturbation parameter $\chi = 0$ and plugging in the results from the linear approximation, we find the value for $V_{11,ss}$, $V_{12,ss}$, $V_{13,ss}$, $V_{22,ss}$, $V_{23,ss}$, $V_{33,ss}$, $c_{1,ss}$, $c_{2,ss}$, $c_{3,ss}$, $k_{1,ss}$, $k_{2,ss}$, $k_{3,ss}$, $l_{1,ss}$, $l_{2,ss}$, $l_{3,ss}$, $f^1_{1,ss}$, $f^2_{2,ss}$, and $f^3_{3,ss}$ Since the first derivatives of the policy function depend only on the first and second derivatives of the value function, the solution of the system is such that $c_{3,ss} = k_{3,ss} = l_{3,ss} = f^1_{3,ss} = 0$, i.e., precautionary behavior depends on the third derivative of the value function, as demonstrated, for example, by Kimball (1990).

To find the cubic approximation to the value function and the second order approximation to the policy function, we take derivatives on the second derivatives found before with respect to the states $\left(\widetilde{k}_t, \widetilde{z}_t\right)$, set the perturbation parameter $\chi = 0$, plug in the results from the previous step, and solve for the unknown variables. Repeating these steps, we can get any arbitrary order approximation.

Our perturbation approach is extremely fast once it has been implemented in the computer, taking only a fraction of a second to find a solution for some particular parameter values.[9] Also, even if it is a local solution method, it is highly accurate for a long range of parameter values (Caldara, Fernández-Villaverde, Rubio-Ramírez, and Yao, 2008). Nevertheless, it is important to keep in mind some of its problems. The most important is that it may handle badly big shocks that push us very far away from the steady state. This may be a relevant concern if we want to explore, for instance, Barro's (2005) claim that rare disasters (i.e., very large shocks) may be a plausible explanation for asset pricing data. Implementing variations of the basic perturbation scheme, as a matched asymptotic expansion, could be a possible solution to this shortcoming.

3.2. Approximation

By following the previous procedure, and given some parameter values, we find a third order approximation for our value function and a second order approximation to the policy function for the controls.

[NOTE: Currently we are working on implementing a fourth order approximation to the value function and a third order expansion to the policy function. We already have the code that computes the solution but not the evaluation of the likelihood based on it.]

A third order approximation to the policy function is important to capture the time-varying risk premium (the first order approximation does not have a term premium and the

[9] There is a fixed cost in finding all the symbolic derivatives but this only has to be done once.

second order approximation has a constant one).

The controls in our model are:

$$controls_t = \left\{\widetilde{c}_t, l_t, \widetilde{k}_{t+1}\right\}'$$

with steady state values $controls_{ss}$. Also, the other endogenous variables of interest for us are:

$$end_t = \left\{\widetilde{y}_t, f_t^1, f_t^4, f_t^8, f_t^{12}, f_t^{16}, f_t^{20}\right\}'$$

since, as we will explain later, we will use as our observables output and bonds of different maturities (1 quarter, and 1, 2, 3, 4, and 5 years). Finally, the states of the model are $\widetilde{s}_t = \left(\widetilde{k}_t, \widetilde{z}_t\right)'$ and three co-states that will be convenient for simplifying our state space representation $(\widetilde{c}_{t-1}, \widetilde{y}_{t-1}, \varepsilon_{zt-1})$. We stack the states and co-states in:

$$S_t = \left(\widetilde{k}_t, \widetilde{z}_t, \widetilde{c}_{t-1}, \widetilde{y}_{t-1}, \varepsilon_{zt-1}\right)'$$

It is easier to express the solution in terms of deviations from steady state. For any variable var_t, let $\widehat{var}_t = var_t - var_{ss}$. Then, the policy function for the control variable is:

$$\begin{aligned}
\widehat{\widetilde{c}}_t &= \Psi_{1,1}\widehat{\widetilde{s}}_t + \frac{1}{2}\widehat{\widetilde{s}}_t'\Psi_{1,2}\widehat{\widetilde{s}}_t + \Psi_{1,3} \\
\widehat{l}_t &= \Psi_{2,1}\widehat{\widetilde{s}}_t + \frac{1}{2}\widehat{\widetilde{s}}_t'\Psi_{2,2}\widehat{\widetilde{s}}_t + \Psi_{2,3} \\
\widehat{\widetilde{k}}_{t+1} &= \Psi_{3,1}\widehat{\widetilde{s}}_t + \frac{1}{2}\widehat{\widetilde{s}}_t'\Psi_{3,2}\widehat{\widetilde{s}}_t + \Psi_{3,3}
\end{aligned}$$

where $\Psi_{i,1}$ is a 1×2 vector, $\Psi_{i,2}$ is a 2×2 matrix, and $\Psi_{i,3}$ a scalar that captures precautionary behavior (the second derivative of the policy function with respect to the perturbation parameter) for $i = \{1, 2, 3\}$.

The approximation function for the endogenous variables is:

$$\widehat{end}_t = \Phi_{j,1}\widehat{\widetilde{s}}_t + \frac{1}{2}\widehat{\widetilde{s}}_t'\Phi_{j,2}\widehat{\widetilde{s}}_t + \Phi_{j,3}$$

for $j = \{1, ..., 7\}$.

3.3. State Space Representation

Once we have our perturbation solution of the model, we write the state space representation that will allow us to evaluate later the likelihood of the model.

We begin with the transition equation. Using the policy and approximation functions from the previous section and the law of motion for productivity, we have:

$$\begin{pmatrix} \widehat{\widetilde{k}}_t \\ \widetilde{z}_t \\ \widehat{\widetilde{c}}_{t-1} \\ \widetilde{y}_{t-1} \\ \varepsilon_{zt-1} \end{pmatrix} = \begin{pmatrix} \Psi_{2,1}\widehat{\widetilde{s}}_{t-1} + \frac{1}{2}\widehat{\widetilde{s}}'_{t-1}\Psi_{2,2}\widehat{\widetilde{s}}_{t-1} + \Psi_{2,3} \\ \exp\left(\lambda + \sigma\varepsilon_{zt}\right) \\ \Psi_{1,1}\widehat{\widetilde{s}}_{t-1} + \frac{1}{2}\widehat{\widetilde{s}}'_{t-1}\Psi_{1,2}\widehat{\widetilde{s}}_{t-1} + \Psi_{1,3} \\ \Phi_{1,1}\widehat{\widetilde{s}}_{t-1} + \frac{1}{2}\widehat{\widetilde{s}}'_{t}\Phi_{1,2}\widehat{\widetilde{s}}_{t-1} + \Phi_{1,3} \\ \frac{1}{\sigma}\left(\log \widetilde{z}_{t-1} - \lambda\right) \end{pmatrix}$$

where we have already normalized $\chi = 1$. In a more compact notation:

$$S_t = h\left(S_{t-1}, W_t\right) \tag{5}$$

where $W_t = \varepsilon_{zt}$ is the system innovation.

Second, we build the measurement equations for our model. We will assume that we observe consumption, output, hours per capita, a 1-quarter bond, and 1, 2, 3, 4, and 5 year bonds. Consumption, output and hours per capita will provide macro information. Of the relevant aggregate variables, we do not include investment because in our model it is the residual of output minus consumption, which we already have. The price of bonds provides us with financial data that are relatively straightforward to observe. We reserve the other asset prices, like the return on equity, as a data set for validation of our estimates in section 5.

Since our prototype DSGE model has only one source of uncertainty, the productivity shock, we need to introduce measurement error to avoid stochastic singularity. While the presence of measurement error is a plausible hypothesis in the case of aggregate variables, we need to provide further justification for the price of bonds. However, it is better to explain this issue in the next section when we talk about data. Suffice it to say at this moment that we have measurement error in the price of bonds because they are the price of synthetic nominal bonds that are not directly observed and which need, moreover, to be deflated into real prices.

We will assume that all the variables are subject to some linear, Gaussian measurement error. Then, our observables $\mathbb{Y}_t$ are:

$$\mathbb{Y}_t = \left(\Delta \log \mathcal{C}_t, \Delta \log \mathcal{Y}_t, L_t, QR_{t,t+1}, QR_{t,t+4}, QR_{t,t+8}, QR_{t,t+12}, QR_{t,t+16}, QR_{t,t+20}\right)$$

where $\Delta \log \mathcal{C}_t$ is the measured first difference of log consumption per capita, $\Delta \log \mathcal{Y}_t$ is the measured first difference of log output per capita, L_t is measured hours, and $QR_{t,t+j}$ is the measured real price of a $j-$period bond.

Since our observables are different from the variables in the model, which are re-scaled, we need to map one into the other. We start with consumption. We observe:

$$\Delta \log c_t = \log c_t - \log c_{t-1}$$

By our definition of re-scaled variables, $c_t = \widetilde{c}_t x_t$, we get:

$$\log c_t = \log \widetilde{c}_t + \log x_t$$

Thus:

$$\log c_t - \log c_{t-1} = \log \widetilde{c}_t + \log x_t - \log \widetilde{c}_{t-1} + \log x_{t-1} = \log \widetilde{c}_t - \log \widetilde{c}_{t-1} + \lambda + \sigma_z \varepsilon_{zt-1}$$

Since $\widehat{\widetilde{c}}_t = \widetilde{c}_t - \widetilde{c}_{ss}$, we can write:

$$\log c_t - \log c_{t-1} = \log \left(\widetilde{c}_{ss} + \widehat{\widetilde{c}}_t \right) - \log \left(\widetilde{c}_{ss} + \widehat{\widetilde{c}}_{t-1} \right) + \lambda + \sigma_z \varepsilon_{zt-1}$$

and by plugging in the policy function for consumption:

$$\Delta \log c_t = \log \left(\widetilde{c}_{ss} + \Psi_{1,1} \widehat{\widetilde{s}}_t + \frac{1}{2} \widehat{\widetilde{s}}_t' \Psi_{1,2} \widehat{\widetilde{s}}_t + \Psi_{1,3} \right) - \log \left(\widetilde{c}_{ss} + \widehat{\widetilde{c}}_{t-1} \right) + \lambda + \sigma_z \varepsilon_{zt-1}$$

Now, the variable reported by the statistical agency is:

$$\Delta \log \mathcal{C}_t = \Delta \log c_t + \upsilon_{1,t} \text{ where } \upsilon_{1,t} \sim \mathcal{N}(0,1)$$

which is equal to

$$\Delta \log \mathcal{C}_t = \log \left(\widetilde{c}_{ss} + \Psi_{1,1} \widehat{\widetilde{s}}_t + \frac{1}{2} \widehat{\widetilde{s}}_t' \Psi_{1,2} \widehat{\widetilde{s}}_t + \Psi_{1,3} \right) - \log \left(\widetilde{c}_{ss} + \widehat{\widetilde{c}}_{t-1} \right) + \lambda + \sigma_z \varepsilon_{zt-1} + \sigma_{\upsilon 1} \upsilon_{1,t}$$

The first difference of log output is also observed with some measurement error:

$$\Delta \log \mathcal{Y}_t = \Delta \log y_t + \upsilon_{2,t} \text{ where } \upsilon_{2,t} \sim \mathcal{N}(0,1)$$

and repeating the same steps as before for output, we get:

$$\Delta \log \mathcal{Y}_t = \log \left(\widetilde{y}_{ss} + \Phi_{1,1} \widehat{\widetilde{s}}_t + \frac{1}{2} \widehat{\widetilde{s}}_t' \Phi_{1,2} \widehat{\widetilde{s}}_t + \Phi_{1,3} \right) - \log \left(\widetilde{y}_{ss} + \widehat{\widetilde{y}}_{t-1} \right) + \lambda + \sigma_z \varepsilon_{zt-1} + \sigma_{\upsilon 2} \upsilon_{2,t}$$

For hours, we have:

$$\widehat{l}_t = \Psi_{3,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}'_t\Psi_{3,2}\widehat{\widehat{s}}_t + \Psi_{3,3}$$

or:

$$L_t = l_t + \upsilon_{3,t} = l_{ss} + \Psi_{3,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}'_t\Psi_{3,2}\widehat{\widehat{s}}_t + \Psi_{3,3} + \sigma_{\upsilon 3}\upsilon_{3,t} \text{ where } \upsilon_{3,t} \sim \mathcal{N}(0,1)$$

For each of the bonds of different maturity we have:

$$\begin{aligned}
QR_{t,t+1} &= qr_{t,t+1} + \upsilon_{4,t} = \widetilde{f}^{1}_{ss} + \Phi_{2,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}'_t\Phi_{2,2}\widehat{\widehat{s}}_t + \Phi_{2,3} + \sigma_{\upsilon 4}\upsilon_{4,t} \text{ where } \upsilon_{4,t} \sim \mathcal{N}(0,1) \\
QR_{t,t+4} &= qr_{t,t+4} + \upsilon_{5,t} = \widetilde{f}^{4}_{ss} + \Phi_{3,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}'_t\Phi_{3,2}\widehat{\widehat{s}}_t + \Phi_{3,3} + \sigma_{\upsilon 5}\upsilon_{5,t} \text{ where } \upsilon_{5,t} \sim \mathcal{N}(0,1) \\
QR_{t,t+8} &= qr_{t,t+8} + \upsilon_{6,t} = \widetilde{f}^{8}_{ss} + \Phi_{4,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}'_t\Phi_{4,2}\widehat{\widehat{s}}_t + \Phi_{4,3} + \sigma_{\upsilon 6}\upsilon_{6,t} \text{ where } \upsilon_{6,t} \sim \mathcal{N}(0,1) \\
QR_{t,t+12} &= qr_{t,t+12} + \upsilon_{7,t} = \widetilde{f}^{12}_{ss} + \Phi_{5,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}'_t\Phi_{5,2}\widehat{\widehat{s}}_t + \Phi_{5,3} + \sigma_{\upsilon 7}\upsilon_{7,t} \text{ where } \upsilon_{7,t} \sim \mathcal{N}(0,1) \\
QR_{t,t+16} &= qr_{t,t+16} + \upsilon_{8,t} = \widetilde{f}^{16}_{ss} + \Phi_{6,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}'_t\Phi_{6,2}\widehat{\widehat{s}}_t + \Phi_{6,3} + \sigma_{\upsilon 8}\upsilon_{8,t} \text{ where } \upsilon_{8,t} \sim \mathcal{N}(0,1) \\
QR_{t,t+20} &= qr_{t,t+20} + \upsilon_{9,t} = \widetilde{f}^{20}_{ss} + \Phi_{7,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}'_t\Phi_{7,2}\widehat{\widehat{s}}_t + \Phi_{7,3} + \sigma_{\upsilon 9}\upsilon_{9,t} \text{ where } \upsilon_{9,t} \sim \mathcal{N}(0,1)
\end{aligned}$$

Putting the different pieces together, we have the measurement equation:

$$\begin{pmatrix} \Delta \log \mathcal{C}_t \\ \Delta \log \mathcal{Y}_t \\ L_t \\ QR_{t,t+1} \\ QR_{t,t+4} \\ QR_{t,t+8} \\ QR_{t,t+12} \\ QR_{t,t+16} \\ QR_{t,t+20} \end{pmatrix} = \begin{pmatrix} \log\left(\frac{\widetilde{c}_{ss}+\Psi_{1,1}\widehat{\widehat{s}}_t+\frac{1}{2}\widehat{\widehat{s}}'_t\Psi_{1,2}\widehat{\widehat{s}}_t+\Psi_{1,3}}{\widetilde{c}_{ss}+\widehat{\widetilde{c}}_{t-1}}\right) + \lambda + \sigma_z\varepsilon_{zt-1} \\ \log\left(\frac{\widetilde{y}_{ss}+\Phi_{1,1}\widehat{\widehat{s}}_t+\frac{1}{2}\widehat{\widehat{s}}'_t\Phi_{1,2}\widehat{\widehat{s}}_t+\Phi_{1,3}}{\widetilde{y}_{ss}+\widehat{\widetilde{y}}_{t-1}}\right) + \lambda + \sigma_z\varepsilon_{zt-1} \\ l_{ss} + \Psi_{3,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}'_t\Psi_{3,2}\widehat{\widehat{s}}_t + \Psi_{3,3} \\ \widetilde{f}^{1}_{ss} + \Phi_{2,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}'_t\Phi_{2,2}\widehat{\widehat{s}}_t + \Phi_{2,3} \\ \widetilde{f}^{4}_{ss} + \Phi_{3,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}'_t\Phi_{3,2}\widehat{\widehat{s}}_t + \Phi_{3,3} \\ \widetilde{f}^{8}_{ss} + \Phi_{4,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}'_t\Phi_{4,2}\widehat{\widehat{s}}_t + \Phi_{4,3} \\ \widetilde{f}^{12}_{ss} + \Phi_{5,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}'_t\Phi_{5,2}\widehat{\widehat{s}}_t + \Phi_{5,3} \\ \widetilde{f}^{16}_{ss} + \Phi_{6,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}'_t\Phi_{6,2}\widehat{\widehat{s}}_t + \Phi_{6,3} \\ \widetilde{f}^{20}_{ss} + \Phi_{7,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}'_t\Phi_{7,2}\widehat{\widehat{s}}_t + \Phi_{7,3} \end{pmatrix} + \begin{pmatrix} \sigma_{\upsilon 1}\upsilon_{1,t} \\ \sigma_{\upsilon 2}\upsilon_{2,t} \\ \sigma_{\upsilon 3}\upsilon_{3,t} \\ \sigma_{\upsilon 4}\upsilon_{4,t} \\ \sigma_{\upsilon 5}\upsilon_{5,t} \\ \sigma_{\upsilon 6}\upsilon_{6,t} \\ \sigma_{\upsilon 7}\upsilon_{7,t} \\ \sigma_{\upsilon 8}\upsilon_{8,t} \\ \sigma_{\upsilon 9}\upsilon_{9,t} \end{pmatrix}$$

In a more compact notation:

$$\mathbb{Y}_t = g\left(S_t, V_t\right) \tag{6}$$

where $V_t = (\upsilon_{1,t}, \upsilon_{2,t}, \upsilon_{3,t}, \upsilon_{4,t}, \upsilon_{5,t}, \upsilon_{6,t}, \upsilon_{7,t}, \upsilon_{8,t}, \upsilon_{9,t})$ is the system noise.

4. Analysis of the Model

In this section, we analyze the behavior of the model by exploiting the convenient structure of our perturbation solution. This study will help us to understand the equilibrium dynamics and how the likelihood function will identify the different parameters from the variation in the data. This step is crucial because it could be plausible to entertain the idea that the richer structure of Epstein-Zin preferences is not identified in the data (as in the example built by Kocherlakota, 1990b). Fortunately, we will show that our model has sufficiently rich dynamics that we can learn from the data. This is not a surprise, though, as it confirms previous, although somehow more limited, theoretical results. In a simpler environment, when output growth follows a Markov process, Wang (1993) shows that the preference parameters of Epstein-Zin preferences are generically recoverable from the price of equity or from the price of bonds. Furthermore, Wang (1993) shows that equity and bond prices are generically unique and smooth with respect to parameters.

