PIER Working Paper 10-011

"The Term Structure of Interest Rates in a DSGE Model with Recursive Preferences"

by

Jules H. van Binsbergen, Jesús Fernández-Villaverde,
Ralph S.J. Koijen, Juan F. Rubio-Ramírez

http://ssrn.com/abstract=1585848

The Term Structure of Interest Rates

in a DSGE Model with Recursive Preferences[*]

Jules H. van Binsbergen   **Jesús Fernández-Villaverde**   **Ralph S.J. Koijen**

Stanford University   University of Pennsylvania   University of Chicago

Graduate School of Business   FEDEA, NBER and CEPR   Booth School of Business

Juan F. Rubio-Ramírez

Duke University

Federal Reserve Bank of Atlanta

FEDEA

March 2010

Abstract

We solve a dynamic stochastic general equilibrium (DSGE) model in which the representative household has Epstein and Zin recursive preferences. The parameters governing preferences and technology are estimated by means of maximum likelihood using macroeconomic data and asset prices, with a particular focus on the term structure of interest rates. We estimate a large risk aversion, an elasticity of intertemporal substitution higher than one, and substantial adjustment costs. Furthermore, we identify the tensions within the model by estimating it on subsets of these data. We conclude by pointing out potential extensions that might improve the model's fit.

1. Introduction

In this paper, we study whether an estimated dynamic stochastic general equilibrium (DSGE) model in which the representative household has Epstein and Zin recursive preferences can match both macroeconomic and yield curve data. After solving the model using perturbation methods, we build the likelihood function with the particle filter and estimate the structural parameters, those describing preferences and technology, via maximum likelihood and macroeconomic and yield curve data. We also estimate the model on subsets of the data to illustrate how the structural parameters are identified.

The motivation for our exercise is that economists are paying increasing attention to recursive utility functions (Kreps and Porteus, 1978, Epstein and Zin, 1989 and 1991, and Weil, 1990).[1] The key advantage of these preferences is that they allow separation between the intertemporal elasticity of substitution (IES) and risk aversion. In the asset pricing literature, researchers have argued that Epstein and Zin preferences account for many patterns in the data, possibly in combination with other features such as long-run risk. Bansal and Yaron (2004) is a prime representative of this line of work. From a policy perspective, recursive preferences generate radically bigger welfare costs of the business cycle than those coming from standard expected utility (Tallarini, 2000). Hence, they may change the trade-offs that policy makers face, as shown by Levin, López-Salido, and Yun (2007). Finally, Epstein and Zin preferences can be reinterpreted, under certain conditions, as a case of robust control preferences (Hansen, Sargent, and Tallarini, 1999). The natural question is, thus, whether we can successfully take an empirically plausible version of these models to the data.

Our paper makes three main contributions. The first contribution is to study the role of Epstein and Zin preferences in a full-fledged production economy with endogenous capital and labor supply and their interaction with the yield curve. Having a production economy is crucial for many questions in which economists are interested. For example, one of the most attractive promises of integrating macroeconomics and finance is to have, in the middle run, richer models for policy advice. Fiscal or monetary policy will have implications for the yield curve because they trigger endogenous responses on the accumulation of capital. These effects on the yield curve may be key for the propagation mechanism of policy. Similarly, we want to learn how to interpret movements in the yield curve as a way to identify the effects of

---

[1] Among many others, Backus, Routledge, and Zin (2004 and 2007), Bansal, Dittman, and Kiku (2007), Bansal, Gallant, and Tauchen (2008), Bansal, Kiku, and Yaron (2007), Bansal and Yaron (2004), Campanale, Castro, and Clementi (2010), Campbell (1993 and 1996), Campbell and Viceira (2001), Chen, Favilukis and Ludvigson (2007), Croce (2006), Dolmas (1996), Gomes and Michaelides (2005), Hansen, Heaton, and Li (2008), Kaltenbrunner and Lochstoer (2008), Lettau and Uhlig (2002), Piazzesi and Schneider (2006), Rudebusch and Swanson (2008), Tallarini (2000), and Uhlig (2007). See also Hansen *et al.* (2008) for a survey of the literature.

policy interventions on variables, such as investment, that are central to the business cycle.

This is particularly easy to see if we concentrate on the consumption process that drives the stochastic discount factor under recursive preferences. Except in a few papers,[2] researchers interested in asset pricing have studied economies in which consumption follows an exogenous process. This is a potentially important shortcoming. First, production economies place tight restrictions on the comovements of consumption with other endogenous variables that exogenous consumption models are not forced to satisfy. Second, in DSGE models, the consumption process itself is not independent of the parameters fixing the IES and risk aversion. In comparison, by fixing the consumption process in endowment economies, a change in preferences implicitly translates to a change in the labor income process. This complicates the interpretation of preference parameters. Finally, considering production economies with labor supply is quantitatively relevant. Uhlig (2007) has shown how, with Epstein and Zin preferences, leisure significantly affects asset pricing through the risk-adjusted expectation operator. Thus, even when leisure enters separately in the period utility function, the recursive formulation will make the stochastic discount factor dependent on leisure.

Unfortunately, working with Epstein and Zin preferences is harder than working with expected utility. Instead of the simple optimality conditions of expected utility, recursive preferences imply necessary conditions that include the value function itself.[3] Therefore, standard linearization techniques cannot be employed. Hence, the literature has resorted to either simplifying the problem by working only with an exogenous flow for consumption, or using costly computational solution algorithms such as value function iteration (Croce, 2006) or projection methods (Campanale, Castro, and Clementi, 2010). The former solution precludes all those exercises in which consumption reacts endogenously to the dynamics of the model. The latter solution makes estimation, a basic goal of our paper, exceedingly challenging because of the time spent in the solution of the model for each set of parameter values.

We get around this obstacle by solving the equilibrium dynamics of our economy with perturbation methods. Thus, we illustrate how this approach is a fast and reliable way to compute models with Epstein and Zin preferences. Our choice is motivated by several considerations. First, we will show how perturbation offers insights into the structure of the

---

[2]Among recent examples, Backus, Routledge, and Zin (2007), Campanale, Castro, and Clementi (2010), Croce (2006), or Rudebusch and Swanson (2008).

[3]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 estimator. 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. See, for instance, Campbell (1996) and Lustig, Van Nieuwerburgh, and Verdehaln (2007).

solution of the model that enhance our understanding of the role of recursive preferences. In particular, we will learn that the first-order approximation to the policy functions of our model with Epstein and Zin preferences is equivalent to that of the model with standard utility and the same IES. The risk aversion parameter does not show up in this first-order approximation. Instead, risk aversion appears in the constant of the second-order approximation that captures precautionary behavior. This 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, the risk aversion parameter influences the average level and the slope of the yield curve. Risk aversion also enters into the coefficients of the third-order approximation that change the slope of the response of the solution to variations in the states of the model.

Furthermore, in practice, perturbation methods are the only computationally feasible procedure to solve the policy-oriented, medium-scale DSGE models that have dozens of state variables (Christiano, Eichenbaum, and Evans, 2005, and Smets and Wouters, 2007). Consequently, our solution approach has an important applicability. Moreover, since implementing a third-order perturbation is feasible with off-the-shelf software such as `Dynare`, which requires minimum programming knowledge by the user, our findings may induce researchers to explore recursive preferences in further detail.

In contemporaneous work, Rudebusch and Swanson (2008) also use perturbation methods to solve a DSGE model with recursive preferences. Their model differs from ours in that they do not include endogenous capital, which in a production model imposes important constraints on the properties of returns. In addition, they rely on an approximation to the returns on bonds through a consol. Andreasen and Zabczyk (2010) show, however, that this may introduce computational biases. In terms of methodology, we estimate the model via maximum likelihood, whereas Rudebusch and Swanson calibrate the parameters.