First, note that the deterministic steady state of the model with Epstein-Zin preferences is the same as the one in the model with standard utility function and IES equal to ψ. To see that, we can evaluate the value function at $\chi = 0$:

$$\begin{aligned} V\left(k_t, z_t; 0\right) &= \max_{c_t, l_t} \left[\left(c_t^{\upsilon} \left(1 - l_t\right)^{1-\upsilon} \right)^{\frac{1-\gamma}{\theta}} + \beta \left(V^{1-\gamma}\left(k_{t+1}, z_{t+1}; 0\right)\right)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}} \\ &= \max_{c_t, l_t} \left[\left(c_t^{\upsilon} \left(1 - l_t\right)^{1-\upsilon} \right)^{\frac{1-\gamma}{\theta}} + \beta \left(V\left(k_{t+1}, z_{t+1}; 0\right)\right)^{\frac{1-\gamma}{\theta}} \right]^{\frac{\theta}{1-\gamma}} \end{aligned}$$

which is just a monotone transformation of the value function of the standard model:

$$\begin{aligned} V\left(k_t, z_t; 0\right)^{\frac{1-\gamma}{\theta}} &= \max_{c_t, l_t} \left(c_t^{\upsilon} \left(1 - l_t\right)^{1-\upsilon} \right)^{\frac{1-\gamma}{\theta}} + \beta \left(V\left(k_{t+1}, z_{t+1}; 0\right)\right)^{\frac{1-\gamma}{\theta}} \Rightarrow \\ \widetilde{V}\left(k_t, z_t; 0\right) &= \max_{c_t, l_t} \left(c_t^{\upsilon} \left(1 - l_t\right)^{1-\upsilon} \right)^{\frac{1-\gamma}{\theta}} + \beta \widetilde{V}\left(k_{t+1}, z_{t+1}; 0\right) \end{aligned}$$

Second, the first order approximation of the policy function:

$$\begin{aligned} \widehat{\widetilde{c}}_t &= \Psi_{1,1} \widehat{\widetilde{s}}_t \\ \widehat{l}_t &= \Psi_{2,1} \widehat{\widetilde{s}}_t \\ \widehat{\widetilde{k}}_{t+1} &= \Psi_{3,1} \widehat{\widetilde{s}}_t \end{aligned}$$

is also equivalent to the first order approximation of the policy function of the model with standard utility function and IES ψ. Putting it in slightly different terms, the matrices $\Psi_{i,1}$'s depend on ψ but not on γ. To obtain this result, we need only to inspect the derivatives involved in the solution of the system and check that risk aversion, γ, cancels out in each equation (we do not include the derivatives here in the interest of brevity). This result is not surprising: the first order approximation to the policy function displays certainty equivalence. Hence, it is independent of the variance of the shock indexed by χ. But if the solution is independent of χ, it must also be independent of γ, because otherwise it will not provide the right answer in the case when $\chi = 0$ and γ is irrelevant. Remember, though, from section 3, that certainty equivalence in policy functions does not imply welfare equivalence. Indeed, welfare is affected by the term $\frac{1}{2}V_{33,ss}$ in the approximation to the value function, which depends on γ. By reading this term, we can evaluate the first order approximation of the welfare cost of aggregate fluctuations as a function of risk aversion.

In the second order approximation,

$$\begin{aligned} \widehat{\widehat{c}}_t &= \Psi_{1,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}_t'\Psi_{1,2}\widehat{\widehat{s}}_t + \Psi_{1,3} \\ \widehat{l}_t &= \Psi_{2,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}_t'\Psi_{2,2}\widehat{\widehat{s}}_t + \Psi_{2,3} \\ \widehat{\widehat{k}}_{t+1} &= \Psi_{3,1}\widehat{\widehat{s}}_t + \frac{1}{2}\widehat{\widehat{s}}_t'\Psi_{3,2}\widehat{\widehat{s}}_t + \Psi_{3,3} \end{aligned}$$

we find, again from inspection of the derivatives, that all the terms in the matrices $\Psi_{i,2}$'s are the same in the model with Epstein-Zin preferences and in the model with standard utility function. The only differences are the constant terms $\Psi_{i,3}$'s. These three constants change with the value of γ. This might seem a small difference, but it influences the equilibrium dynamics in an interesting way. The terms $\Psi_{i,3}$'s capture precautionary behavior and they move the ergodic distribution of states, affecting, through this channel, allocations, prices, and welfare up to a first order. More concretely, by changing the mean of capital in the ergodic distribution, risk aversion influences the average level of the yield curve.[10]

To illustrate this channel, we simulate the model for a benchmark calibration. We fix the discount factor $\beta = 0.9985$ to a number slightly smaller than one that ensures the existence of equilibrium with growth in the deterministic case (Kocherlakota, 1990a). The parameter that governs labor supply, $\upsilon = 0.34$, matches the microeconomic evidence of labor supply to around one-third of available time in the deterministic steady state. We set $\zeta = 0.3$ to match labor share of national income. The depreciation rate $\delta = 0.0294$ fixes the investment/output ratio, and $\lambda = 0.0047$ matches the average per capita output growth in the data sample that

[10] Our result generalizes previous findings in the literature. For example, Piazzesi and Schneider (2006) also derive, using an endowment economy and fixing the IES to one, that the average yield depends on having recursive or expected preferences, but that the dynamics will be the same for both types of preferences. Dolmas (2006) checks that, for a larger class of utility functions that includes Epstein-Zin preferences, the business cycle dynamics of the model are esentially identical to the ones of the model with standard preferences.

we will use below. Finally, $\sigma = 0.007$ follows the stochastic properties of the Solow residual of the U.S. economy. Table 4.1 gathers the different values of the calibration.

4.1: Calibrated Parameters

Parameter	β	υ	δ	ζ	λ	σ
Value	0.9985	0.34	0.0294	0.3	0.0047	0.007

Since we do not have a tight prior on the values for γ and ψ, and we want to explore the dynamics of the model for a reasonable range of values, we select 3 values for γ, 2, 20, and 100, and three values for ψ, 0.5, 1, and 2, which bracket most of the values used in the literature (although many authors prefer even smaller values for ψ, we found that the results for smaller ψ's do not change much from the case when $\psi = 0.5$).[11] We then compute the model for all nine combinations of values of γ and ψ, i.e., $\{2, 0.5\}$, $\{20, 0.5\}$, $\{100, 0.5\}$, and so on.

We simulate the model, starting from the deterministic steady state, for 1000 periods, and using the policy functions for each of the nine combinations of risk aversion and IES discussed above. To make the comparison meaningful, the shocks are common across all paths. We disregard the first 800 periods as a burn-in. The burn-in is important because it eliminates the transition from the deterministic steady state of the model to the middle regions of the ergodic distribution of capital. This is usually achieved in many fewer periods than the ones in our burn-in, but we want to be conservative in our results.

Figure 4.1 plots the simulated paths for capital. We have nine panels. In the first column, we vary between panels the risk aversion, γ, and plot within each panel, the realizations for all three different values of the IES. In the second column, we invert the order, changing the IES across panels and plotting different values of γ within each plot. With our plotting strategy, we can gauge the effect of changing one parameter at a time more easily. We extract two lessons from figure 4.1. First, all nine paths are very similar, except in their level. Figure 4.2 shows this point more clearly: we follow the same convention as in figure 4.1 except that we subtract the mean from each series. The second lesson is that while changing the IES from 0.5 to 1 and from 1 to 2 has a significant effect on the level of capital (around 7.6 percent when we move from 0.5 to 1 and 3.8 percent when we move from 1 to 2), changing risk aversion has a much smaller effect. Increasing γ from 2 to 20 increases the mean of capital by less than 1 percent and moving from 20 to 100 raises capital by around 3 percent.

These patterns repeat themselves in our next figures, where we plot output (figures 4.3 and 4.4), consumption (figures 4.5 and 4.6), the price of a 1-quarter bond (figures 4.7 and 4.8),

[11] When the IES is equal to 1, the utility from current consumption and leisure collapses to a logarithmic form.

labor (figures 4.9 and 4.10), and dividends (figures 4.11 and 4.12). There is a third aspect of interest that is difficult to see in the figures for capital or output but which is clearer in the figures for labor or dividends. Once we remove the mean of the series, all the simulated paths are nearly identical as we change γ but we keep the IES constant (second column of figures 4.1 to 4.12). This basically tells us that there is no information in the macro data about γ beyond a weak effect on the level of the variables. However, when we keep γ constant but we change the IES, there is some variation in the simulated paths, indicating that the variables have information, beyond the mean, about the value of IES.

Figure 4.13, which follows the same plotting convention as the previous figures, draws the yield curve evaluated at the deterministic steady state of the economy. A striking feature of the second column is how the IES affects only the level of the yield curve but not its curvature. Raising the IES only has the effect of translating down the curve. Interestingly, we can see how with our benchmark calibration and values of γ from 20 to 100 we can account for the mean level of the yields reported in the data (we will report later that the mean of the 1-quarter yield in our sample is 2.096 percent). It is in the first column, as we vary risk aversion, that we appreciate how the slope of the curve can change.

All the yield curves in figure 4.13 have a negative slope, which seems to contradict the empirical observations. However, we need to remember that the model produces a yield curve that is a function of the states. Figure 4.13 plots only the curves for some particular values of the state, in this case the deterministic steady state. Because of risk aversion and the precautionary saving it induces, capital will usually take values above the deterministic steady state (the constant terms $\Psi_{i,3}$'s are positive). Hence, if we are at the steady state, capital is relatively scarce with respect to its mean and the real interest rate relatively high. Since $\Psi_{3,3}$ is positive, the next periods will be times of accumulation of capital. Consequently, interest rates will fall in the next periods, inducing the negative slope of the curve.

To illustrate this point, we repeat our exercise in figure 4.14, where we plot the yield curves in the mean of the ergodic distribution, and in figure 4.15, for values of capital that are 0.6 percent higher than the steady state. Of special interest is figure 4.15, where capital is sufficiently above the steady state that the curves become positive for all nine combinations of parameter values. Similar figures can easily be generated by variations in the productivity shock: a positive productivity shock tends to make the curve downward sloping and a negative shock, upward sloping.[12] These variations in the level and slope of the yield curve will be key to the success of our estimation exercise: the likelihood will be shaped by all the information about states and parameter values encoded in the financial data.

[12] Which some argue contradicts the observed cyclicality of the yield curves and hence uncovers one further weakness of our model. We will revisit this point in our results section.

5. Likelihood

The parameters of the model are stacked in the vector:

$$\Upsilon = \{\beta, \gamma, \psi, \lambda, \zeta, \delta, \tau, \sigma_{\varepsilon}, \sigma_{\omega}, \sigma_{1\upsilon}, \sigma_{2\upsilon}, \sigma_{3\upsilon}, \sigma_{4\upsilon}, \sigma_{5\upsilon}, \sigma_{6\upsilon}, \sigma_{7\upsilon}, \sigma_{8\upsilon}, \sigma_{9\upsilon}\}$$

and we can define the likelihood function $\mathcal{L}\left(\mathbb{Y}^T; \Upsilon\right)$ as the probability of the observations given some parameter values, where $\mathbb{Y}^T = \{\mathbb{Y}_t\}_{t=1}^T$ is the sequence of observations. We will adopt the notation $\mathbb{X}^T = \{\mathbb{X}_t\}_{t=1}^T$ for an arbitrary variable $\mathbb{X}$ below.

Unfortunately, this likelihood is difficult to evaluate. Our procedure to address this problem is to use a sequential Monte Carlo method. First, we factorize the likelihood into its conditional components:

$$\mathcal{L}\left(\mathbb{Y}^T; \Upsilon\right) = \prod_{t=1}^{T} \mathcal{L}\left(\mathbb{Y}_t | \mathbb{Y}^{t-1}; \Upsilon\right)$$

where $\mathcal{L}\left(\mathbb{Y}_1 | \mathbb{Y}^0; \Upsilon\right) = \mathcal{L}\left(\mathbb{Y}_1; \Upsilon\right)$.

Now, we can condition on the states and integrate with respect to them:

$$\mathbb{Y}_t = g\left(S_t, V_t\right) \tag{7}$$

$$\mathcal{L}\left(\mathbb{Y}^T; \Upsilon\right) = \int \mathcal{L}\left(\mathbb{Y}_1 | S_0; \Upsilon\right) dS_0 \prod_{t=2}^{T} \int \mathcal{L}\left(\mathbb{Y}_t | S_t; \Upsilon\right) p\left(S_t | \mathbb{Y}^{t-1}; \Upsilon\right) dS_t \tag{8}$$

This expression illustrates how the knowledge of $p\left(S_0; \Psi\right)$ and the sequence $\{p\left(S_t | \mathbb{Y}^{t-1}; \Upsilon\right)\}_{t=2}^T$ is crucial. If we know S_t, computing $\mathcal{L}\left(\mathbb{Y}_t | S_t; \Psi\right)$ is relatively easy. Conditional on S_t, the measurement equation (6) is a change of variables from V_t to $\mathbb{Y}_t$. Since we know that the noise is normally distributed, an application of the change of variable formula allows us to evaluate the required probabilities. In the same way, if we have S_0, we can compute $\mathcal{L}\left(\mathbb{Y}_1 | S_0; \Upsilon\right)$ by employing (5) and the measurement equation (6).

However, given our model, we cannot characterize the sequence $\{p\left(S_t | \mathbb{Y}^{t-1}; \Upsilon\right)\}_{t=2}^T$ analytically. Even if we could, these two previous computations still leave open the issue of how to solve for the T integrals in (8).

A common solution to the problems of knowing the state distributions and how to compute the required integrals is to substitute $\{p\left(S_t | \mathbb{Y}^{t-1}; \Upsilon\right)\}_{t=1}^T$and $p\left(S_0; \Upsilon\right)$ by an empirical distribution of draws from them. If we have such draws, we can approximate the likelihood as:

$$\mathcal{L}\left(\mathbb{Y}^T; \Upsilon\right) \simeq \frac{1}{N} \sum_{i=1}^{N} \mathcal{L}\left(\mathbb{Y}_1 | s_{0|0}^i; \Upsilon\right) \prod_{t=2}^{T} \frac{1}{N} \sum_{i=1}^{N} \mathcal{L}\left(\mathbb{Y}_t | s_{t|t-1}^i; \Upsilon\right)$$

where $s^i_{0|0}$ is the draw i from $p(S_0;\Upsilon)$ and $s^i_{t|t-1}$ is the draw i from $p(S_t|\mathbb{Y}^{t-1};\Upsilon)$. Del Moral and Jacod (2002) and Künsch (2005) provide weak conditions under which the right-hand side of the previous equation is a consistent estimator of $\mathcal{L}\left(\mathbb{Y}^T;\Upsilon\right)$ and a central limit theorem applies. A Law of Large Numbers will ensure that the approximation error goes to 0 as the number of draws, N, grows.

Drawing from $p(S_0;\Upsilon)$ is straightforward in our model. Given parameter values, we solve the model and simulate from the ergodic distribution of states. Santos and Peralta-Alva (2005) show that this procedure delivers the empirical distribution of s^i_0 that we require. Drawing from $\{p(S_t|\mathbb{Y}^{t-1};\Upsilon)\}_{t=2}^T$ is more challenging. A popular approach to do so is to apply the particle filter (see Fernández-Villaverde and Rubio-Ramírez, 2005 and 2007 for a more detailed explanation and examples of how to implement the filter to the estimation of non-linear and/or non-normal DSGE models).[13]

The basic idea of the filter is to generate draws through sequential importance resampling (SIR), which extends importance sampling to a sequential environment. The following proposition, due to Rubin (1998), formalizes the idea:

Proposition 1. *Let* $\left\{s^i_{t|t-1}\right\}_{i=1}^N$ *be a draw from* $p(S_t|\mathbb{Y}^{t-1};\Upsilon)$. *Let the sequence* $\{\widetilde{s}^i_t\}_{i=1}^N$ *be a draw with replacement from* $\left\{s^i_{t|t-1}\right\}_{i=1}^N$ *where the resampling probability is given by*

$$q^i_t = \frac{\mathcal{L}\left(\mathbb{Y}_t|s^i_{t|t-1};\Upsilon\right)}{\sum_{i=1}^N \mathcal{L}\left(\mathbb{Y}_t|s^i_{t|t-1};\Upsilon\right)},$$

Then $\{\widetilde{s}^i_t\}_{i=1}^N$ *is a draw from* $p(S_t|\mathbb{Y}^t;\Upsilon)$.

Proposition 1, a direct application of Bayes' theorem, shows how we can take a draw $\left\{s^i_{t|t-1}\right\}_{i=1}^N$ from $p(S_t|\mathbb{Y}^{t-1};\Upsilon)$ to get a draw $\left\{s^i_{t|t}\right\}_{i=1}^N$ from $p(S_t|\mathbb{Y}^t;\Upsilon)$ by building importance weights depending on $\mathbb{Y}_t$. This result is crucial because it allows us to incorporate the information in $\mathbb{Y}_t$ to change our current estimate of S_t. Thanks to SIR, the Monte Carlo method achieves sufficient accuracy in a reasonable amount of time (Arulampalam *et al.*, 2002). A naïve Monte Carlo, in comparison, would just draw simultaneously a whole sequence of $\left\{\left\{s^i_{t|t-1}\right\}_{i=1}^N\right\}_{t=1}^T$ without resampling. Unfortunately, this naïve scheme diverges because all the sequences become arbitrarily far away from the true sequence of states, which is a zero measure set. Then, the sequence of simulated states that is closer to the true state

[13] See also the collection of papers in Doucet, de Freitas, and Gordon (2001), which includes improved sequential Monte Carlo algorithms, like Pitts and Shephard's (1999) auxiliary particle filter.

in probability dominates all the remaining ones in weight. Simple simulations shows that the degeneracy appears even after very few steps.

Given $\left\{s_{t|t}^i\right\}_{i=1}^N$, we draw N exogenous productivity shocks from the normal distribution, apply the law of motion for states that relates the $s_{t|t}^i$ and the shocks ε_{t+1}^i to generate $\left\{s_{t+1|t}^i\right\}_{i=1}^N$. This transition step puts us back at the beginning of proposition 1, but with the difference that we have moved forward one period in our conditioning, from $t|t-1$ to $t+1|t$.

The following pseudocode, copied from Fernández-Villaverde and Rubio-Ramírez (2007), summarizes the algorithm:

Step 0, Initialization: `Set` $t \rightsquigarrow 1$. `Sample` N `values` $\left\{s_{0|0}^i\right\}_{i=1}^N$ from $p(S_0; \Upsilon)$.

Step 1, Prediction: `Sample` N `values` $\left\{s_{t|t-1}^i\right\}_{i=1}^N$ `using` $\left\{s_{t-1|t-1}^i\right\}_{i=1}^N$, `the law of motion for states and the distribution of shocks` ε_t.

Step 2, Filtering: `Assign to each draw` $\left(s_{t|t-1}^i\right)$ `the weight` q_t^i `in proposition 1.`

Step 3, Sampling: `Sample` N `times with replacement from` $\left\{s_{t|t-1}^i\right\}_{i=1}^N$ `using the probabilities` $\{q_t^i\}_{i=1}^N$. `Call each draw` $\left(s_{t|t}^i\right)$. `If` $t < T$ `set` $t \rightsquigarrow t+1$ `and go to step 1. Otherwise stop.`

---

6. Estimation

6.1. Data

We take as our sample the period 1984:1 to 2006:4. The end of the sample is determined by the most current release of data. The start of the sample aims to cover only the period after the "Great Moderation" (Stock and Watson, 2002). We eliminate the previous observations because our model has little to say about the change in the volatility of output (see, among others, the discussions of Justiniano and Primiceri, 2007, or Sims and Zha, 2006). Moreover, elsewhere (Fernández-Villaverde and Rubio-Ramírez, 2007 and 2008), we have argued that attempting to account for the observations before and after the Great Moderation using a DSGE model without additional sources of variation (either in structural parameters or in shock variances) may induce misleading estimates. In our application, this causes particular concern because of the substantial changes in interest rates and asset pricing behavior during the 1970s and early 1980s. Also, Stock and Watson (2007) detect a change in the process for inflation around 1984, and they recommend splitting the inflation sample using the same criterion that we use. Since we follow their inflation specification, we find it natural to be consistent with their choice.

Our output and consumption data come from the Bureau of Economic Analysis NIPA data. We define nominal consumption as the sum of personal consumption expenditures on non-durable goods and services. We define nominal gross investment as the sum of personal consumption expenditures on durable goods, private non-residential fixed investment, and private residential fixed investment. Per capita nominal output is defined as the ratio between our nominal output series and the civilian non-institutional population over 16. Finally, the hours worked per capita series is constructed with the index of average weekly hours in the private sector and the civilian non-institutional population over 16. Since our model implies that hours worked per capita are between 0 and 1, we express hours as the percentage of available week time, such that the average of hours worked is one-third. For inflation, we use the price index for gross domestic product.

The data on bond yields with maturities one year and longer are from CRSP Fama-Bliss discount bond files, which have fully taxable, non-callable, non-flower bonds. Fama and Bliss construe their data by interpolating observations from traded Treasuries. This procedure introduces measurement error, possibly correlated across time.[14] The 1-quarter yield is from the CRSP Fama risk-free rate file. The mean of this yield is 2.096 percent, which is a bit higher than the postwar mean but nearly the same that the mean of 2.020 percent for the period 1891-1998 reported by Campbell (2003), table 1, page 812. To match the frequency of the CRSP data set with NIPA data, we transform the monthly yield observations into quarterly observations with an arithmetic mean.

6.2. Estimation of the Process for Inflation

A preliminary step in our empirical work is to estimate the process for inflation:

$$\log \frac{p_{t+1}}{p_t} - \log \frac{p_t}{p_{t-1}} = \omega_{t+1} - \tau\omega_t \text{ where } \omega_t \sim \mathcal{N}(0, \sigma_\omega) \tag{9}$$

We report in table 1 our estimates (with the standard deviations in parenthesis), and for comparison purposes, Stock and Watson's point estimates in table 3, page 13 of their 2007

[14]Furthermore, Piazzesi (2003) alerts us that there were data entry errors at the time at the CRSP, some rather obvious. This warning raises the possibility that some less obvious errors may still persist in the database.

paper, which use a sample that stops in 2004.[15]

Table 5.1: Estimates of the MA(1)

	Our Estimate	Stock and Watson
τ	0.6469 (0.0877)	0.656 (0.088)
σ_ω	0.0019 (5.8684e-007)	0.0019 (4.375e-007

To gauge the performance of Stock and Watson's specification, we plot in figure 5.1, in the top panel, the one period ahead expected inflation versus current inflation, and in the bottom panel, the expected inflation versus realized inflation. This second draw illustrates the good performance of the MA representation in terms of fit.

Once we have the inflation process, we can find the real bond yields, as described in the second section. In figure 5.2, we plot the evolution over time of the yields of the six bonds that we include in our observables. In figure 5.3, we plot the median real yield curve, which, as documented repeatedly by the literature, has a positive slope.

6.3. Estimation Algorithms

Our paper emphasizes likelihood estimation of DSGE models. Consequently, we will show results both for maximum likelihood and for Bayesian estimation.

Obtaining the maximum likelihood point estimate is complicated because of the shape of the likelihood function. We employ a mixed of a grid search and a simulated annealing algorithm.

With respect to Bayesian inference, we mentioned in the main part of the text that the posterior of the model:

$$p\left(\Psi|\,\mathbb{Y}^T\right) \propto \mathcal{L}\left(\mathbb{Y}^T;\Psi\right) p\left(\Psi\right)$$

is difficult, if not impossible, to characterize. However, we can draw from it and build its empirical counterpart using a Metropolis-Hastings algorithm. We follow the standard practice and choose a random walk proposal, $\Psi_i^* = \Psi_{i-1} + \kappa_i$, $\kappa_i \sim \mathcal{N}\left(0, \Sigma_\kappa\right)$, where Σ_κ is a scaling matrix. This matrix is selected to get the appropriate acceptance ratio of proposals (Roberts, Gelman and Gilks, 1997).

To reduce the "chatter" of the problem, we will keep the innovations in the particle filter (i.e., the draws from the exogenous shock distributions and the resampling probabilities) constant across different passes of the Metropolis-Hastings algorithm. As pointed out

[15] Stock and Watson multiply inflation for 400. The point estimate and standard error of τ is unaffected. We adjust their point estimate of σ_ω accordingly.

by McFadden (1989) and Pakes and Pollard (1989), this is required to achieve stochastic equicontinuity, and even if this condition is not strictly necessary in a Bayesian framework, it reduces the numerical variance of the procedure.

6.4. Results

Figures 5.4 to 5.12 [TO BE COMPLETED].[TO BE COMPLETED]

Ravena and Seppälä (2007) find that inflation risk premia are very small and display little volatility.

6.5. Implications for Other Assets

We can explore the implications of the model and the point estimates for other assets. For instance, we can study the valuation of equity. A preliminary step is to think about all the net payments to capital in the economy, either dividends to shares or coupons to corporate bonds. To compute this net payments to capital, note that, with complete markets, the cash-flow of the firm is

$$d_t + i_t = y_t - w_t l_t$$

where d_t is the net payments to capital. Now, use the observation that $y_t - w_t l_t = r_t k_t$ to find $d_t + i_t = r_t k_t$.

Second, from the aggregate resource constraint:

$$c_t + i_t = w_t l_t + r_t k_t$$

Then:

$$c_t + i_t = w_t l_t + d_t + i_t$$

and consequently, we have:

$$d_t = c_t - w_t l_t = r_t k_t - i_t$$

Given our solution method, computing d_t is straightforward. Once we have the process for d_t is also rather direct to price a claim to it.

Also, we have a claim to equity return:

$$r_t^e = \frac{k_{t+1} + d_t}{k_t} = \frac{k_{t+1} + r_t k_t - i_t}{k_t} = \frac{r_t k_t + (1-\delta) k_t}{k_t} = r_t + 1 - \delta$$

[TO BE COMPLETED]

7. Conclusions

We have shown how to perform likelihood inference in DSGE models with Epstein-Zin preferences. We presented a simple perturbation method that computes the solution of the model quickly and accurately and explained how the particle filter allows the evaluation of the likelihood function.

This paper points out to many lines of future research. Among them, we highlight two. First, we want to take this model more seriously and enrich it with mechanisms like long-run risk and habit persistence. Can these mechanisms explain the empirical problems of the model when we have recursive preferences? Which mechanisms are and which are not necessary in this context? Second, we want to further explore the computation and estimation of DSGE models with exotic preferences (see the review of Backus, Routledge, and Zin, 2004). The implementation and empirical validation of all those new specifications of preferences is an open field were much work is required.

8. Appendix

This appendix offers further details in some aspects of the paper. First, we show how to compute expectations for price from the process for inflation.

8.1. Computing Expectations for Prices

As explained in the main text, we follow Stock and Watson in postulating an IMA(1,1) process for inflation:

$$\log \frac{p_{t+1}}{p_t} - \log \frac{p_t}{p_{t-1}} = \omega_{t+1} - \tau\omega_t$$

This process implies that:

$$\frac{p_{t+1}}{p_t} = \frac{p_t}{p_{t-1}} \exp\left(\omega_{t+1} - \tau\omega_t\right)$$

or, equivalently,

$$\frac{p_t}{p_{t+1}} = \frac{p_{t-1}}{p_t} \exp\left(-\omega_{t+1} + \tau\omega_t\right)$$

i.e., the law of motion of the inverse of the ratio of prices between periods t and $t+1$.

We are interested in deriving expressions for ratios of prices between periods t and $t+j$. To illustrate the procedure, we proceed recursively. We start with one period ahead:

$$\begin{aligned}
\frac{p_t}{p_{t+2}} &= \frac{p_{t+1}}{p_{t+2}} \frac{p_t}{p_{t+1}} \\
&= \frac{p_t}{p_{t+1}} \exp\left(-\omega_{t+2} + \tau\omega_{t+1}\right) \frac{p_t}{p_{t+1}} \\
&= \left(\frac{p_t}{p_{t+1}}\right)^2 \exp\left(-\omega_{t+2} + \tau\omega_{t+1}\right) \\
&= \left(\frac{p_{t-1}}{p_t}\right)^2 \exp\left(-2\omega_{t+1} + 2\tau\omega_t - \omega_{t+2} + \tau\omega_{t+1}\right) \\
&= \left(\frac{p_{t-1}}{p_t}\right)^2 \exp\left(-\omega_{t+2} + (\tau - 2)\,\omega_{t+1} + 2\tau\omega_t\right)
\end{aligned}$$

We move to two periods ahead:

$$\begin{aligned}
\frac{p_t}{p_{t+3}} &= \frac{p_{t+2}}{p_{t+3}}\frac{p_t}{p_{t+2}} \\
&= \frac{p_{t+1}}{p_{t+2}}\exp\left(-\omega_{t+3}+\tau\omega_{t+2}\right)\frac{p_t}{p_{t+2}} \\
&= \frac{p_t}{p_{t+1}}\exp\left(-\omega_{t+2}+\tau\omega_{t+1}-\omega_{t+3}+\tau\omega_{t+2}\right)\frac{p_t}{p_{t+2}} \\
&= \frac{p_{t-1}}{p_t}\exp\left(-\omega_{t+1}+\tau\omega_t-\omega_{t+2}+\tau\omega_{t+1}-\omega_{t+3}+\tau\omega_{t+2}\right)\frac{p_t}{p_{t+2}} \\
&= \frac{p_{t-1}}{p_t}\exp\left(-\omega_{t+3}+(\tau-1)\,\omega_{t+2}+(\tau-1)\,\omega_{t+1}+\tau\omega_t\right)\left(\frac{p_{t-1}}{p_t}\right)^2\exp\left(-\omega_{t+2}+(\tau-2)\,\omega_{t+1}+2\tau\omega_t\right) \\
&= \left(\frac{p_{t-1}}{p_t}\right)^3\exp(-\omega_{t+3}+(\tau-2)\,\omega_{t+2}+(2\tau-3)\,\omega_{t+1}+3\tau\omega_t)
\end{aligned}$$

Iterating, we get to the general formula:

$$\frac{p_t}{p_{t+j}} = \left(\frac{p_{t-1}}{p_t}\right)^j \exp(\sum_{i=0}^{j-1}(i\tau-(i+1))\,\omega_{t+j}+j\tau\omega_t)$$

With the formula finding expectations is direct:

$$\begin{aligned}
\mathbb{E}_t\frac{p_t}{p_{t+j}} &= \mathbb{E}_t\left(\frac{p_{t-1}}{p_t}\right)^j \exp(\sum_{i=0}^{j-1}(i\tau-(i+1))\,\omega_{t+j}+j\tau\omega_t) \\
&= \left(\frac{p_{t-1}}{p_t}\right)^j \mathbb{E}_t\exp(\sum_{i=0}^{j-1}(i\tau-(i+1))\,\omega_{t+j}+j\tau\omega_t)
\end{aligned}$$

Note that in the exponent we have the sum of $j-1$ independent normals with zero mean and variance $(i\tau-(i+1))^2\,\sigma_\omega^2$ plus a constant $j\tau\omega_t$. Then, we have a normal random variable with mean $j\tau\omega_t$ and variance:

$$\sum_{i=0}^{j-1}(i\tau-(i+1))^2\,\sigma_\omega^2$$

Hence, the exponent has a lognormal distribution with expectation:

$$\exp\left(j\tau\omega_t+\frac{1}{2}\sigma_\omega^2\sum_{i=0}^{j-1}(i\tau-(i+1))^2\right)$$

and:

$$\mathbb{E}_t\frac{p_t}{p_{t+j}} = \left(\frac{p_{t-1}}{p_t}\right)^j \exp\left(j\tau\omega_t+\frac{1}{2}\sigma_\omega^2\sum_{i=0}^{j-1}(i\tau-(i+1))^2\right)$$

8.2. Computation

The estimation of the MA process for inflation is undertaken with Matlab's `garchset` function.

The estimation of the model was done with mixed-programming as follows. `Mathematica` computed the analytical derivatives of the value function of the representative household and generated Fortran 95 code with expressions. The derivatives depend on the parameters as symbolic variables. Then, we link the output into a Fortran 95 code that evaluates the solution of the model for each parameter value as implied by the maximization algorithm or by the Metropolis-Hastings algorithm The Fortran 95 code was compiled in Intel Visual Fortran 10.3 to run on Windows-based machines. We use a Xeon Processor 5160 EMT64 at 3.00 GHz with 16 GB of RAM.

References

[1] Arulampalam, A.S., S. Maskell, N. Gordon, and T. Clapp (2002). "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking." *IEEE Transactions on Signal Processing* 50, 174-188.

[2] Backus, D.K., B.R. Routledge, and S.E. Zin (2004). "Exotic Preferences for Macroeconomists." *NBER Macroeconomics Annual* 2004, 319-390.

[3] Backus, D.K., B.R. Routledge, and S.E. Zin (2004). "Asset Pricing in Business Cycle Analysis." *Mimeo*, New York University.

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

[5] Barro, R.J. (2005). "Rare Disasters and Asset Markets in the Twentieth Century." *Mimeo*, Harvard University.

[6] Benigno, P. and M. Woodford (2006). "Linear-Quadratic Approximation of Optimal Policy Problems." *Mimeo*, Columbia University.

[7] Caldara, D., J. Fernández-Villaverde, J.F. Rubio-Ramírez, and W. Yao (2008). "Computing Models with Recursive Preferences." *Mimeo*, University of Pennsylvania.

[8] Campaneli, C., R. Castro, and G. Clementi (2007). "Asset Pricing in a Production Economy with Chew-Dekel Preferences." *Working Papers* 07-13, New York University, Leonard N. Stern School of Business, Department of Economics.

[9] Campbell, J.Y. (1993). "Intertemporal Asset Pricing without Consumption Data." *American Economic Review* 83, 487-512.

[10] Campbell, J.Y. (1996). "Understanding Risk and Return." *Journal of Political Economy* 104, 298-345.

[11] Campbell, J.Y. (2003). "Consumption-Based Asset Pricing." In G.M. Constantinides, M. Harris, and R.M. Stulz (eds). *Handbook of the Economics of Finance*, Vol. 1B, North-Holland.

[12] Cho, J-O. and T.C. Cooley (2000). "Business Cycle Uncertainty and Economic Welfare." *Mimeo*, New York University.

[13] Cochrane, J.H. (2008). "Financial Markets and the Real Economy." in R. Mehra (ed), *Handbook of the Equity Premium Puzzle*, North-Holland.

[14] Croce, M. (2006). "Welfare Costs and Long-Run Consumption Risk in a Production Economy." *Mimeo*, University of North Carolina.

[15] Del Moral P., and J. Jacod (2002), "The Monte-Carlo Method for Filtering with Discrete Time Observations. Central Limit Theorems." in T. J. Lyons and T. S. Salisbury (eds), *Numerical Methods and Stochastics.* The Fields Institute Communication*s,* American Mathematical Society.

[16] Dolmas, J. (2007). "Real Business Cycle Dynamics under First Order Risk Aversion." *Mimeo*, Federal Reserve Bank of Dallas.

[17] Doucet A., N. de Freitas and N. Gordon (2001). *Sequential Monte Carlo Methods in Practice*. Springer Verlag.

[18] Epstein, L., and S.E. Zin (1989). "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: a Theoretical Framework." *Econometrica* 57, 937-969.

[19] Epstein, L., and S.E. Zin (1991). "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: an Empirical Analysis." *Journal of Political Economy* 99, 263-286.

[20] Fernández-Villaverde, J. and J.F. Rubio-Ramírez (2005). "Estimating Dynamic Equilibrium Economies: Linear versus Nonlinear Likelihood." *Journal of Applied Econometrics,* 20, 891-910.

[21] Fernández-Villaverde, J. and J.F. Rubio-Ramírez (2007). "Estimating Macroeconomic Models: A Likelihood Approach." *Review of Economic Studies* 74, 1059-1087.

[22] Fernández-Villaverde, J. and J.F. Rubio-Ramírez (2008). "How Structural Are Structural Parameters?" *NBER Macroeconomics Annual* 2007, 83-137.

[23] Hall, R.E. (1988). "Intertemporal Substitution in Consumption." *Journal of Political Economy* 96, 339-357.

[24] Hansen, L.P., J. Heaton, and N. Li (2008). "Consumption Strikes Back?: Measuring Long Run Risk." *Mimeo*, University of Chicago.

[25] Hansen, L.P., and T.J. Sargent (1995). "Discounted Linear Exponential Quadratic Gaussian Control." *IEEE Transactions on Automatic Control* 40, 968-971.

[26] Hansen, L.P., T.J. Sargent, and T.D. Tallarini (1999). "Robust Permanent Income and Pricing." *Review of Economic Studies* 66, 873-907.

[27] Judd, K.L. (1998). *Numerical Methods in Economics.* MIT Press.

[28] Judd, K.L. and S.M. Guu (1992). "Perturbation Solution Methods for Economic Growth Models." In H. Varian (ed.), *Economic and Financial Modelling with Mathematica.* Springer Verlag.

[29] Judd, K.L. and S.M. Guu (1997). "Asymptotic Methods for Aggregate Growth Models." *Journal of Economic Dynamics and Control* 21, 1025-1042.

[30] Justiniano A. and G.E. Primiceri (2007). "The Time Varying Volatility of Macroeconomic Fluctuations." *American Economic Review*, forthcoming.

[31] Kaltenbrunner, G. and L. Lochstoer (2007). "Long-Run Risk through Consumption Smoothing." *Mimeo*, London Business School.

[32] Kimball, M.S. (1990). "Precautionary Saving in the Small and in the Large." *Econometrica* 58, 53-73.

[33] Kocherlakota, N.R. (1990a). "On the 'Discount' Factor in Growth Economies." *Journal of Monetary Economics* 25, 43-47.

[34] Kocherlakota, N.R. (1990b). "Disentangling the Coefficient of Relative Risk Aversion from the Elasticity of Intertemporal Substitution: An Irrelevance Result." *Journal of Finance* 45, 175-190.

[35] Kreps, D.M. and E.L. Porteus (1978). "Temporal Resolution of Uncertainty and Dynamic Choice Theory." *Econometrica* 46, 185-200.

[36] Künsch, H.R. (2005). "Recursive Monte Carlo Filters: Algorithms and Theoretical Analysis." *Annals of Statistics* 33, 1983-2021.

[37] Lettau, M. and S. Ludvigson (2001). "Consumption, Aggregate Wealth, and Expected Stock Returns." *Journal of Finance* 56, 815-849.

[38] Levin, A., D. López-Salido, and T. Yun (2006). "Risk-Sensitive Monetary Policy." *Mimeo*, Federal Reserve Board.

[39] Levine, P., J. Pearlman, and R. Pierse (2007). "Linear-Quadratic Approximation, External Habit, and Targeting Rules." *Mimeo*, University of Surrey.

[40] Lustig, H., S. Van Nieuwerburgh, and A. Verdehaln (2007). "The Wealth-Consumption Ratio: A Litmus Test for Consumption-based Asset Pricing Models." *Mimeo*, UCLA.

[41] McFadden, D.L. (1989). "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration." *Econometrica* 57, 995-1026.

[42] Pakes, A. and D. Pollard (1989). "Simulation and the Asymptotics of Optimization Estimators." *Econometrica* 57, 1027-1057.

[43] Piazzesi, M. (2003). "Affine Term Structure Models." *Mimeo*, University of Chicago.

[44] Piazzesi, M. and M. Schneider (2006). "Equilibrium Yield Curves." *NBER Macroeconomics Annual* 2006, 389-442.