The second contribution of our paper is to show how to estimate the model by maximum likelihood. In studying the asset pricing implications of full-fledged equilibrium models, it is common practice to calibrate the parameters.[4] While this approach illuminates the main economic mechanism at work, it might overlook some restrictions implied by the model. This is relevant, since various asset pricing models can explain the same set of moments, but the economic mechanism generating the results, be it habits, long-run risks, or rare disasters, is quite different and implies diverse equilibrium dynamics. Our likelihood-based inference imposes all cross-equation restrictions implied by the equilibrium model and is, therefore,

---

[4]Famous examples are Campbell and Cochrane (1999) and Bansal and Yaron (2004). A notable exception is Chen, Favilukis, and Ludvigson (2007), who estimate an endowment economy in which the representative agent has habit-type preferences.

much more powerful in testing asset pricing models.

The third contribution of our paper is to the fast-growing literature on term structure models. These models are successful in fitting the term structure of interest rates, but this is typically accomplished using latent variables.[5] Even though some papers include macroeconomic or monetary policy variables, such variables still enter in a reduced-form way. Our approach imposes much additional structure on such models, but the restrictions directly follow from the assumptions we make about preferences and technology. Such models obviously underperform the statistical models,[6] but they improve our understanding as to which preferences and technology processes induce a realistic term structure of interest rates. Furthermore, as we have argued before, macroeconomists require a structural model to design and evaluate economic policies that might affect the term structure of interest rates in an environment with recursive preferences.

Summarizing, this paper is the first one to show how to use perturbation techniques in combination with the particle filter to overcome the difficulties in estimating production models with recursive preferences. To do so, we rely on a prototype real business cycle economy with Epstein and Zin preferences and long-run growth through a unit root in the law of motion for technological progress.

As our first step, we perturb the value function formulation of the household problem to obtain a third-order approximation to the solution of the model given some parameter values in a trivial amount of time.[7] Given our econometric goals, an additional advantage of our solution technique is that we do not limit ourselves to the case with unitary IES, as Tallarini (2000) and others are forced to do.[8] There are three reasons why this flexibility might be important. First, because restricting the IES to one seems an unreasonably tight restriction that is hard to reconcile with previous findings. Second, a value of the IES equal to one implies that the consumption-wealth ratio is constant over time. This implication of the model is hard to verify because total wealth is not directly observable, since it includes human wealth. However, different attempts at measurement, such as Lettau and Ludvigson (2001) or Lustig, van Nieuwerburgh, and Verdelhan (2007), reject the hypothesis that the ratio of consumption to wealth is constant. Third, the debate between Campbell (1996) and

---

[5] See, among others, Dai and Singleton (2000 and 2002), Duffee (2002), Cochrane and Piazzesi (2005 and 2008), and Ang, Bekaert, and Wei (2008).

[6] Campbell and Cochrane (2000) make a similar point in relation to consumption-based and reduced-form asset pricing models.

[7] In companion work, Caldara *et al.* (2010) document that this solution is highly accurate and compare it with alternative computational approaches.

[8] 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.

Bansal and Yaron (2004) about the usefulness of the Epstein and Zin approach pertains to the right value of the IES. By directly estimating this parameter using all economic restrictions implied by production economies, we contribute to this conversation.

The second step in our procedure is to use the particle filter to evaluate the likelihood function of the model (Fernández-Villaverde and Rubio-Ramírez, 2007).[9] Evaluating the likelihood function of a DSGE 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. These non-linearities make the conditional distribution of states intractable and prevent the application 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.

We estimate the model with US data on consumption growth, output growth, five bond yields, and inflation over the period 1953.Q1 to 2008.Q4. The point estimates imply a high coefficient of risk aversion, an IES well above one, and substantial adjustment costs of capital. However, we find that the model barely generates a bond risk premium and substantially underestimates the volatility of bond yields. The model is able, however, to reproduce the autocorrelation patterns in consumption growth, the 1-year bond yield, and inflation. To better understand the model's shortcomings and how the parameters are identified, we re-estimate the model based on subsets of our data. First, we omit inflation from our sample. The estimates we then find imply a bond risk premium that is comparable to the one we measure in the data, and the model reproduces the empirical bond yield volatility. However, this "success" is explained by the fact that, in this case, the volatility of inflation is too high. Finally, we estimate our model based only on bond yields. The estimates and implications are remarkably similar to the previous case in which we omit the observations on inflation. This leads us to conclude that the parameters are mostly identified from yield and inflation data. This also illustrates the large amount of information regarding structural parameters in finance data and the importance of incorporating asset pricing observations into the estimation of DSGE models.

The rest of the paper is structured as follows. In section 2, we present our model. In section 3, we explain how we solve the model with perturbation and what we learn about the structure of the solution. Section 4 describes the likelihood-based method based on the use of the particle filter. Section 5 reports the data and our empirical findings. Section 6 outlines several extensions and section 7 concludes. Three appendices offer further details.

---

[9]A recent application of the particle filter in finance includes Binsbergen and Koijen (2010), who use the particle filter to estimate the time series of expected returns and expected growth rates using a present-value model.

2. A Production Economy with Recursive Preferences

In this section, we present a simple production economy that we will later take to the data and use it to price nominal bonds at different maturities. The only deviation from the standard stochastic neoclassical growth model is that we consider Epstein and Zin preferences, instead of standard state-separable constant relative risk aversion (CRRA). In addition, we add a process for inflation that captures well the dynamics of price increases in the data and that will allow us to value nominal bonds.

2.1. Preferences

There is a representative household whose utility function over streams of consumption c_t and leisure $1 - l_t$ is:

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

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

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

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 $\gamma = \frac{1}{\psi}$, we have that $\theta = 1$ and the recursive preferences collapse to the standard CRRA case. The Epstein and Zin framework implies that the household has preferences for the timing of the resolution of uncertainty. In our notation, if $\gamma > \frac{1}{\psi}$, the household prefers an early resolution of uncertainty, and if $\gamma < \frac{1}{\psi}$, a later resolution. The discount factor is β and one period corresponds to one quarter.

2.2. Technology

There is a representative firm with access to a technology described by a neoclassical production function $y_t = k_t^\zeta \left(z_t l_t \right)^{1-\zeta}$, where output y_t is produced with capital, k_t, labor, l_t, and technology z_t. This technology evolves as a random walk in logs with drift λ:

$$\log z_{t+1} = \lambda + \log z_t + \chi \sigma_\varepsilon \varepsilon_{zt+1}, \tag{1}$$

where $\varepsilon_{zt} \sim \mathcal{N}(0, 1)$. The parameter χ scales the standard deviation of the productivity shock, σ_ε. This parameter, also called the perturbation parameter, will facilitate the presentation of

our solution method later on. We pick this specification over trend stationarity motivated by Tallarini (2000), who shows that a unit root representation such as (1) facilitates matching the observed market price of risk in a model close to ours. Similarly, Álvarez and Jermann (2005) calculate that most of the unconditional variation in the pricing kernel comes from the permanent component. Part of the reason, as emphasized by Rouwenhorst (1995), is that period-by-period unit root shifts of the long-run growth path of the economy increase the variance of future paths of the variables and, hence, the utility cost of risk.

2.3. Budget and Resource Constraints

The budget constraint of the household is:

$$c_t + i_t + \frac{b_{t+1}}{p_t} \frac{1}{R_t} = r_t k_t + w_t l_t + \frac{b_t}{p_t}, \tag{2}$$

where p_t is the price level of the final good at time t, i_t is investment in period t, k_t is capital in period t, b_t is the number of one-period uncontingent bonds held in period t that pay one nominal unit in period $t+1$, R_t^{-1} is their unit price at time t, w_t is the real wage at time t, and r_t is the real rental price of capital at time t, both measured in units of the final good. In the interest of clarity, we include in the budget constraint only the one-period uncontingent bond we just described. Using the pricing kernel, in section 2.6, we will write the set of equations that determine the prices of nominal bonds at any maturity. In any case, their price in equilibrium will be such that the representative agent will hold a zero amount of them. The aggregate resource constraint is

$$y_t = c_t + i_t. \tag{3}$$

2.4. Dynamics of the Capital Stock

Capital depreciates at rate δ. Thus, the dynamics of the capital stock is given by:

$$k_{t+1} = (1 - \delta) k_t + G\left(\frac{i_t}{k_t}\right) k_t, \tag{4}$$

in which:

$$G\left(\frac{i_t}{k_t}\right) = \frac{a_1}{1 - \tau} \left(\frac{i_t}{k_t}\right)^{1 - \frac{1}{\tau}} + a_2,$$

denotes the adjustment cost of capital as in Jermann (1998). We normalize:

$$a_1 = \frac{e^\lambda - 1 + \delta}{1 - \tau},$$

9

and

$$a_2 = \left(e^\lambda - 1 + \delta\right)^{\frac{1}{\tau}},$$

such that adjustment costs do not affect the steady state of the model.