[45] Pitt, M.K. (2002). "Smooth Particle Filters for Likelihood Evaluation and Maximization". *Mimeo*, University of Warwick.

[46] Pitt, M.K. and N. Shephard (1999). "Filtering via Simulation: Auxiliary Particle Filters." *Journal of the American Statistical Association* 94, 590-599.

[47] Restoy, F. and P. Weil (1998). "Approximate Equilibrium Asset Prices." *NBER Working Papers* 6611.

[48] Roberts, G., A. Gelman, and W. Gilks (1997). "Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorthims." *Annals of Applied Probability* 7, 110-120.

[49] Rouwenhost, K.G. (1993). "Asset Pricing Implications of Equilibrium Business Cycle Models." In Cooley, T. F. (Eds.), *Frontiers of Business Cycle Research*. Princeton University Press, Princeton, 293-330.

[50] Rubin, D.B. (1998). "Using the SIR Algorithm to Simulate Posterior Distributions". in J.M. Bernardo, M.H. DeGroot, D.V. Lindley, and A.F.M. Smith (eds), *Bayesian Statistics 3*, 395-402, Oxford University Press.

[51] Santos, M.S. and A. Peralta-Alva (2005). "Accuracy of Simulations for Stochastic Dynamic Models". *Econometrica* 73, 1939-1976.

[52] Schmitt-Grohé, S. and M. Uribe (2004). "Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function". *Journal of Economic Dynamics and Control* 28, 755-775.

[53] Schmitt-Grohé, S. and M. Uribe (2005). "Optimal Fiscal and Monetary Policy in a Medium Scale Macroeconomic Model". *NBER Macroeconomic Annual 2005*, 382-425.

[54] Sims, C.A. and T. Zha (2006). "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review* vol 96, 54-81.

[55] Stock, J.H. and M.W. Watson (2002). "Has the Business Cycle Changed, and Why?" *NBER Macroeconomics Annual* 2002, 159-218.

[56] Stock, J.H. and M.W. Watson (2007). "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking* 39, 3-33.

[57] Tallarini., T.D. (2000). "Risk-Sensitive Real Business Cycles." *Journal of Monetary Economics* 45, 507-532.

[58] Uhlig, H. (2007). "Leisure, Growth, and Long Run Risk." *Mimeo*, University of Chicago.

[59] Wang, S. (1993). "The Local Recoverability of Risk Aversion and Intertemporal Substitution." *Journal of Economic Theory* 59, 333-363.

[60] Weil, P. (1990). "Nonexpected Utility in Macroeconomics." *Quarterly Journal of Economics* 105, 29-42.

Figure 4.1: Capital

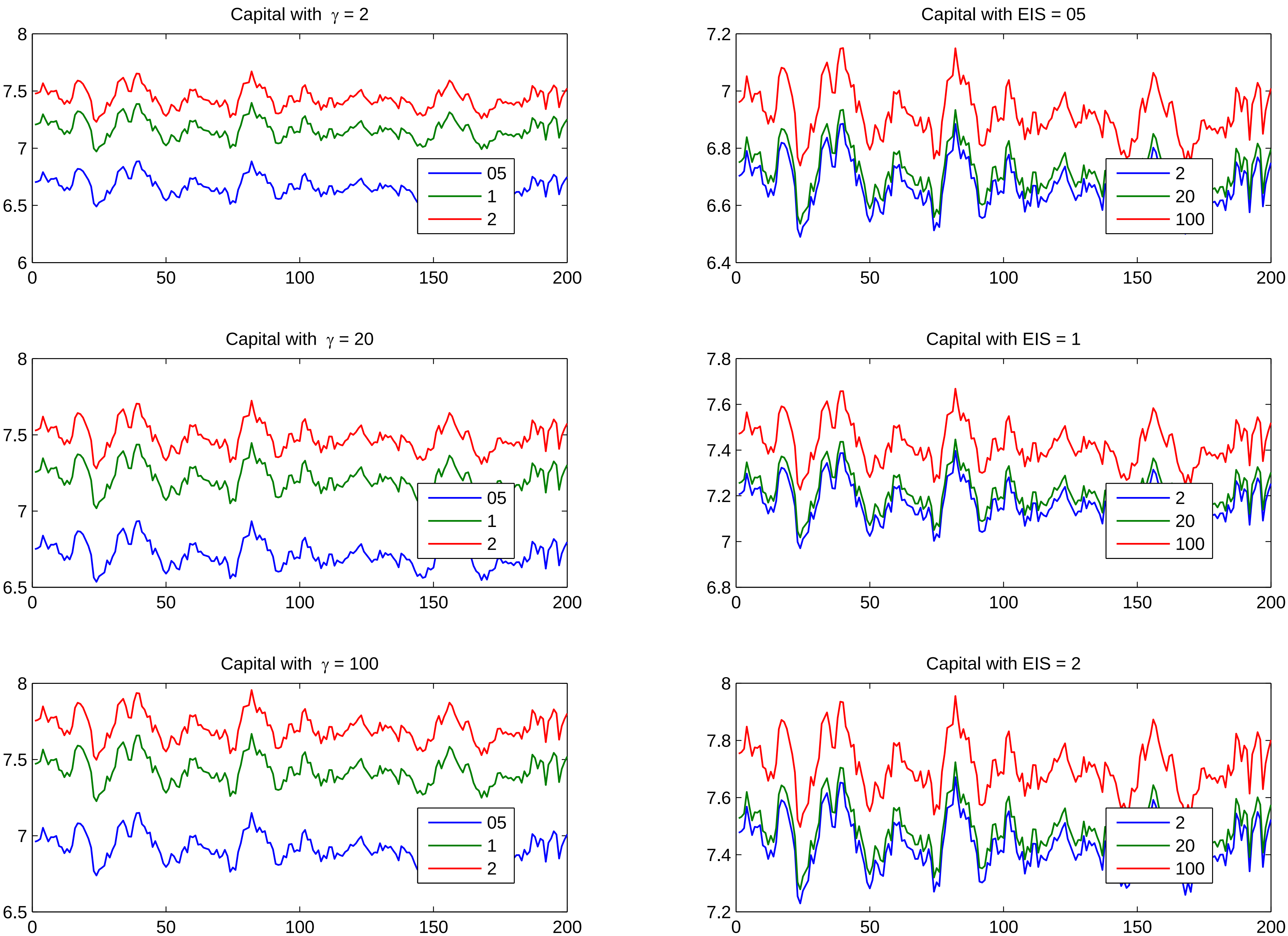


Figure 4.2: Capital, re-centered

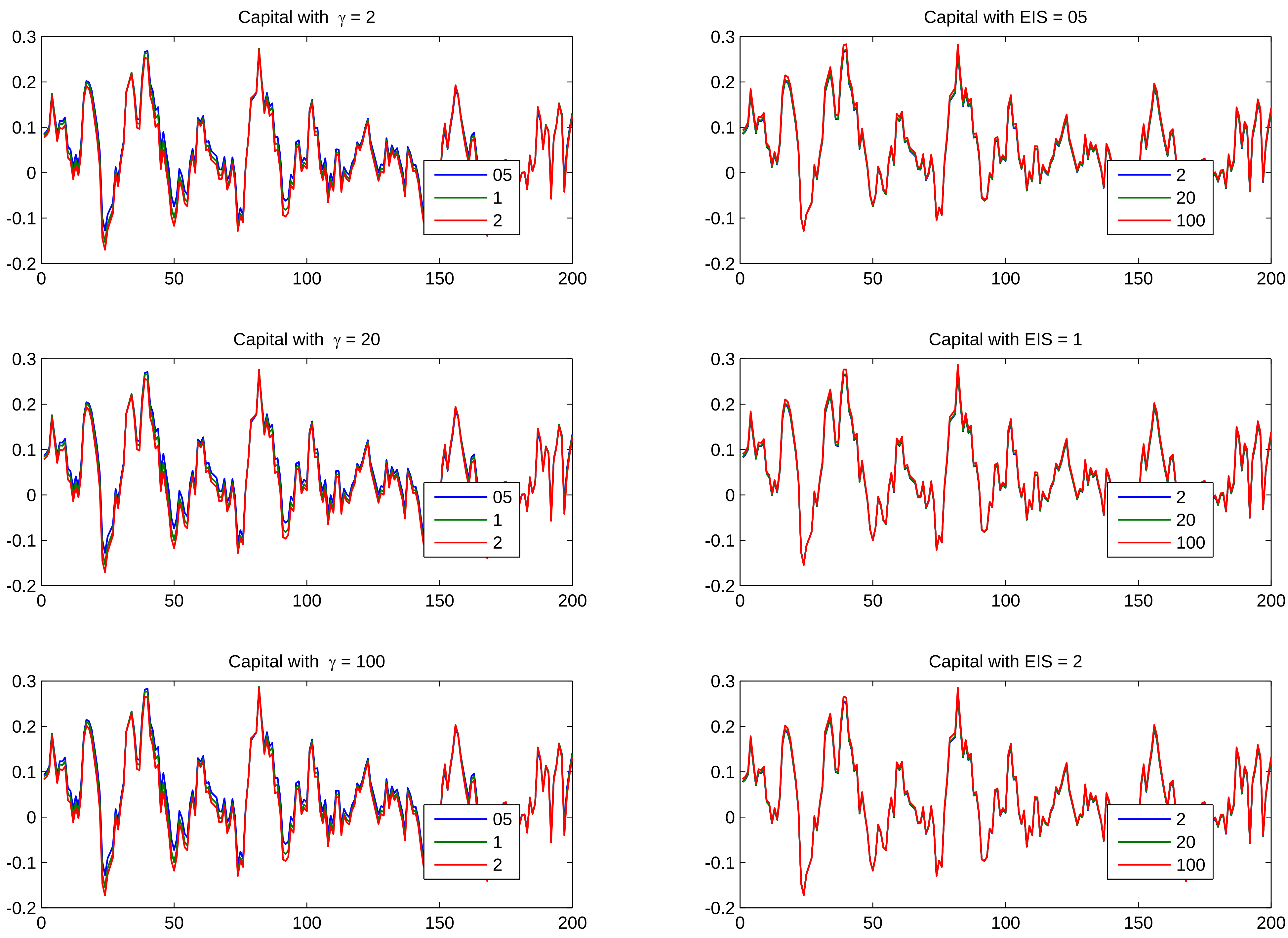


Figure 4.3: Output

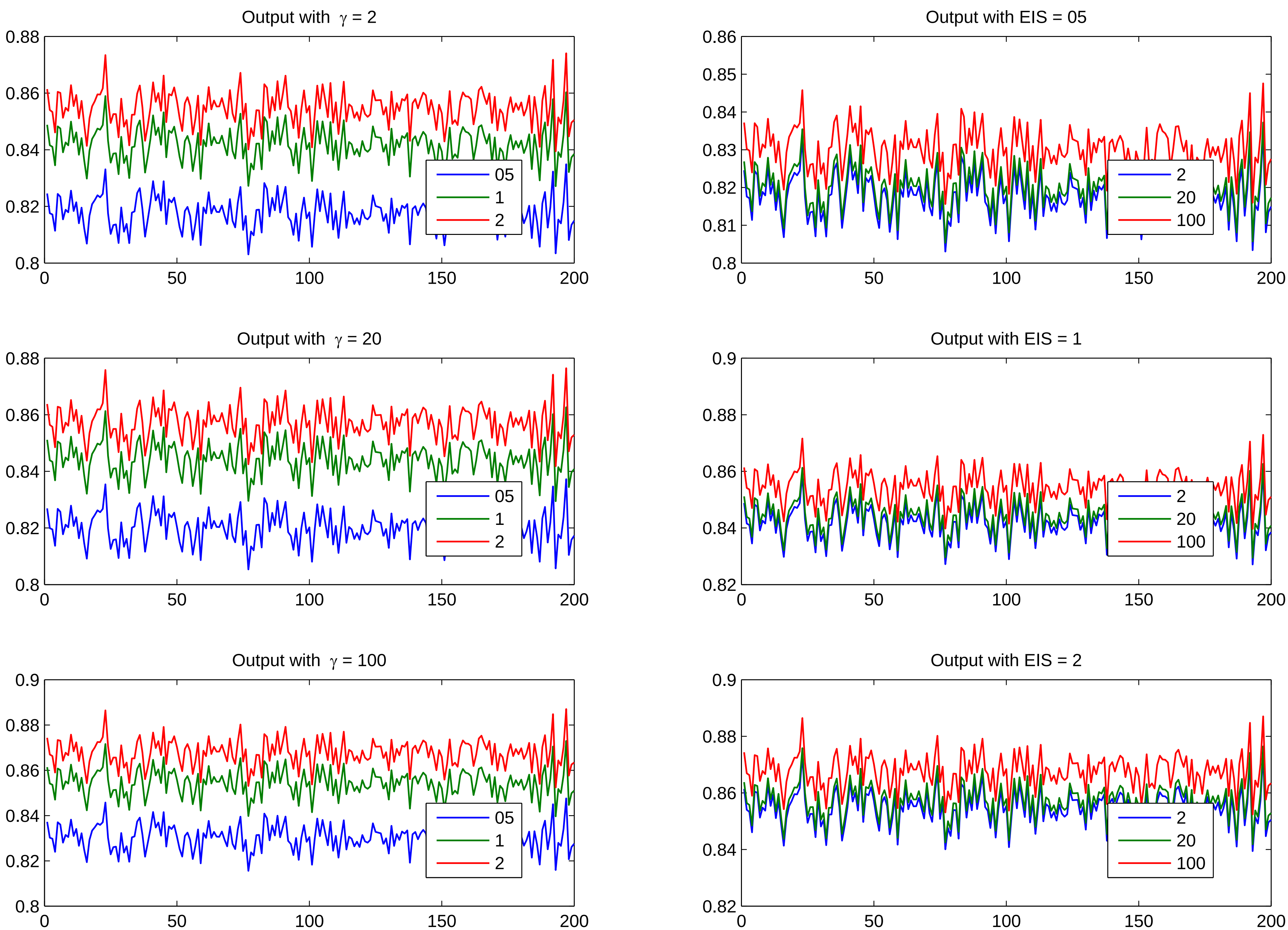


Figure 4.4: Output, re-centered

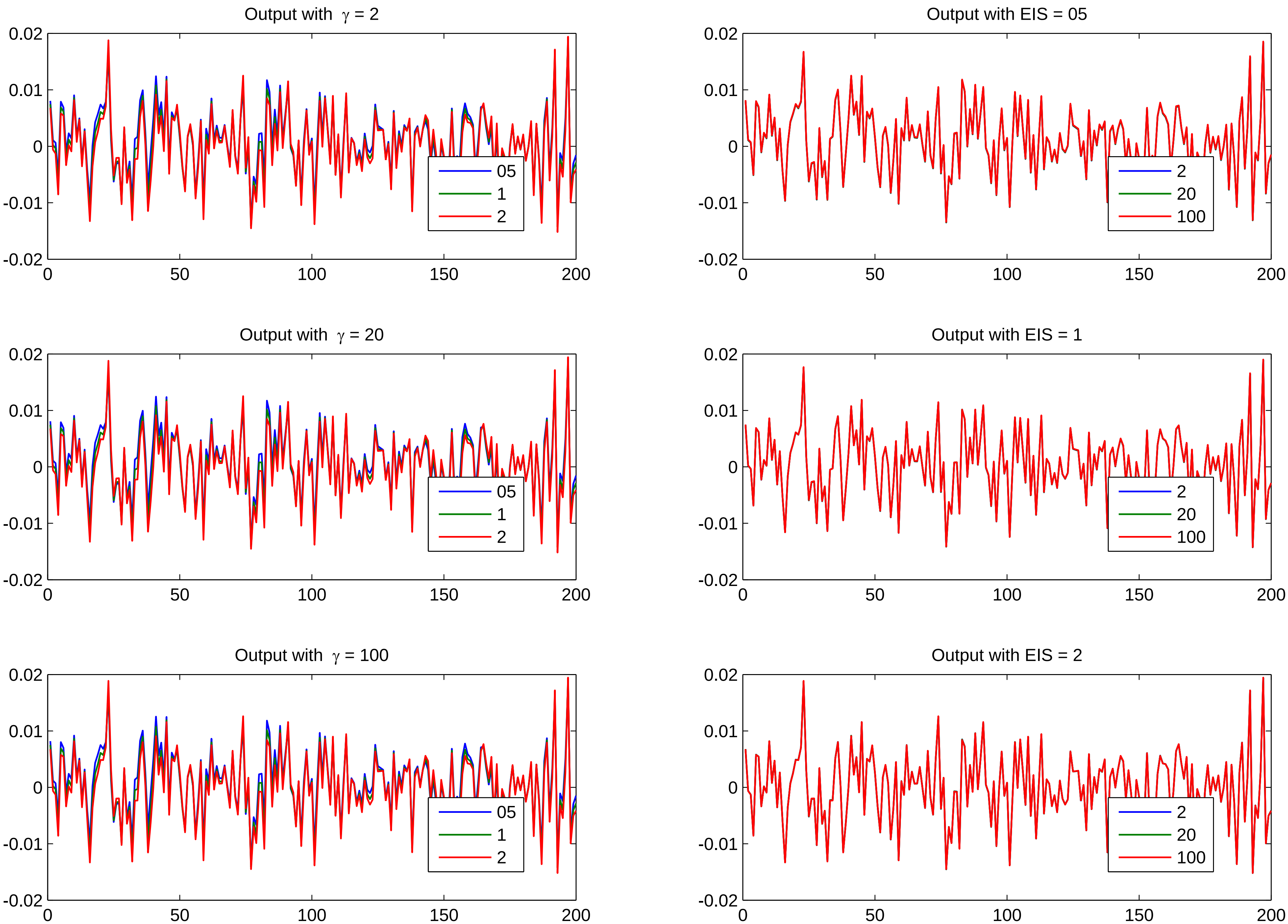


Figure 4.5: Price 1 Quarter Bond

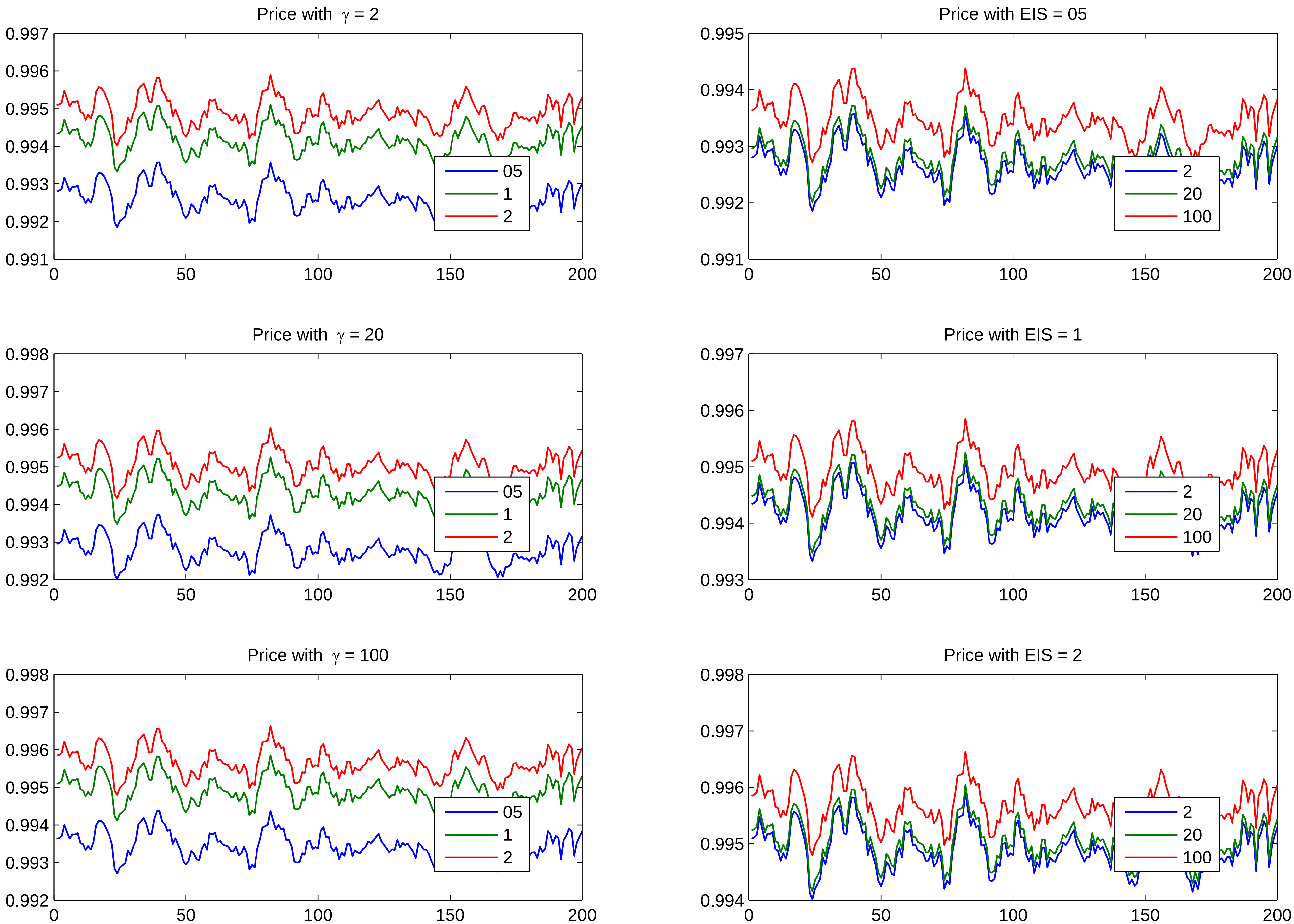


Figure 4.6: Price 1 Quarter Bond, re-centered

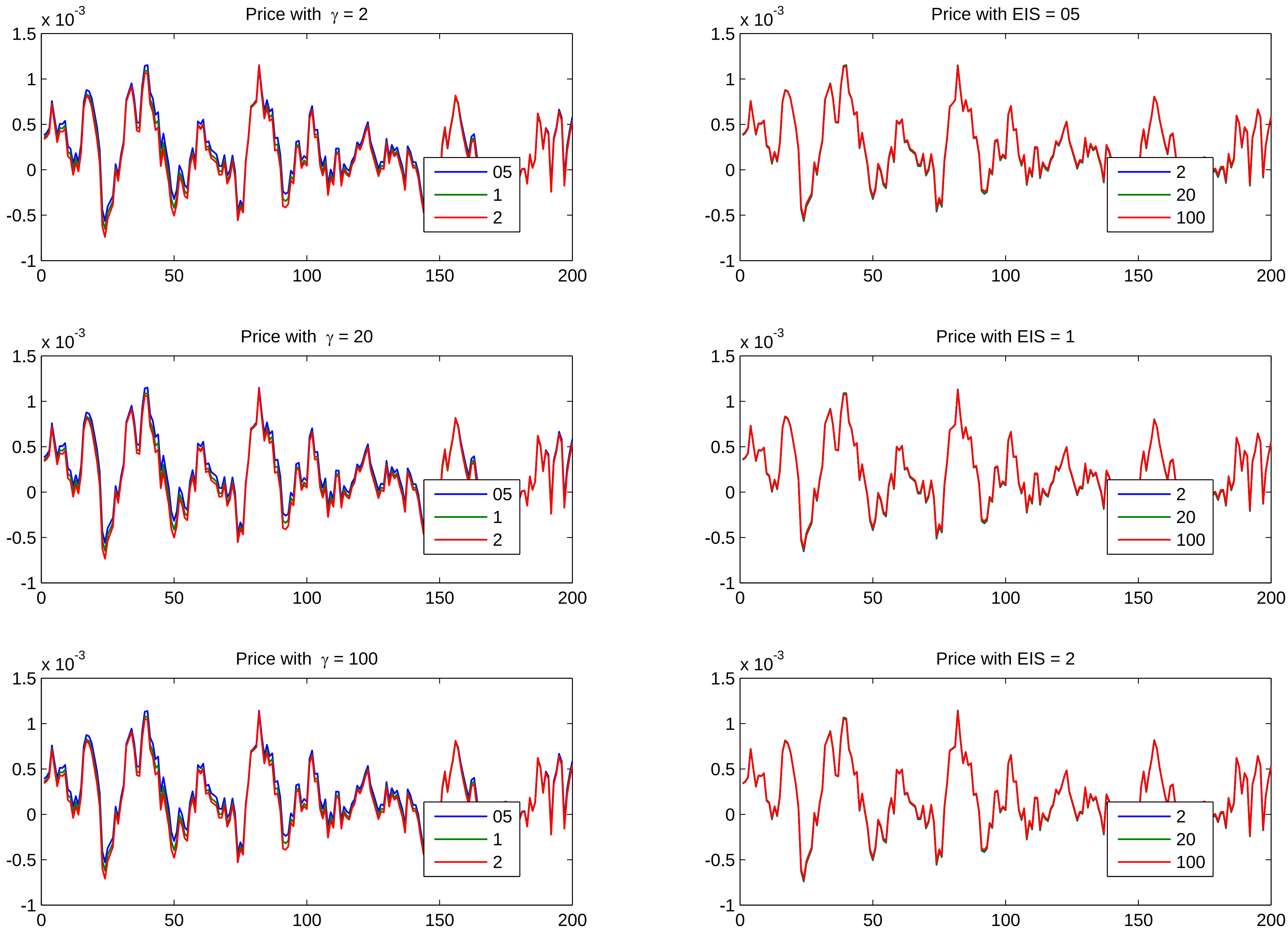


Figure 4.7: Consumption

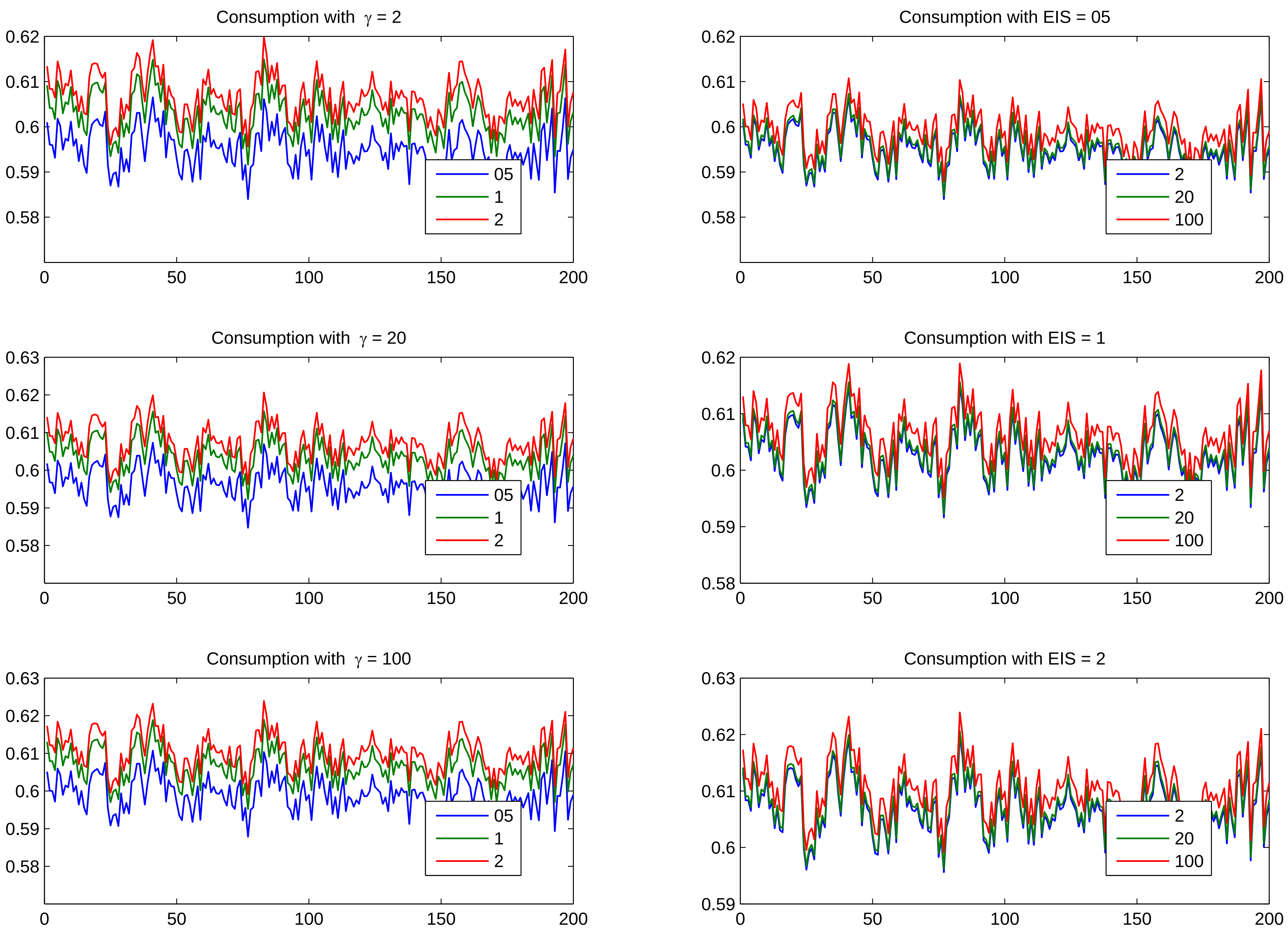