2.5. Inflation Dynamics

In our data, we will include nominal bond yields at different maturities as part of our observables. Hence, we need to take a stand on how inflation, $\log \pi_t$, evolves over time. Since we want to keep the model as stylized as possible, we take inflation as an exogenous process that does not affect allocations. Therefore, money is neutral in our economy.[10] Also, the representative household has rational expectations about the process.

Following Campbell and Viceira (2001), among others, inflation is assumed to follow the process:

$$\log \pi_{t+1} = \log \bar{\pi} + \rho \left(\log \pi_t - \log \bar{\pi}\right) + \chi \left(\sigma_\omega \omega_{t+1} + \kappa_0 \sigma_\varepsilon \varepsilon_{zt+1}\right) + \iota \left(\sigma_\omega \omega_t + \kappa_1 \sigma_\varepsilon \varepsilon_{zt}\right),$$

where $\omega_t \sim \mathcal{N}(0,1)$, $\omega_t \perp \varepsilon_{zt}$, and $\log \pi_t \equiv \log p_t - \log p_{t-1}$. The parameters κ_0 and κ_1 capture the correlation of unexpected and expected inflation with innovations to the technology process, ε_{zt+1} and ε_{zt} respectively. As before, χ is the perturbation parameter.

This specification allows us to accomplish two objectives. First, it lets us consider a correlation between innovations to inflation expectations and innovations to the stochastic discount factor. This implies that bond prices do not move one to one with expected inflation and that we have an inflation premium. Second, the MA components capture the negative first-order autocorrelation and the small higher order autocorrelations of inflation growth reported by Stock and Watson (2007). These authors prefer an IMA(1,1) process for inflation instead of our ARMA specification. Unfortunately, we cannot handle a unit root process for inflation because the perturbation method to be used to solve the model requires that inflation have a steady-state value. To minimize the effects of our stationarity assumption, we will calibrate ρ to be 0.955 (the highest possible value for ρ such that we do not suffer from numerical instabilities) and $\bar{\pi}$ to 1.009 to match the observed average inflation. Our choice of ρ is close to the value estimated by Stock and Watson (2007) when an ARMA(1,1) model related to the one described here is estimated over a sample period similar to ours.

---

[10] This is not as restrictive an assumption as it might seem. Nominal rigidities, while important to capture business cycle dynamics, are not very useful for matching asset pricing properties (see, for instance, De Paoli, Scott, and Weeken, 2007, or Doh, 2009). This is particularly true once we account for, as we do, part of the relation between price changes and technology shocks through our process for inflation.

2.6. Pricing Nominal Bonds

Given our process for inflation, we now move to price nominal bonds. In Appendix 8.1, we show that the stochastic discount factor (SDF) for our economy is given by:

$$M_{t+1} = \beta \left(\frac{c_{t+1}^v \left(1 - l_{t+1}\right)^{1-v}}{c_t^v \left(1 - l_t\right)^{1-v}} \right)^{\frac{1-\gamma}{\theta}} \frac{c_t}{c_{t+1}} \left(\frac{V_{t+1}^{1-\gamma}}{\mathbb{E}_t V_{t+1}^{1-\gamma}} \right)^{1-\frac{1}{\theta}}.$$

where the value function V_t is defined as:

$$V_t = \max_{c_t, l_t, i_t} U_t,$$

subject to (3) and (4). We switch notation to V_t because it is convenient to distinguish between the utility function of the household, U_t, and the value function that solves the household's problem V_t. Note that since the welfare theorems hold in our model, this value function is also equal to the solution of the social planner's problem, a result we use in the appendices in a couple of steps. Nothing of substance depends on working with the social planner's problem except that the notation is easier to handle.

Hence, the Euler equation for the one-period nominal bonds is:

$$\mathbb{E}_t \left(M_{t+1} \frac{1}{\pi_{t+1}} \right) = \frac{1}{R_t},$$

which can be written as:

$$\mathbb{E}_t \left[\beta \left(\frac{c_{t+1}^v \left(1 - l_{t+1}\right)^{1-v}}{c_t^v \left(1 - l_t\right)^{1-v}} \right)^{\frac{1-\gamma}{\theta}} \frac{c_t}{c_{t+1}} \left(\frac{V_{t+1}^{1-\gamma}}{\mathbb{E}_t V_{t+1}^{1-\gamma}} \right)^{1-\frac{1}{\theta}} \frac{1}{\pi_{t+1}} \right] = \frac{1}{R_t}.$$

But what is more important for us, we can also compute bond prices recursively using the following formula:

$$\mathbb{E}_t \left(M_{t+1} \frac{1}{\pi_{t+1}} \frac{1}{R_{t+1,t+m}} \right) = \frac{1}{R_{t,t+m}}, \tag{5}$$

with $R_{t,t+m}^{-1}$ being the time-t price of an m-periods nominal bond. Note that we write $R_{t,t+1} = R_t$ and $R_{t+1,t+1} = 1$.

Disappointingly, we do not have any analytic expression for the equilibrium dynamics of the model. In the next two sections, we will explain, first, how to use perturbation methods to solve for these dynamics. Second, we will show how to exploit the output of the perturbation to write a state-space representation of the model and how to exploit this representation to evaluate the associated likelihood function.

11

3. Solving the Model Using Perturbation

As we will explain in detail momentarily, we solve our economy by perturbing the value function of the household plus the equilibrium conditions of the model defined by optimality and feasibility. The advantage of perturbation over other methods such as value function iteration or projection is that it produces an answer in a sufficiently fast manner as to make likelihood estimation feasible.[11]

We are not the first to explore the perturbation of value functions. Judd (1998) proposes the idea but does not elaborate much on the topic. More recently, Schmitt-Grohé and Uribe (2005) use a second-order approximation to the value function to rank different fiscal and monetary policies in terms of welfare.

Our solution approach is also linked with that of Benigno and Woodford (2006) and Hansen and Sargent (1995). Benigno and Woodford (2006) 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.[12] Thanks to this alternative approximation, the authors find the correct local welfare ranking of different policies. Our method, as theirs, can deal with non-linear constraints and obtain the correct local approximation. One advantage of our method is that it is easily generalizable to higher-order approximations without complication. 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, imposes 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. This step is time consuming and it is not obvious that the required fixed-point exists or that the recentering converges. Our method does not suffer from those limitations.

In this paper, we find a third-order approximation to the value function and decision rules that yields outstanding accuracy as measured by Euler equation errors. We stop at order three because third-order terms allow for a time-varying risk premium, an important feature of the data that we want to capture, while keeping computational costs at a reasonable level.[13] Since we need to solve the model hundreds of times for different parameter values in our estimation, speed is of the utmost importance. Obtaining higher-order approximations is conceptually straightforward but computationally cumbersome.

---

[11] Also, as documented by Caldara *et al.* (2010) while exploring how to compute a model similar to ours, the accuracy of a high-order perturbation is excellent even far away from the steady state of the model.

[12] See also Levine, Pearlman, and Pierse (2007) for a similar treatment of the problem.

[13] See also Binsbergen (2009)

In our exposition, we use a concise notation to illustrate the required steps. Otherwise, the algebra becomes too involved to be developed explicitly in the paper in all its detail. In our application, the symbolic algebra is undertaken by a computer employing `Mathematica`, which automatically generates `Fortran 95` code that we can evaluate numerically.

3.1. Basic Structure

Since our model is non-stationary, we need to make it stationary by rescaling the variables by z_{t-1}. Hence, for any variable x_t, we denote its normalized value by $\tilde{x}_t = x_t/z_{t-1}$. Also, remember that the stochastic processes are written in terms of a perturbation parameter χ. When $\chi = 1$, we are dealing with the stochastic version of the model and when $\chi = 0$ we are dealing with the deterministic case with steady state $\tilde{k}_{ss}$ and $\log \tilde{z}_{ss} = \lambda$.