Figure 4.8: Consumption, re-centered

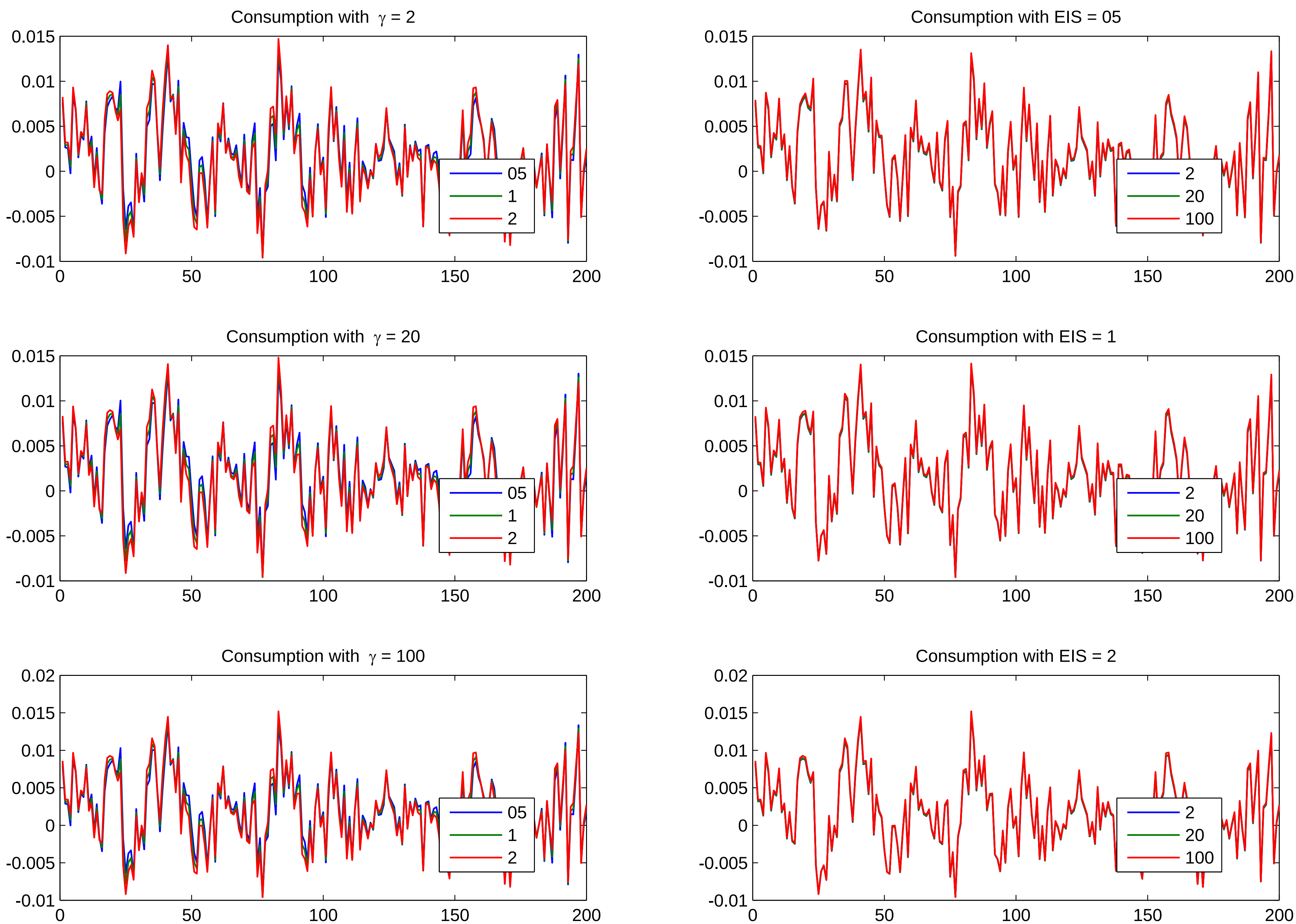


Figure 4.9: Labor

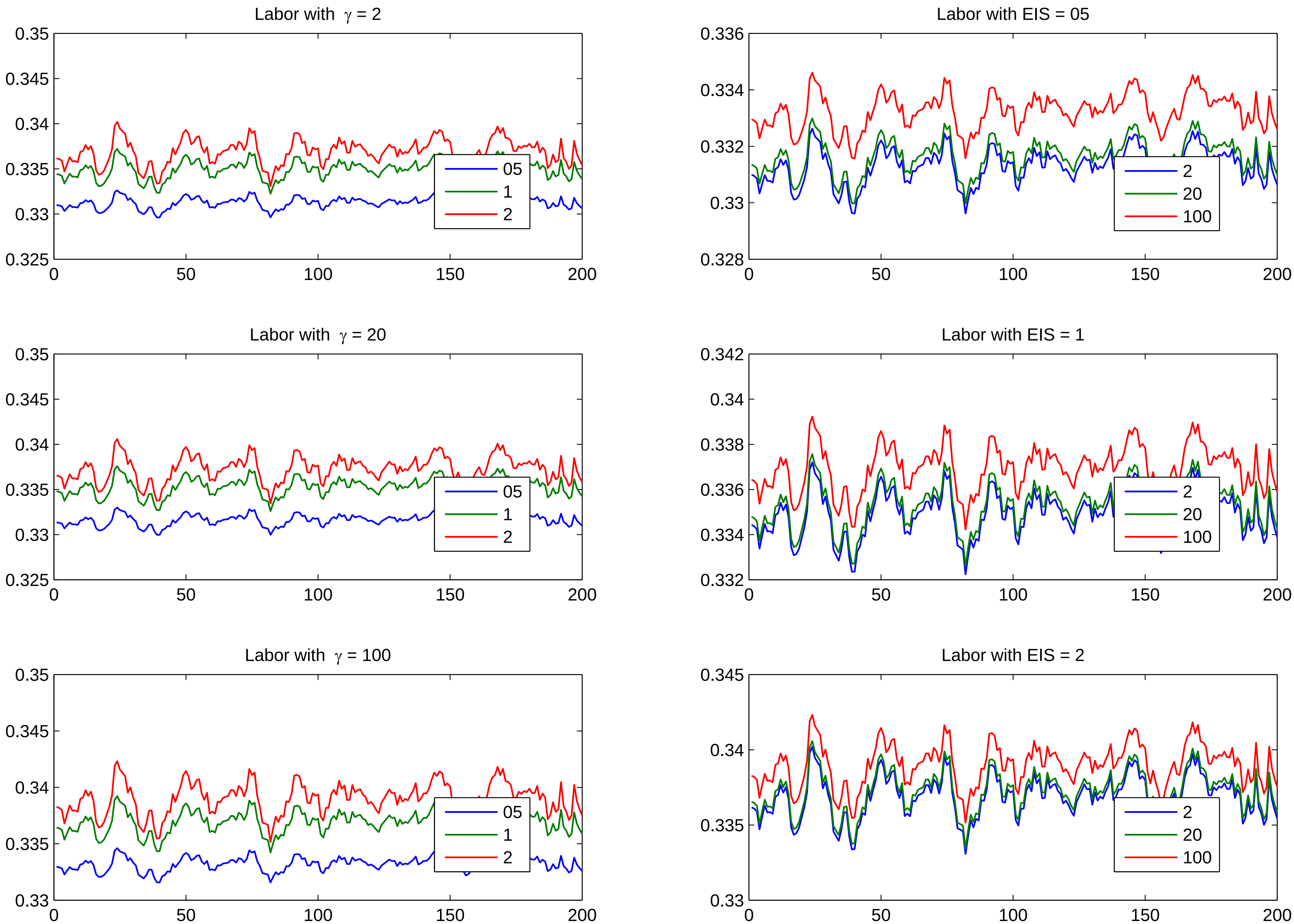


Figure 4.10: Labor, re-centered

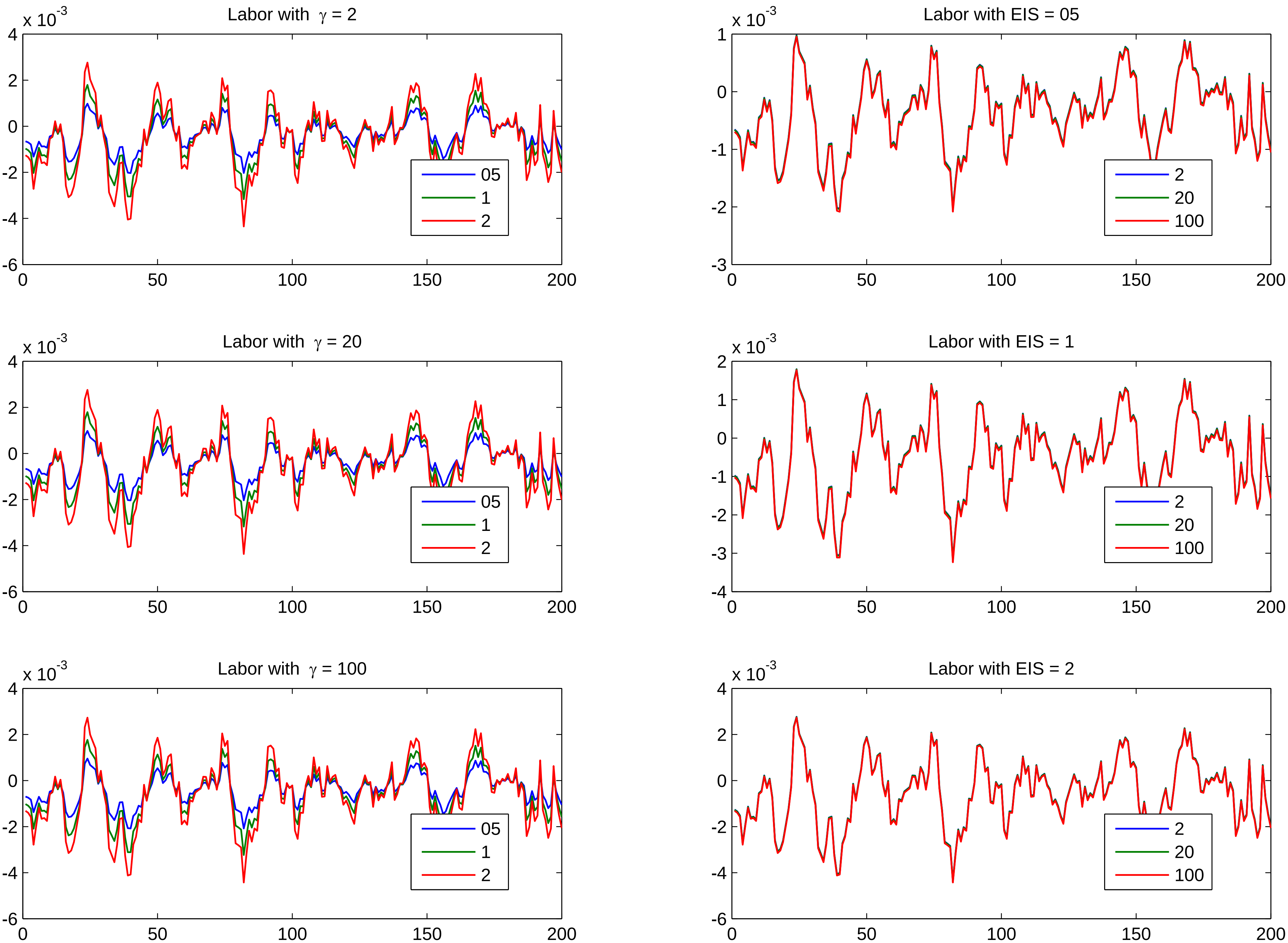


Figure 4.11: Dividends

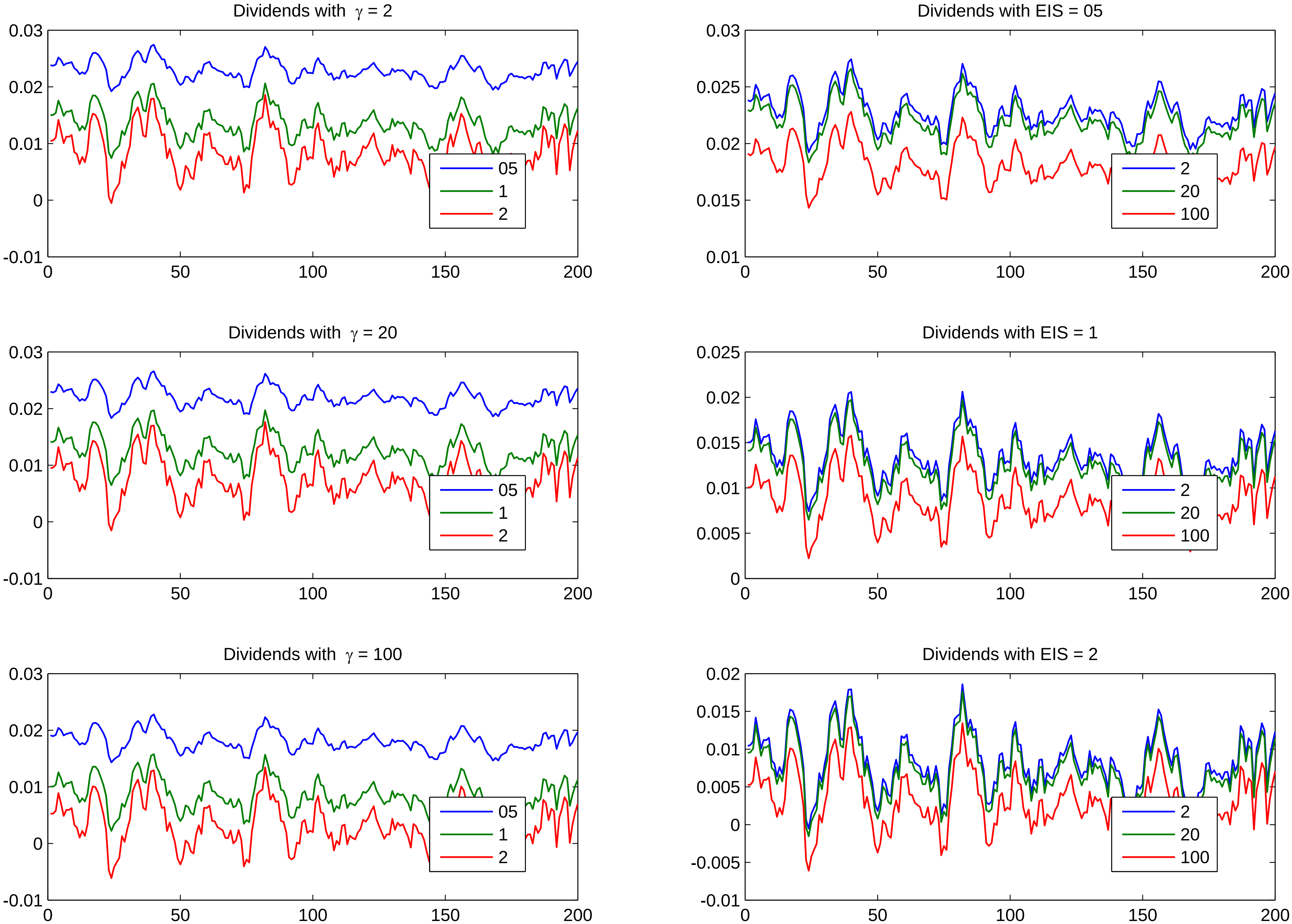


Figure 4.12: Dividends, re-centered

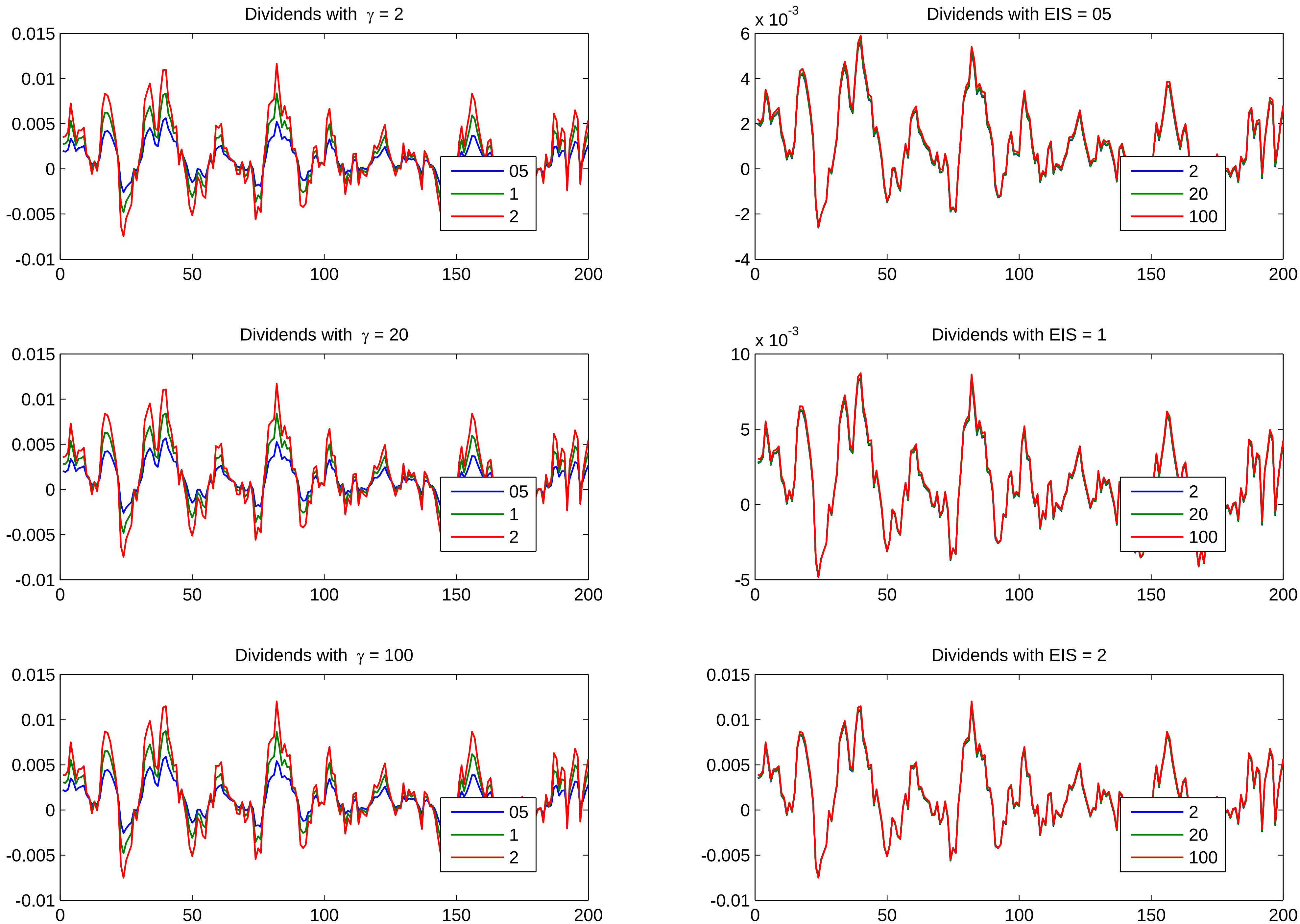


Figure 4.13: Yield Curves, Steady State

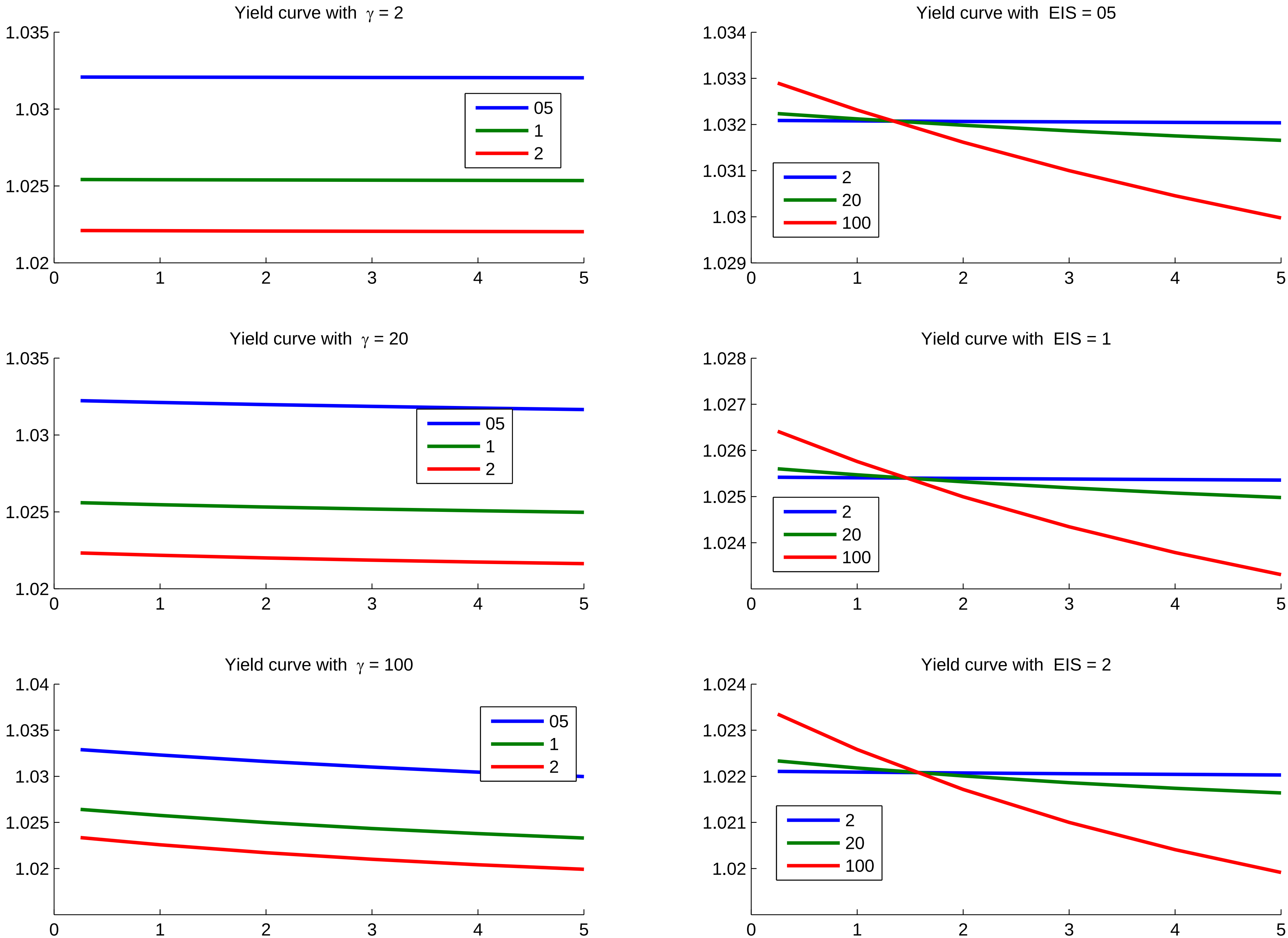


Figure 4.14: Yield Curves, Mean Ergodic Distribution

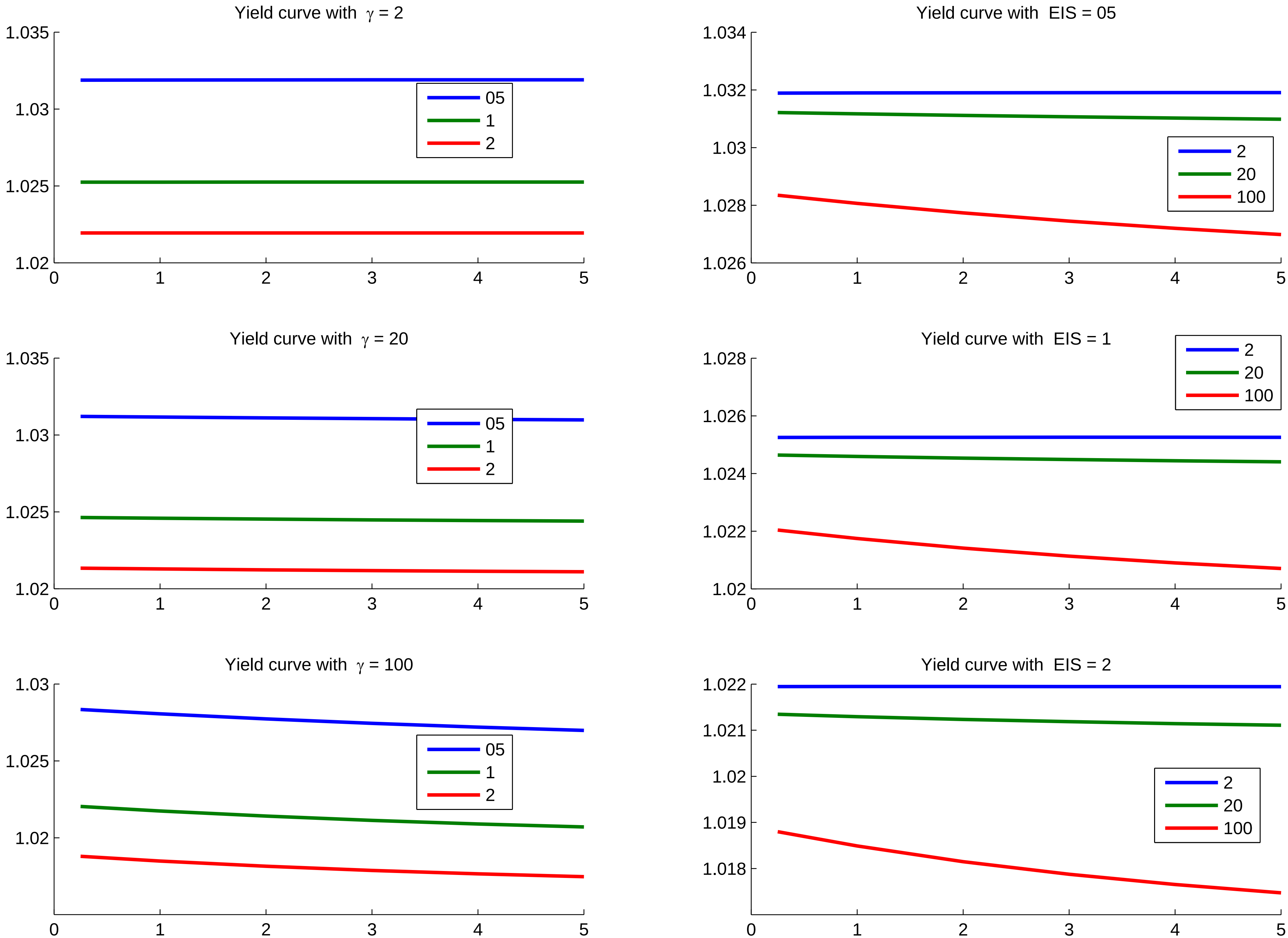