Thus, we can write the value function, $V\left(\tilde{k}_t, \log \tilde{z}_t; \chi\right)$, and the decision rules for consumption, $c\left(\tilde{k}_t, \log \tilde{z}_t; \chi\right)$, investment, $i\left(\tilde{k}_t, \log \tilde{z}_t; \chi\right)$, capital, $k\left(\tilde{k}_t, \log \tilde{z}_t; \chi\right)$, and labor, $l\left(\tilde{k}_t, \log \tilde{z}_t; \chi\right)$, as a function of the two rescaled states, $\tilde{k}_t$ and $\log \tilde{z}_t$ and the perturbation parameter, χ. Since money is neutral in this model, the above-described value function and decision rules do not depend on inflation. This is extremely helpful because it allows us to first solve for them without considering inflation and, only then, using the previous findings, to solve for nominal bond prices that do depend on inflation.

Now, we find the approximations to the value function and the decision rules for consumption, labor, investment, and capital. Later, we will find the approximations to the bond yields. Define $s_t = \left(\tilde{k}_t - \tilde{k}_{ss}, \log \tilde{z}_t - \log \tilde{z}_{ss}; 1\right)$ as the vector of states in differences with respect to the steady state, where s_{it} is the $i-$component of this vector at time t for $i = \{1, 2, 3\}$. Under differentiability conditions, the third-order Taylor approximation of the value function, evaluated at $\chi = 1$, around the steady state is

$$V\left(\tilde{k}_t, \log \tilde{z}_t; 1\right) \simeq V_{ss} + V_{i,ss} s_t^i + \frac{1}{2} V_{ij,ss} s_t^i s_t^j + \frac{1}{6} V_{ijl,ss} s_t^i s_t^j s_t^l, \tag{6}$$

where each term $V_{\ldots,ss}$ is a scalar equal to a derivative of the value function evaluated at the steady state:

$$V_{ss} \equiv V\left(\tilde{k}_{ss}, \log \tilde{z}_{ss}; 0\right),$$

$$V_{i,ss} \equiv V_i\left(\tilde{k}_{ss}, \log \tilde{z}_{ss}; 0\right) \text{ for } i = \{1, 2, 3\},$$

$$V_{ij,ss} \equiv V_{ij}\left(\tilde{k}_{ss}, \log \tilde{z}_{ss}; 0\right) \text{ for } i, j = \{1, 2, 3\},$$

and

$$V_{ijl,ss} \equiv V_{ijl}\left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}; 0\right) \text{ for } i,j,l = \{1,2,3\},$$

where we have followed the tensor notation:

$$
\begin{aligned}
V_{i,ss} s_t^i &= \sum_{i=1}^{3} V_{i,ss} s_{i,t}, \\
V_{ij,ss} s_t^i s_t^j &= \sum_{i=1}^{3}\sum_{i=1}^{3} V_{ij,ss} s_{i,t} s_{j,t},
\end{aligned}
$$

and

$$V_{ijl,ss} s_t^i s_t^j s_t^l = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{l=1}^{3} V_{ijl,ss} s_{i,t} s_{j,t} s_{l,t},$$

that eliminates the symbol $\sum_{i=1}^{3}$ when no confusion arises.

Expression (6) has interesting properties. For example, when we evaluate it at $\left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}; 1\right)$ (the values of capital and productivity growth of the steady state and positive variance of shocks), all terms will drop, except V_{ss}, $V_{3,ss}$, $V_{33,ss}$, and $V_{333,ss}$. But it turns out that all the terms in odd powers of χ (in this case, $V_{3,ss}$ and $V_{333,ss}$) are identically equal to zero. Therefore, a third-order approximation of the value function evaluated in $\left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}; 1\right)$ is:

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

where $\frac{1}{2}V_{33,ss}$ is a measure of the welfare cost of the business cycle, that is, of how much utility changes when the variance of the productivity shocks is σ^2 instead of zero.[14] This quantity

---

[14] This welfare cost can easily be transformed into consumption equivalent units by computing the decrease in consumption φ that makes the household indifferent between consuming $(1 - \varphi)c_{ss}$ units per period with certainty or c_t units with risk. The steady-state value function is:

$$V_{ss} = \left(\frac{1}{1 - \beta\widetilde{z}_{ss}^{\frac{v(1-\gamma)}{\theta}}}\right)^{\frac{\theta}{1-\gamma}} \widetilde{c}_{ss}^{v}(1 - l_{ss})^{1-v},$$

which implies:

$$\widetilde{c}_{ss}^{v}(1 - l_{ss})^{1-v} + \frac{1}{2}\left(1 - \beta\widetilde{z}_{ss}^{\frac{v(1-\gamma)}{\theta}}\right)^{\frac{\theta}{1-\gamma}} V_{33,ss} = \left(\widetilde{c}_{ss}(1 - \varphi)\right)^{v}(1 - l_{ss})^{1-v}.$$

Then:

$$\varphi = 1 - \left[1 + \frac{1}{2}\frac{\left(1 - \beta\widetilde{z}_t^{\frac{v(1-\gamma)}{\theta}}\right)^{\frac{\theta}{1-\gamma}}}{\widetilde{c}_{ss}^{v}(1 - l_{ss})^{1-v}} V_{33,ss}\right]^{\frac{1}{v}}.$$

is not necessarily negative, both because of the Jensen's inequality (our productivity process is in logs) and because concave utility functions may have convex indirect utility functions (see Cho and Cooley, 2000, for an example in a real business cycle model). The term $V_{33,ss}$ is also a key difference of perturbation from the standard linear-quadratic approximation, where constants are dropped because they are irrelevant for the optimization. This is yet another advantage of perturbation: $V_{33,ss}$ measures the effects of risk in the value function and we want to keep it for welfare evaluation.

Following the same tensor notation as before, the decision rule for any control variable var (consumption, labor, investment, and capital) can be approximated in a similar way as

$$var\left(\widetilde{k}_t, \log \widetilde{z}_t; 1\right) \simeq var_{ss} + var_{i,ss} s_t^i + \frac{1}{2} var_{ij,ss} s_t^i s_t^j + \frac{1}{6} var_{ijl,ss} s_t^i s_t^j s_t^l,$$

in which we define

$$var_{ss} \equiv var\left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}; 0\right),$$

$$var_{i,ss} \equiv var_i\left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}; 0\right) \text{ for } i = \{1, 2, 3\},$$

$$var_{ij,ss} \equiv var_{ij}\left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}; 0\right) \text{ for } i, j = \{1, 2, 3\},$$

and

$$var_{ijl,ss} \equiv var_{ijl}\left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}; 0\right) \text{ for } i, j, l = \{1, 2, 3\}.$$

The problem is that the derivatives $V_{...,ss}$ and $var_{...,ss}$ are not known. A perturbation method finds them by taking derivatives of a set of equations describing the equilibrium of the model and applying an implicit function theorem to solve for these unknown derivatives. But once we have reached this point, there are two paths we can follow to obtain a set of equations to perturb. The first path, the one in this paper, is to write down the equilibrium conditions of the model plus the definition of the value function. Then, we take successive derivatives with respect to states in this augmented set of equilibrium conditions and solve for the unknown coefficients, which happen to be the derivatives of the value function and decision rules that we need to get our higher-order approximations. This approach, which we call equilibrium conditions perturbation (ECP), allows us to get, after n iterations, the n-th-order approximation to the value function and to the decision rules.

A second path would be to take derivatives of the value function with respect to states and controls and use those derivatives to find the unknown coefficients. This approach, which we call value function perturbation (VFP), delivers after $(n + 1)$ steps, the $(n + 1)$-th-order

approximation to the value function and the n-th-order approximation to the decision rules. This alternative may be more convenient when it is difficult to eliminate levels or derivatives of the value function from the equilibrium conditions or when the value function is smoother than other equilibrium conditions.