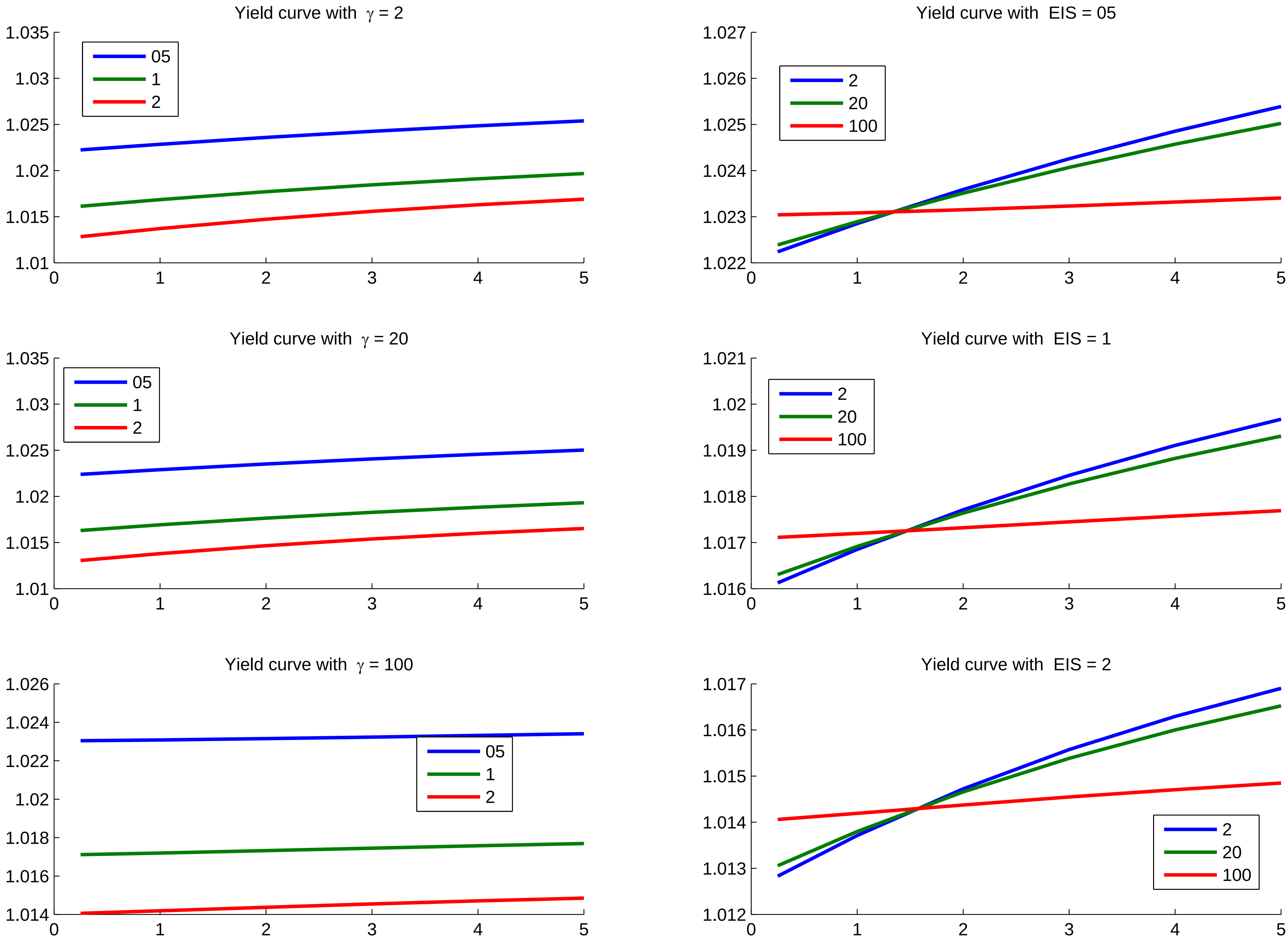
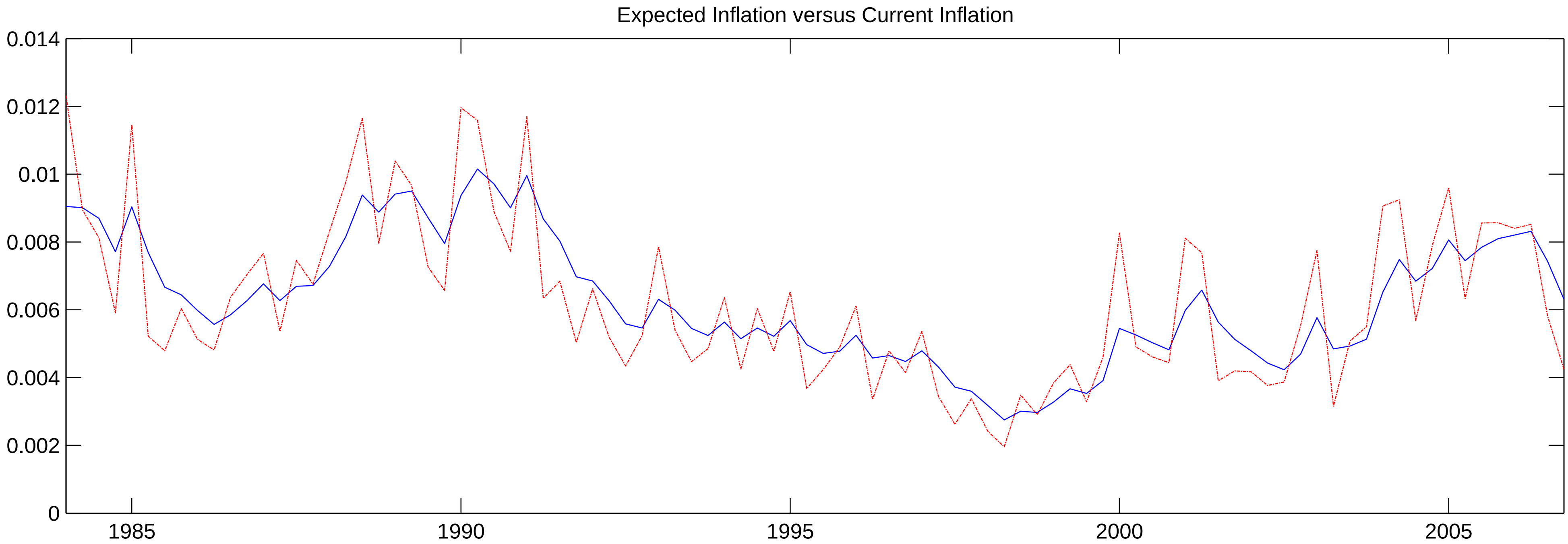
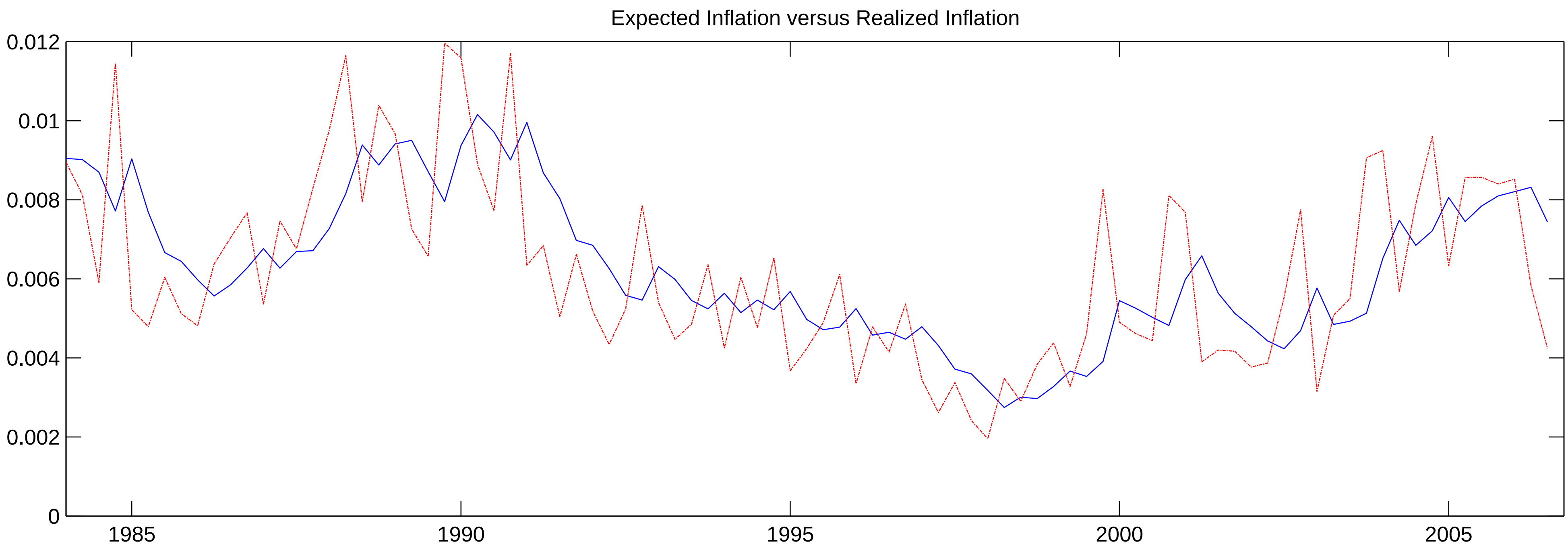
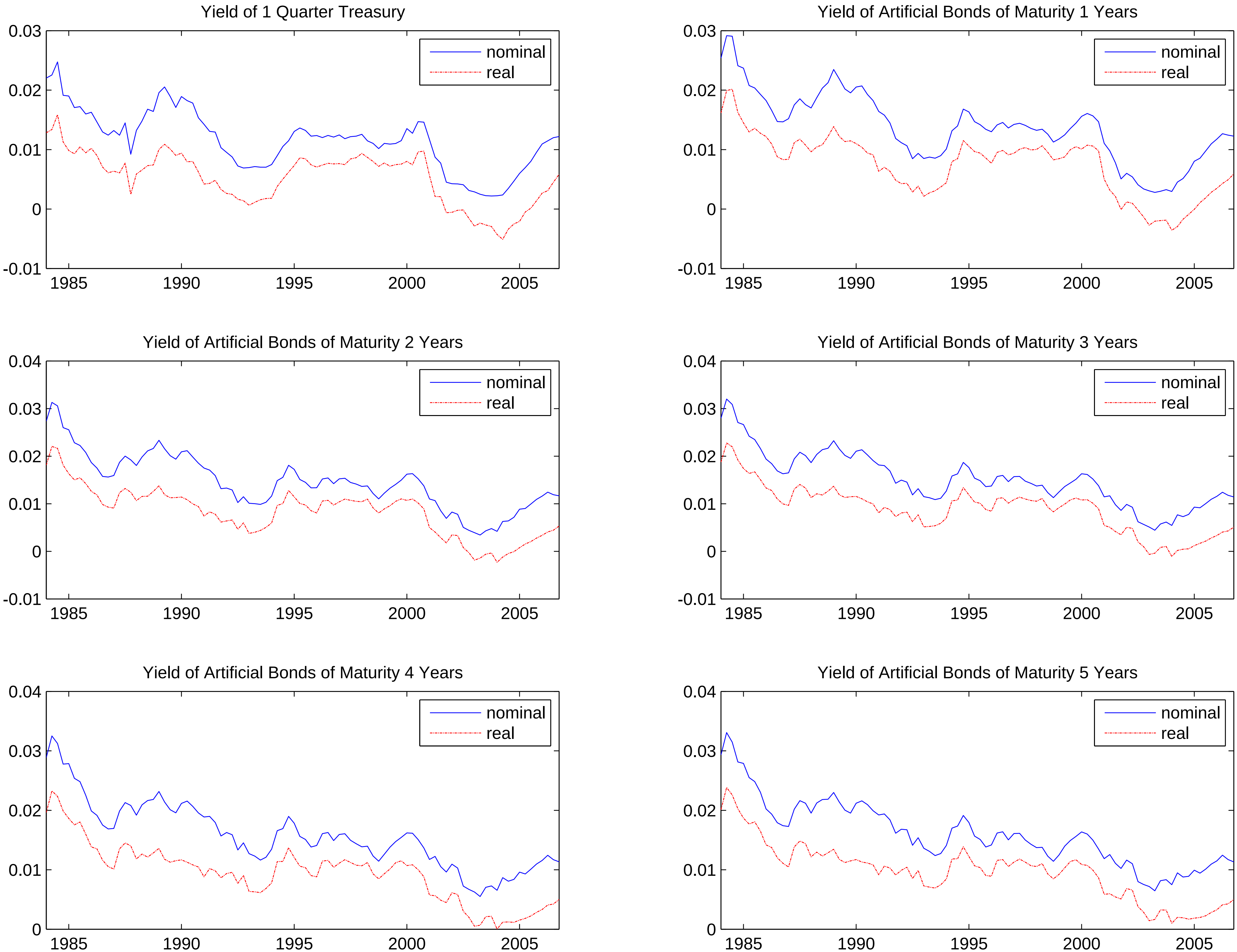


Figure 4.15: Yield Curves, Capital 0.6 over Steady State

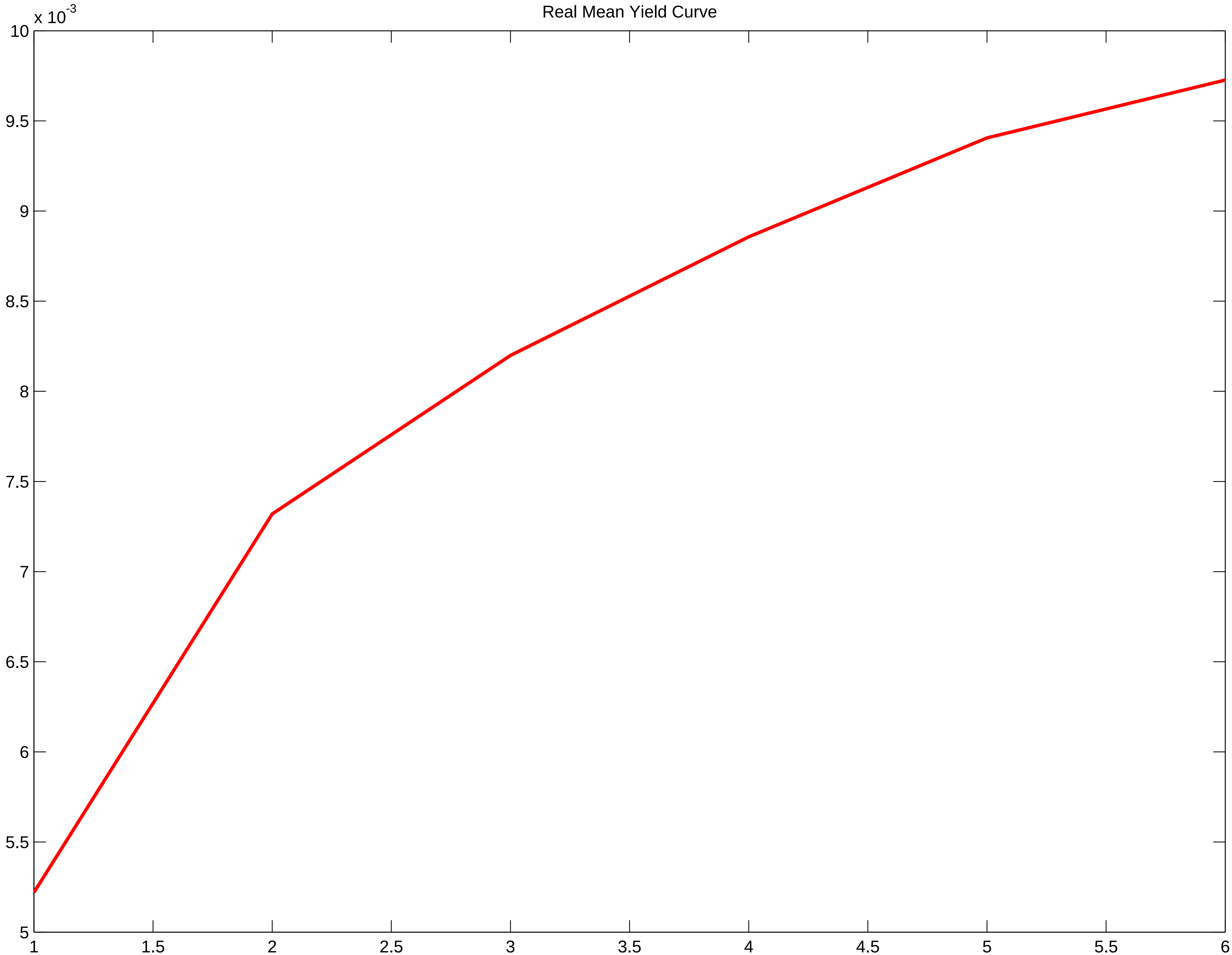



Expected Inflation versus Current Inflation
0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
1985
1990
1995
2000
2005


Expected Inflation versus Realized Inflation
0.012
0.01
0.008
0.006
0.004
0.002
0
1985
1990
1995
2000
2005