3.1.1. Augmented Equilibrium Conditions

We derive now the set of augmented equilibrium conditions to implement the ECP approach. The household's problem is given by:

$$V_t = \max_{c_t, l_t, k_{t+1}, i_t} \left[\left(c_t^v (1 - l_t)^{1-v} \right)^{\frac{1-\gamma}{\theta}} + \beta \left(\mathbb{E}_t V_{t+1}^{1-\gamma} \right)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}},$$

subject to the budget and resource constraint in (2) and (3), and the dynamics of the capital stock in (4).[15]

We find the first-order conditions of the household with respect to consumption, labor, capital, and investment:

$$\frac{\partial V_t}{\partial c_t} = \xi_t,$$

$$\frac{\partial V_t}{\partial l_t} = -\xi_t w_t,$$

$$q_t = \mathbb{E}_t \left[\xi_{t+1} r_{t+1} + q_{t+1} \left(1 - \delta + \frac{\partial G\left(\frac{i_{t+1}}{k_{t+1}} \right) k_{t+1}}{\partial k_{t+1}} \right) \right],$$

and

$$\xi_t = q_t \frac{\partial G\left(\frac{i_t}{k_t} \right) k_t}{\partial i_t},$$

where ξ_t is the Lagrangian multiplier associated with the resource constraint and q_t is the Lagrangian multiplier associated with the law of motion of capital. Using the last two first-order conditions and the fact that $M_{t+1} = \xi_{t+1}/\xi_t$, the Euler equation for capital can be written as:

$$\left(\frac{i_t}{k_t} \right)^{\frac{1}{\tau}} = \mathbb{E}_t \left[\begin{array}{l} \beta \left(\frac{c_{t+1}^v (1-l_{t+1})^{1-v}}{c_t^v (1-l_t)^{1-v}} \right)^{\frac{1-\gamma}{\theta}} \frac{c_t}{c_{t+1}} \left(\frac{V_{t+1}^{1-\gamma}}{\mathbb{E}_t V_{t+1}^{1-\gamma}} \right)^{1-\frac{1}{\theta}} \times \\ \left(a_2 r_{t+1} + \left(\frac{i_{t+1}}{k_{t+1}} \right)^{\frac{1}{\tau}} \left(1 - \delta + a_1 + \frac{a_2}{\tau - 1} \left(\frac{i_{t+1}}{k_{t+1}} \right)^{1-\frac{1}{\tau}} \right) \right) \end{array} \right].$$

---

[15] We could reduce the number of controls by substituting out of the budget constraints. We choose not to do so, since we will employ these first-order conditions in the solution method.

If $\tau \to \infty$, that is, if we do not have adjustment costs, this equation simplifies to:

$$1 = \mathbb{E}_t \left[\frac{\xi_{t+1}}{\xi_t} \left(r_{t+1} + 1 - \delta \right) \right],$$

which is the standard equation without adjustment costs. Also, Tobin's Q is defined as:

$$\frac{q_t}{\xi_t} = \frac{1}{a_2} \left(\frac{i_t}{k_t} \right)^{\frac{1}{\tau}}$$

$$= \frac{1}{(e^\lambda - 1 + \delta)^{\frac{1}{\tau}}} \left(\frac{i_t}{k_t} \right)^{\frac{1}{\tau}},$$

which also equals 1 if $\tau \to \infty$.

Using the first-order conditions with respect to consumption and leisure, we can relate consumption, leisure, and the wage rate:

$$\frac{1-\upsilon}{\upsilon} \frac{c_t}{1-l_t} = (1-\zeta) k_t^\zeta z_t^{1-\zeta} l_t^{-\zeta}.$$

Hence, the augmented set of equilibrium conditions to be used to solve for V_t, i_t, k_{t+1}, c_t, and l_t is:

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

$$\left(\frac{i_t}{k_t} \right)^{\frac{1}{\tau}} = \mathbb{E}_t \left[\begin{array}{c} \beta \left(\frac{c_{t+1}^\upsilon (1-l_{t+1})^{1-\upsilon}}{c_t^\upsilon (1-l_t)^{1-\upsilon}} \right)^{\frac{1-\gamma}{\theta}} \frac{c_t}{c_{t+1}} \left(\frac{V_{t+1}^{1-\gamma}}{\mathbb{E}_t V_{t+1}^{1-\gamma}} \right)^{1-\frac{1}{\theta}} \times \\ \left(a_2 r_{t+1} + \left(\frac{i_{t+1}}{k_{t+1}} \right)^{\frac{1}{\tau}} \left(1 - \delta + a_1 + \frac{a_2}{\tau - 1} \left(\frac{i_{t+1}}{k_{t+1}} \right)^{1-\frac{1}{\tau}} \right) \right) \end{array} \right],$$

$$\frac{1-\upsilon}{\upsilon} \frac{c_t}{1-l_t} = (1-\zeta) k_t^\zeta z_t^{1-\zeta} l_t^{-\zeta},$$

$$c_t + i_t = k_t^\zeta z_t^{1-\zeta} l_t^{1-\zeta},$$

and

$$k_{t+1} = (1-\delta) k_t + G \left(\frac{i_t}{k_t} \right) k_t,$$

together with the law of motion for $\log z_t$.

After normalizing the set of equilibrium conditions as described in Appendix 8.3, we write them in more compact notation:

$$F \left(\widetilde{k}_t, \log \widetilde{z}_t; \chi \right) = \mathbf{0},$$

17

where F is a 5-dimensional function (and where all the endogenous variables in the previous equation are not represented explicitly because they are functions themselves of $\widetilde{k}_t$, $\log \widetilde{z}_t$ and χ) and $\mathbf{0}$ is the vectorial zero.

3.1.2. Approximating the Value Function and Decision Rules

The task now is to take successive derivatives of F and solve for the unknown coefficients of the Taylor expansions of the value function and decision rules that we presented before. These unknown coefficients appear in these derivatives because the augmented equilibrium conditions are expressed in terms of the different variables and we need to differentiate them with respect to the states. Those are precisely the terms in the Taylor expansions.

The first step is to evaluate F at $\chi = 0$. This gives us the steady-state values for $\widetilde{V}_{ss}$, $\widetilde{i}_{ss}$, $\widetilde{k}_{ss}$, $\widetilde{c}_{ss}$, and $\widetilde{l}_{ss}$. This steady state is identical to the steady state of the stochastic neoclassical growth model with a standard CRRA utility function.

Next, to find the first-order approximation to the value function and the decision rules, we take first derivatives of the function F with respect to $\left(\widetilde{k}_t, \log \widetilde{z}_t; \chi\right)$ and evaluate them at the steady state $\left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}; 0\right)$ that we just found to get

$$F_i \left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}; 0\right) = 0 \text{ for } i \in \{1, 2, 3\}.$$

This step gives us 15 different derivatives (5 equilibrium conditions times the 3 variables of F). With this quadratic 15-equations system, we find the first-order derivatives of the value function and of 4 decision rules (V_t, i_t, k_{t+1}, c_t, and l_t) evaluated at the steady state. We pick the stable solution by checking the appropriate eigenvalue.

To find the second-order derivatives to the value function and decision rules, we derive F_i with respect to $\left(\widetilde{k}_t, \log \widetilde{z}_t; \chi\right)$ and evaluate them at the steady state $\left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}; 0\right)$ from the first step to obtain:

$$F_{ij} \left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}; 0\right) = 0 \text{ for } i, j \in \{1, 2, 3\}.$$

This gives us yet a new system of equations. Then, we plug in the terms that we already know from the steady state and from the first-order approximation and we get that the only unknowns left are the second-order terms of the value function and of the decision rules. Quite conveniently, this system is linear and it can be solved quickly. Indeed, all the systems resulting from derivatives higher than the first one will be linear in the unknown coefficients. Repeating these steps, we can get any arbitrary-order approximation to the value function and decision rules. As mentioned above, in this paper we stop at order three.

3.1.3. Approximating Nominal Bonds Yields

To complete our computation, we also need to approximate the yield of nominal bonds. To do so, we take advantage of our recursive bond price equation (5). First, define:

$$sa_t = \left(\widetilde{k}_t - \widetilde{k}_{ss}, \log \widetilde{z}_t - \log \widetilde{z}_{ss}, \log \pi_t - \log \bar{\pi}, \omega_t; 1 \right)$$

which is the state vector in deviations with respect to the mean augmented with the difference of inflation with respect to its mean and the inflation innovation ω_t (sa stands for states augmented).

Then, in similar fashion to the value function and the decision rules, a third-order Taylor approximation to the yields is:

$$R_m \left(\widetilde{k}_t, \log \widetilde{z}_t, \log \pi_t, \omega_t; 1 \right) \simeq R_{m,ss} + R_{m,i,ss} sa_t + \frac{1}{2} R_{m,ij,ss} sa_t^i sa_t^j + \frac{1}{6} R_{m,ijl,ss} sa_t^i sa_t^j sa_t^l$$

for all m, in which we define:

$$R_{m,ss} \equiv R_{m,ss} \left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}, \log \bar{\pi}, 0; 0 \right),$$

$$R_{m,i,ss} \equiv R_{m,i} \left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}, \log \bar{\pi}, 0; 0 \right) \text{ for } i = \{1, 2, 3, 4, 5\},$$

$$R_{m,ij,ss} \equiv R_{m,ij} \left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}, \log \bar{\pi}, 0; 0 \right) \text{ for } i, j = \{1, 2, 3, 4, 5\},$$

and:

$$R_{m,ijl,ss} \equiv R_{m,ijl} \left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}; 0 \right) \text{ for } i, j, l = \{1, 2, 3, 4, 5\}.$$

Since in our data set we observe bond yields up to 20 quarters, we need to consider

$$\mathbb{E}_t \left(M_{t+1} \frac{1}{\pi_{t+1}} \frac{1}{R_{t+1,t+m}} \right) = \frac{1}{R_{t,t+m}},$$

for $m \in \{1, \ldots, 20\}$. This set of 20 first-order conditions can also be written, in more compact notation,

$$\widetilde{F} \left(\widetilde{k}_t, \log \widetilde{z}_t, \log \pi_t, \omega_t; \chi \right) = 0.$$

We can use $\widetilde{F}$ evaluated at $\chi = 0$ and the steady-state value $\widetilde{V}_{ss}$, $\widetilde{i}_{ss}$, $\widetilde{k}_{ss}$, $\widetilde{c}_{ss}$, and $\widetilde{l}_{ss}$ found above to find the steady-state values for $R_{t,t+j}$ for $m \in \{1, \ldots, 20\}$, $\log \pi_t$, and ω_t. These last two are, obviously, $\log \bar{\pi}$ and 0.

To find the first-order approximation to the nominal bond yields, we proceed as we did for the perturbation of the value function and decision rules. We take first derivatives of

the function $\widetilde{F}$ with respect to $\left(\widetilde{k}_t, \log \widetilde{z}_t, \log \pi_t, \omega_t; \chi \right)$ and evaluate them at the steady state $\left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}, \log \bar{\pi}, 0; 0 \right)$ to get:

$$\widetilde{F}_i \left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}, \log \bar{\pi}, 0; 0 \right) = 0 \text{ for } i \in \{1, 2, 3, 4, 5\}.$$

After substituting for the steady-state values and the first-order derivatives of the value function and decision rules that we found above, this step gives us a system of equations on the first-order derivatives of the yields evaluated at the steady state.

To find the second-order approximation to the nominal bond yields, we take derivatives on $\widetilde{F}_i$ with respect to $\left(\widetilde{k}_t, \log \widetilde{z}_t, \log \pi_t, \omega_t; \chi \right)$ and evaluate them at the steady state $\left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}, \log \bar{\pi}, 0; 0 \right)$ to get

$$\widetilde{F}_{ij} \left(\widetilde{k}_{ss}, \log \widetilde{z}_{ss}, \log \bar{\pi}, 0; 0 \right) = 0 \text{ for } i, j \in \{1, 2, 3, 4, 5\}.$$

We can use this linear system (together with the steady-state values, the first- and second-order derivatives of the value function and decision rules, and the first-order derivatives of the nominal bond yields that we found above) to find the second-order derivatives of the yields evaluated at the steady state. Repeating these steps, we can get any arbitrary-order approximation. We stop at order three because that would be enough to get the time-varying risk premium that we are interested in.

3.2. Role of γ

Direct inspection of the derivatives that we presented before (since the expressions are inordinately long, we cannot include them in the paper) reveals that:

1. The constant terms V_{ss}, var_{ss}, or $R_{m,ss}$ do **not** depend on γ, the parameter that controls risk aversion.

2. **None** of the terms in the first-order approximation, $V_{.,ss}$, $var_{.,ss}$, or $R_{m,.,ss}$ (for all m) depend on γ.

3. **None** of the terms in the second-order approximation, $V_{..,ss}$, $var_{..,ss}$, or $R_{m,..,ss}$ depend on γ, except $V_{33,ss}$, $var_{33,ss}$, and $R_{m,33,ss}$ (for all m). This last term is a constant that captures precautionary behavior caused by the presence of productivity shocks.

4. In the third-order approximation only the terms of the form $V_{33.,ss}$, $V_{3.3,ss}$, $V_{.33,ss}$ and $var_{33.,ss}$, $var_{3.3,ss}$, $var_{.33,ss}$ and $R_{m,33.,ss}$, $R_{m,3.3,ss}$, $R_{m,.33,ss}$ (for all m) that is, terms on functions of χ^2, depend on γ.

These observations tell us three important facts. First, a linear approximation to the decision rules does not depend on the risk aversion parameter or on the variance level of the productivity shock. In other words, it is certainty equivalent. Therefore, if we are interested in recursive preferences, we need to go at least to a second-order approximation. Second, given some fixed parameter values, the difference between the second-order approximation to the decision rules of a model with CRRA preferences and a model with recursive preferences is just a constant. This constant generates a second, indirect effect, because it changes the ergodic distribution of the state variables and, hence, the points where we evaluate the decision rules along the equilibrium path. In the third-order approximation, all of the terms on functions of χ^2 depend on γ. Thus, we can use them to further identify the risk aversion parameter, which is only weakly identified in the second-order approximation as it shows up only in one term and is not identified at all in the first-order approximation. These arguments also demonstrate how perturbation methods can provide analytic insights beyond computational advantages and help in understanding the numerical results in Tallarini (2000), who implements a recentering scheme that incorporates into the first-order approximation an effect similar to the second-order approximation constant.[16]

4. Estimation

Once we have our solution from the previous section, we use it to write a state-space representation of the dynamics of the states and observables that will allow us to evaluate the likelihood function of the model. For this last step, and since our solution is inherently non-linear (remember that the risk aversion parameter only affects the second- and third-order coefficients of the approximation), we will rely on the particle filter as described in Fernández-Villaverde and Rubio-Ramírez (2007).

4.1. State-Space Representation

As econometricians, we will observe per capita consumption growth, per capita output growth, the 1-, 2-, 3-, 4-, and 5-year nominal bond yields, and inflation. Per capita consumption growth and per capita output growth will provide macro information. The price of the

---

[16]This characterization is also crucial because it is plausible to entertain the idea that the richer structure of Epstein and Zin preferences is not identified (as in the example built by Kocherlakota, 1990). Fortunately, the second- and third-order terms allow us to learn from the observations. 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 and Zin preferences are generically recoverable from the price of equity or from the price of bonds. Furthermore, equity and bond prices are generically unique and smooth with respect to parameters.

nominal bonds provides us with financial data. Later, we will find that including finance data is key for the success of our empirical strategy.