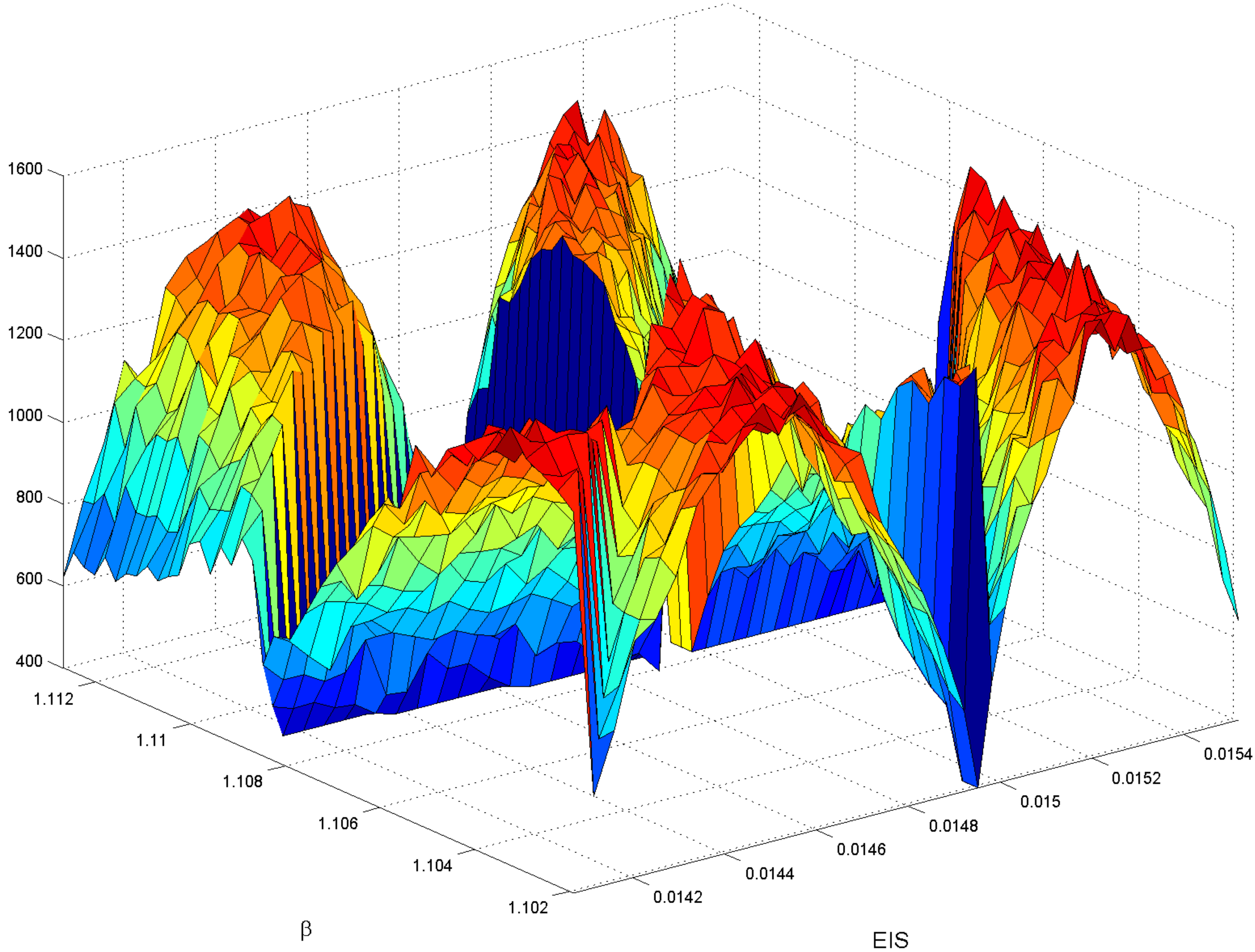


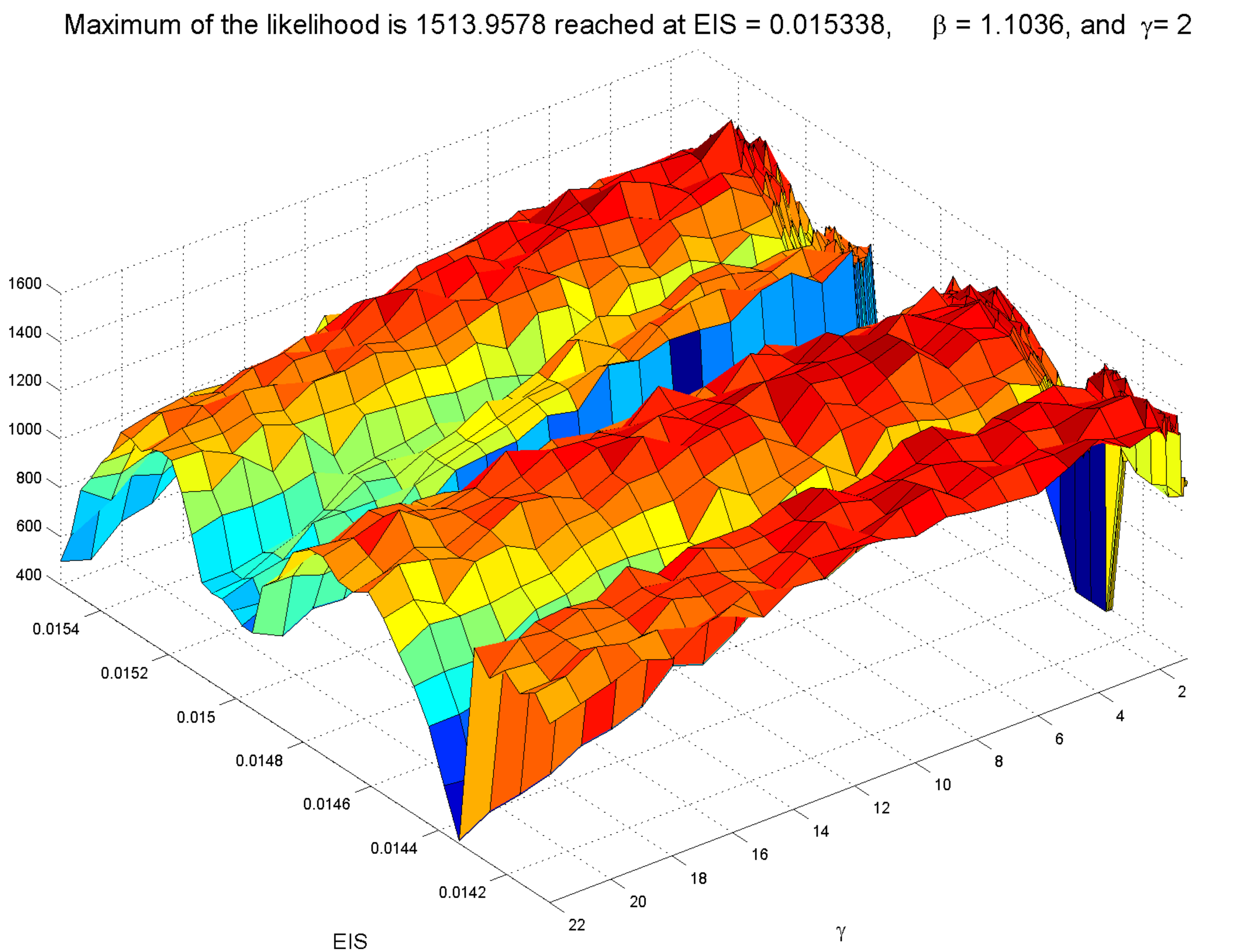
Yield of 1 Quarter Treasury
0.03
0.02
0.01
0
-0.01
1985
1990
1995
2000
2005
nominal
real
Yield of Artificial Bonds of Maturity 1 Years
0.03
0.02
0.01
0
-0.01
1985
1990
1995
2000
2005
nominal
real
Yield of Artificial Bonds of Maturity 2 Years
0.04
0.03
0.02
0.01
0
-0.01
1985
1990
1995
2000
2005
nominal
real
Yield of Artificial Bonds of Maturity 3 Years
0.04
0.03
0.02
0.01
0
-0.01
1985
1990
1995
2000
2005
nominal
real
Yield of Artificial Bonds of Maturity 4 Years
0.04
0.03
0.02
0.01
0
1985
1990
1995
2000
2005
nominal
real
Yield of Artificial Bonds of Maturity 5 Years
0.04
0.03
0.02
0.01
0
1985
1990
1995
2000
2005
nominal
real



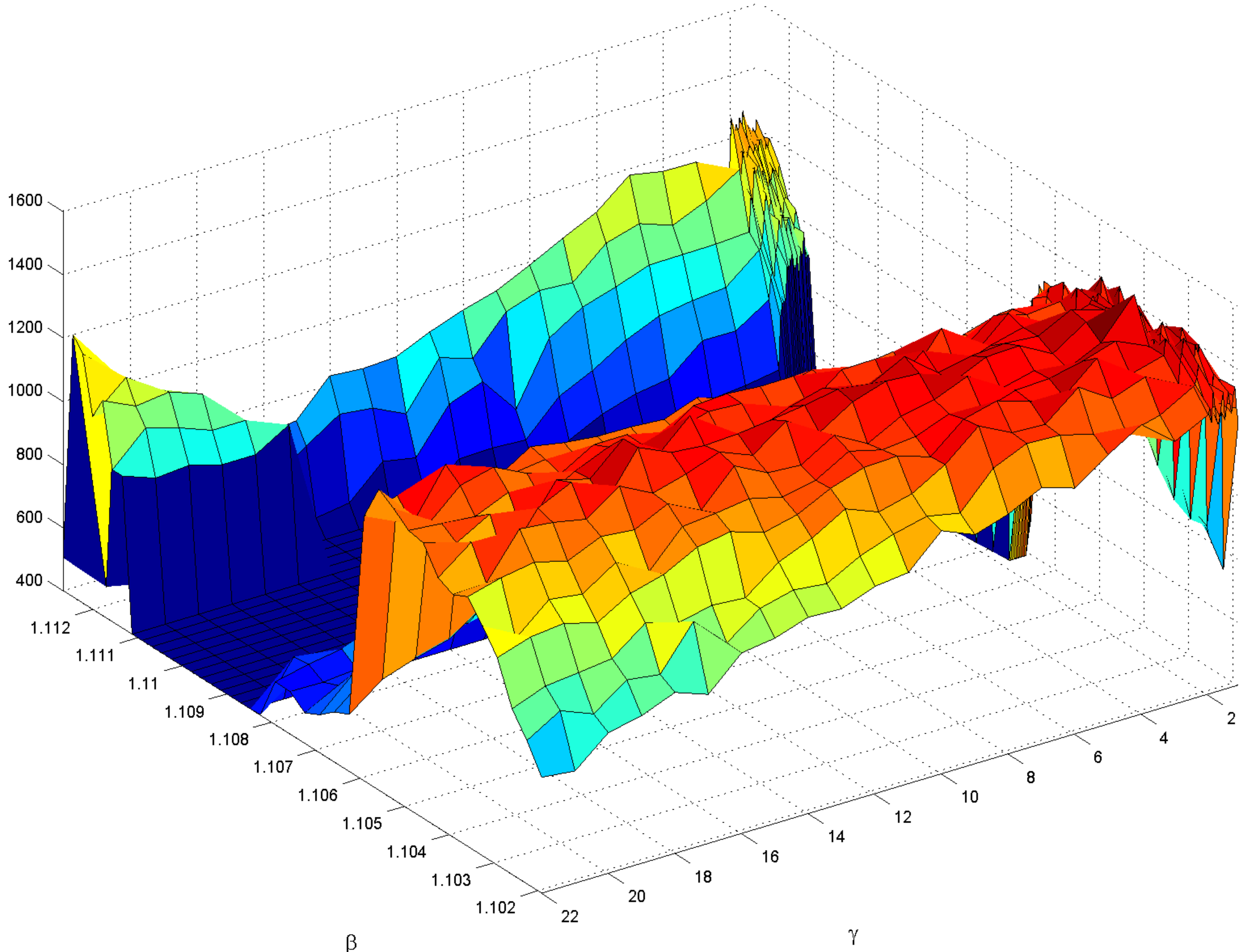
Real Mean Yield Curve
x 10^-3
10
9.5
9
8.5
8
7.5
7
6.5
6
5.5
5
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
6

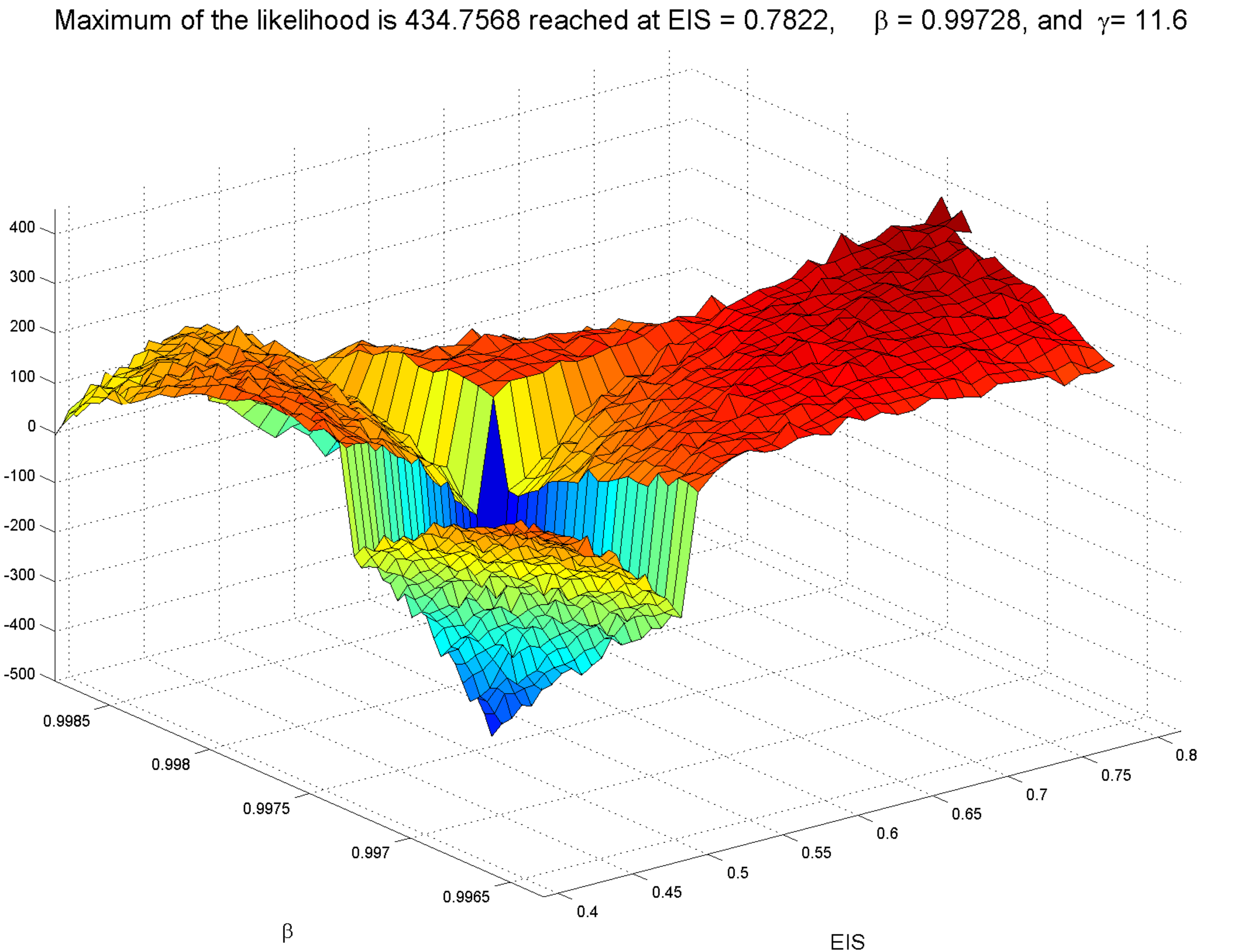
Maximum of the likelihood is 1513.9578 reached at EIS = 0.015338, β = 1.1036, and γ= 2




Maximum of the likelihood is 1513.9578 reached at EIS = 0.015338, β = 1.1036, and γ= 2
1600
1400
1200
1000
800
600
400
0.0154
0.0152
0.015
0.0148
0.0146
0.0144
0.0142
EIS
2
4
6
8
10
12
14
16
18
20
22
γ

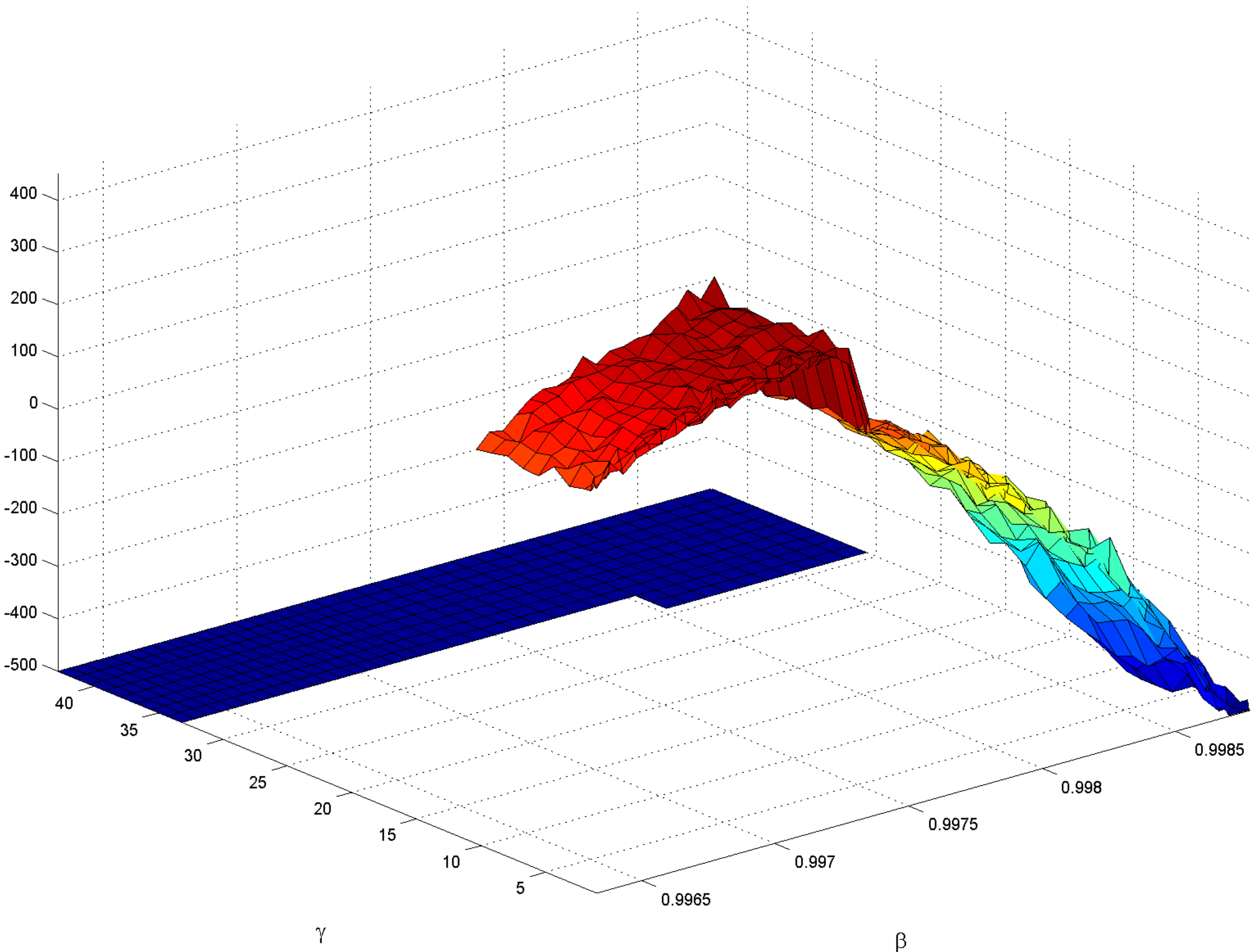
Maximum of the likelihood is 1513.9578 reached at EIS = 0.015338, β = 1.1036, and γ= 2

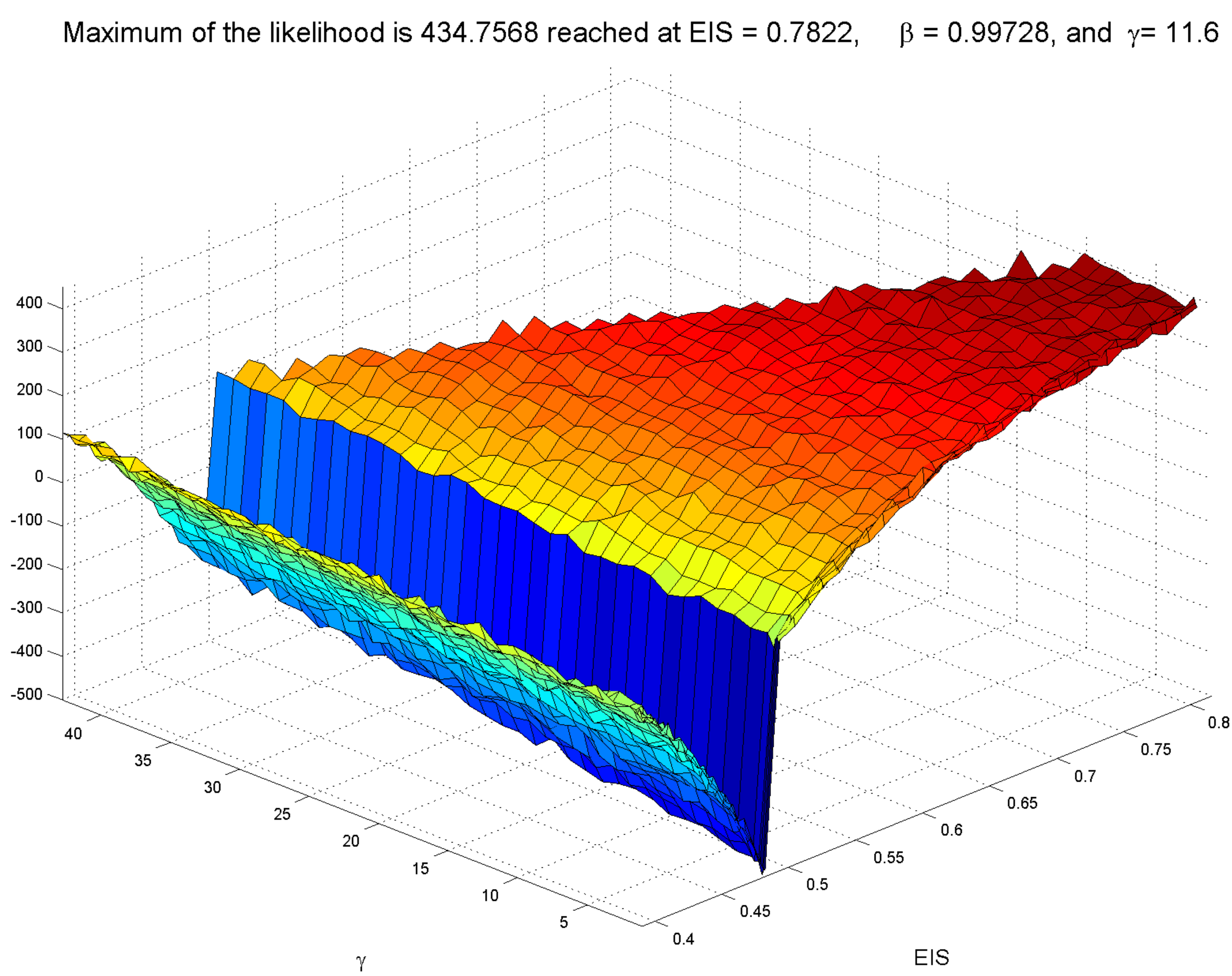




Maximum of the likelihood is 434.7568 reached at EIS = 0.7822, β = 0.99728, and γ= 11.6
400
300
200
100
0
-100
-200
-300
-400
-500
0.9985
0.998
0.9975
0.997
0.9965
β
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
EIS

Maximum of the likelihood is 434.7568 reached at EIS = 0.7822, β = 0.99728, and γ= 11.6





Maximum of the likelihood is 434.7568 reached at EIS = 0.7822, β = 0.99728, and γ= 11.6
400
300
200
100
0
-100
-200
-300
-400
-500
40
35
30
25
20
15
10
5
γ
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
EIS