Since our DSGE model has only two sources of uncertainty, the productivity shock and the inflation shock, we need to introduce measurement error to avoid stochastic singularity. It is common to have measurement error in term structure models. The justification comes from the idea that we do not observe zero coupon bonds. Instead, we observe the market prices of bonds with coupons and we need some procedure to back out the zero coupon bonds. This procedure induces measurement error. Similarly, National Income and Product Accounts (NIPA) can only provide researchers with an approximated estimate of output and consumption. Therefore, we will assume that all the variables (except inflation) are observed subject to a measurement error.[17]

It is easier to express the solution of our model in terms of deviations from steady state. Thus, for any variable var_t, we let $\widehat{var}_t = var_t - var_{ss}$. Also, we introduce a constant to keep track of means. Then, the law of motion for the states is

$$
\begin{pmatrix}
\widehat{\widehat{k}}_{t+1} \\
\widehat{\log \widetilde{z}_{t+1}} \\
\widehat{\log \pi_{t+1}} \\
\omega_{t+1} \\
1
\end{pmatrix}
=
\begin{pmatrix}
k_{i,ss}s_t^i + \frac{1}{2}k_{ij,ss}s_t^i s_t^j + \frac{1}{6}k_{ijl,ss}s_t^i s_t^j s_t^l \\
\sigma_\varepsilon \varepsilon_{zt+1} \\
\rho\widehat{\log \pi}_t + \left(\sigma_\omega \omega_{t+1} + \kappa_0 \sigma_\varepsilon \varepsilon_{zt+1}\right) + \iota\left(\sigma_\omega \omega_t + \kappa_1 \sigma_\varepsilon \widehat{\log \widetilde{z}_t}\right) \\
\omega_{t+1} \\
1
\end{pmatrix} .
$$

Since our observables are

$$
\mathbb{Y}_t = \left(\Delta \log c_t, \Delta \log y_t, R_{t,t+4}, R_{t,t+8}, R_{t,t+12}, R_{t,t+16}, R_{t,t+20}, \log \pi_t\right)' ,
$$

we need to map $\Delta \log c_t$ and $\Delta \log y_t$ into the model-scaled variables $\widetilde{c}_t$ and $\widetilde{c}_{t-1}$ and $\widetilde{y}_t$ and $\widetilde{y}_{t-1}$. We start with consumption. We observe that $\Delta \log c_t = \log c_t - \log c_{t-1}$ and we have that $c_t = \widetilde{c}_t z_{t-1}$ by our definition of re-scaled variables. Thus:

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

---

[17]Our exogenous process for inflation already has a linear additive innovation ω_{t+1}, which will make an additional measurement error difficult to identify.

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

$$\Delta \log c_t = \log\left(\widehat{\widetilde{c}}_t + \widetilde{c}_{ss}\right) - \log\left(\widehat{\widetilde{c}}_{t-1} + \widetilde{c}_{ss}\right) + \widehat{\log \widetilde{z}_{t-1}} + \lambda.$$

Equivalently,

$$\Delta \log y_t = \log\left(\widehat{\widetilde{y}}_t + \widetilde{y}_{ss}\right) - \log\left(\widehat{\widetilde{y}}_{t-1} + \widetilde{y}_{ss}\right) + \widehat{\log \widetilde{z}_{t-1}} + \lambda.$$

Hence, in order to simplify our state-space representation, it is convenient to consider $\left(\widehat{\widetilde{c}}_{t-1}, \widehat{\widetilde{y}}_{t-1}, \widehat{\log \widetilde{z}_{t-1}}\right)$ as additional (pseudo-)state variables. It is also the case that we need to map $\log \pi_t$ into our states. Since the law of motion of inflation is

$$\log \pi_t - \log \bar{\pi} = \rho\left(\log \pi_{t-1} - \log \bar{\pi}\right) + \left(\sigma_\omega \omega_t + \kappa_0 \sigma_\varepsilon \varepsilon_{zt}\right) + \iota\left(\sigma_\omega \omega_{t-1} + \kappa_1 \sigma_\varepsilon \varepsilon_{zt-1}\right),$$

we need to also consider $\left(\widehat{\log \pi_{t-1}}, \omega_{t-1}\right)$ as additional (pseudo-)state variables. We use the notation $\mathbb{S}_t$ to refer to the vector of augmented state variables.

Once this is done, our state-space representation can be written as a transition equation

$$\mathbb{S}_{t+1} = \begin{pmatrix} \widehat{\widetilde{k}}_{t+1} \\ \widehat{\log \widetilde{z}_{t+1}} \\ \widehat{\log \pi_{t+1}} \\ \omega_{t+1} \\ 1 \\ \widehat{\widetilde{c}}_t \\ \widehat{\widetilde{y}}_t \\ \widehat{\log \widetilde{z}_t} \\ \widehat{\log \pi_t} \\ \omega_t \end{pmatrix} = \begin{pmatrix} k_{i,ss} s_t^i + \frac{1}{2} k_{ij,ss} s_t^i s_t^j + \frac{1}{6} k_{ijl,ss} s_t^i s_t^j s_t^l \\ \sigma_\varepsilon \varepsilon_{zt+1} \\ \rho \widehat{\log \pi_t} + \left(\sigma_\omega \omega_{t+1} + \kappa_0 \sigma_\varepsilon \varepsilon_{zt+1}\right) + \iota\left(\sigma_\omega \omega_t + \kappa_1 \sigma_\varepsilon \widehat{\log \widetilde{z}_t}\right) \\ \omega_{t+1} \\ 1 \\ c_{i,ss} s_t^i + \frac{1}{2} c_{ij,ss} c_t^i c_t^j + \frac{1}{6} c_{ijl,ss} c_t^i c_t^j c_t^l \\ y_{i,ss} s_t^i + \frac{1}{2} y_{ij,ss} s_t^i s_t^j + \frac{1}{6} y_{ijl,ss} s_t^i s_t^j s_t^l \\ \widehat{\log \widetilde{z}_t} \\ \widehat{\log \pi_t} \\ \omega_t \end{pmatrix},$$

and a measurement equation

$$
\mathbb{Y}_t =
\begin{pmatrix}
\log\left(\widetilde{c}_{ss} + c_{i,ss}s_t^i + \frac{1}{2}c_{ij,ss}c_t^i c_t^j + \frac{1}{6}c_{ijl,ss}c_t^i c_t^j c_t^l\right) - \log\left(\widehat{\widetilde{c}}_{t-1} + \widetilde{c}_{ss}\right) + \widehat{\log \widetilde{z}_{t-1}} + \lambda \\
\log\left(\widetilde{y}_{ss} + y_{i,ss}s_t^i + \frac{1}{2}y_{ij,ss}s_t^i s_t^j + \frac{1}{6}y_{ijl,ss}s_t^i s_t^j s_t^l\right) - \log\left(\widehat{\widetilde{y}}_{t-1} + \widetilde{y}_{ss}\right) + \widehat{\log \widetilde{z}_{t-1}} + \lambda \\
R_{4,ss} + R_{i,4,ss}^i sa_t + \frac{1}{2}R_{ij,4,ss}sa_t^i sa_t^j + \frac{1}{6}R_{ijl,4,ss}sa_t^i sa_t^j sa_t^l \\
R_{8,ss} + R_{i,8,ss}^i sa_t + \frac{1}{2}R_{ij,8,ss}sa_t^i sa_t^j + \frac{1}{6}R_{ijl,8,ss}sa_t^i sa_t^j sa_t^l \\
R_{12,ss} + R_{i,12,ss}^i sa_t + \frac{1}{2}R_{ij,12,ss}sa_t^i sa_t^j + \frac{1}{6}R_{ijl,12,ss}sa_t^i sa_t^j sa_t^l \\
R_{16,ss} + R_{i,16,ss}^i sa_t + \frac{1}{2}R_{ij,16,ss}sa_t^i sa_t^j + \frac{1}{6}R_{ijl,16,ss}sa_t^i sa_t^j sa_t^l \\
R_{20,ss} + R_{i,20,ss}^i sa_t + \frac{1}{2}R_{ij,20,ss}sa_t^i sa_t^j + \frac{1}{6}R_{ijl,20,ss}sa_t^i sa_t^j sa_t^l \\
\log \bar{\pi} + \rho\widehat{\log \pi_{t-1}} + \kappa_0\widehat{\log \widetilde{z}_t} + \iota\left(\sigma_\omega\omega_{t-1} + \kappa_1\sigma_\varepsilon\widehat{\log \widetilde{z}_{t-1}}\right)
\end{pmatrix}
+
\begin{pmatrix}
\sigma_{v1}v_{1,t} \\
\sigma_{v2}v_{2,t} \\
\sigma_{v3}v_{3,t} \\
\sigma_{v4}v_{4,t} \\
\sigma_{v5}v_{5,t} \\
\sigma_{v6}v_{6,t} \\
\sigma_{v7}v_{7,t} \\
\sigma_\omega\omega_t
\end{pmatrix},
$$

where $\begin{pmatrix} \sigma_{v1}v_{1,t} & \sigma_{v2}v_{2,t} & \sigma_{v3}v_{3,t} & \sigma_{v4}v_{4,t} & \sigma_{v5}v_{5,t} & \sigma_{v6}v_{6,t} & \sigma_{v7}v_{7,t} & 0 \end{pmatrix}'$ is the measurement error vector. We assume that $v_{i,t} \sim \mathcal{N}(0,1)$ for all $i \in \{1,\ldots,7\}$ and $v_{i,t} \perp v_{j,t}$ for $i \neq j$ and $i,j \in \{1,\ldots,7\}$. The eighth element, the one corresponding to inflation, is missing since we assume no measurement error for inflation.

If we define $W_{t+1} = (\varepsilon_{zt+1}, \omega_{t+1})'$ and $\mathcal{V}_t = \begin{pmatrix} v_{1,t} & v_{2,t} & v_{3,t} & v_{4,t} & v_{5,t} & v_{6,t} & v_{7,t} \end{pmatrix}'$, we can write our transition and measurement equations more compactly as

$$
\mathbb{S}_{t+1} = h\left(\mathbb{S}_t, W_{t+1}\right), \tag{7}
$$

and

$$
\mathbb{Y}_t = g\left(\mathbb{S}_t, \mathcal{V}_t\right). \tag{8}
$$

4.2. Likelihood

We stack the set of structural parameters in our model in the vector:

$$
\Upsilon = (\beta, \gamma, \psi, \upsilon, \lambda, \zeta, \delta, \tau, \kappa_0, \iota, \kappa_1, \sigma_\varepsilon, \sigma_\omega, \sigma_{1v}, \sigma_{2v}, \sigma_{3v}, \sigma_{4v}, \sigma_{5v}, \sigma_{6v}, \sigma_{7v})'.
$$

The likelihood function $\mathcal{L}\left(\mathbb{Y}^T; \Upsilon\right)$ is the probability of the observations given some parameter values, where $\mathbb{Y}^t = \{\mathbb{Y}_s\}_{s=1}^t$ for $t \in \{1,\ldots,T\}$ is the history of observations up to time t.

Unfortunately, this likelihood is difficult to evaluate since we do not even have an analytic expression for our state-space representation. We tackle this problem by using a sequential

24

Monte Carlo.[18] 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 $\mathcal{L}\left(\mathbb{Y}_1 | \mathbb{Y}^0; \Upsilon\right) = \mathcal{L}\left(\mathbb{Y}_1; \Upsilon\right)$. Then, we condition on the states and integrate with respect to them to get

$$\mathcal{L}\left(\mathbb{Y}_t | \mathbb{Y}^{t-1}; \Upsilon\right) = \int \int \int \mathcal{L}\left(\mathbb{Y}_t | W_1^t, W_2^{t-1}, \mathbb{S}_0; \Upsilon\right) p\left(W_1^t, W_2^{t-1}, \mathbb{S}_0 | \mathbb{Y}^{t-1}; \Upsilon\right) dW_1^t dW_2^{t-1} d\mathbb{S}_0,$$

(9)

for $t \in \{2, \ldots, T\}$ where $W_{1,t} = \varepsilon_{zt}$, $W_{2,t} = \omega_t$, $W_i^t = \{W_{i,s}\}_{s=1}^t$ for $i = 1, 2$ and $t \in \{1, \ldots, T\}$, and

$$\mathcal{L}\left(\mathbb{Y}_1; \Upsilon\right) = \int \int \int \mathcal{L}\left(\mathbb{Y}_1 | W_1^1, \mathbb{S}_0; \Upsilon\right) p\left(W_1^1, \mathbb{S}_0; \Upsilon\right) dW_1^1 d\mathbb{S}_0. \qquad (10)$$

These expressions illustrate how the knowledge of $p\left(W_1^1, \mathbb{S}_0; \Upsilon\right)$ and of the sequence

$$\left\{p\left(W_1^t, W_2^{t-1}, \mathbb{S}_0 | \mathbb{Y}^{t-1}; \Upsilon\right)\right\}_{t=2}^{T}, \qquad (11)$$

is crucial for our procedure. If we know $\left(W_1^t, W_2^{t-1}, \mathbb{S}_0\right)$, computing $\mathcal{L}\left(\mathbb{Y}_t | W_1^t, W_2^{t-1}, \mathbb{S}_0; \Upsilon\right)$ is relatively easy; it is a change of variables from $W_{2,t}$ and V_t to $\mathbb{Y}_t$. The same is true for $\mathcal{L}\left(\mathbb{Y}_1 | W_1^1, \mathbb{S}_0; \Upsilon\right)$ if we know $\left(W_1^1, \mathbb{S}_0\right)$. However, given our model, we cannot characterize neither $p\left(W_1^1, \mathbb{S}_0; \Upsilon\right)$ nor the sequence (11) analytically. Even if we could, these two previous computations still leave open the issue of how to solve for the integrals in (9) and (10).

A common solution to these problems is to substitute $p\left(W_1^1, \mathbb{S}_0; \Upsilon\right)$ and

$$\left\{p\left(W_1^t, W_2^{t-1}, \mathbb{S}_0 | \mathbb{Y}^{t-1}; \Upsilon\right)\right\}_{t=2}^{T},$$

by an empirical distribution of draws from them. If we have such draws, we can approximate the likelihood using

$$\mathcal{L}\left(\mathbb{Y}_t | \mathbb{Y}^{t-1}; \Upsilon\right) \simeq \frac{1}{N} \sum_{i=1}^{N} \mathcal{L}\left(\mathbb{Y}_t | w_1^{t,i}, w_2^{t-1,i}, s_0^i; \Upsilon\right),$$

---

[18]This is not the only possible algorithm to do so, although it is a procedure that we have found useful in previous work. Alternatives include DeJong *et al.* (2007), Kim, Shephard, and Chib (1998), Fiorentini, Sentana, and Shephard (2004), and Fermanian and Salanié (2004).

where $w_1^{t,i}, w_2^{t-1,i}, s_0^i$ is the draw i from $p\left(W_1^t, W_2^{t-1}, \mathbb{S}_0 | \mathbb{Y}^{t-1}; \Upsilon\right)$ and

$$\mathcal{L}\left(\mathbb{Y}_1; \Upsilon\right) \simeq \frac{1}{N} \sum_{i=1}^N \mathcal{L}\left(\mathbb{Y}_1 | w_1^{1,i}, s_0^i; \Upsilon\right),$$

where $w_1^{1,i}, s_0^i$ is the draw i from $p\left(W_1^1, \mathbb{S}_0; \Upsilon\right)$.

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 $p\left(W_1^1, \mathbb{S}_0; \Upsilon\right)$ 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 $\left(w_1^{1,i}, s_0^i\right)$ that we require. Drawing from $\left\{p\left(W_1^t, W_2^{t-1}, \mathbb{S}_0 | \mathbb{Y}^{t-1}; \Upsilon\right)\right\}_{t=2}^T$ is more challenging. A popular approach to do so is to apply the particle filter (see Fernández-Villaverde and Rubio-Ramírez, 2007, for a more detailed explanation and references).

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, formulated by Rubin (1998), formalizes the idea:

Proposition 1. *Let* $\left\{w_1^{t,i}, w_2^{t-1,i}, s_0^i\right\}_{i=1}^N$ *be a draw from* $p\left(W_1^t, W_2^{t-1}, \mathbb{S}_0 | \mathbb{Y}^{t-1}; \Upsilon\right)$. *Let the sequence* $\left\{\widetilde{w}_1^{t,i}, \widetilde{w}_2^{t-1,i}, \widetilde{s}_0^i\right\}_{i=1}^N$ *be a draw with replacement from* $\left\{w_1^{t,i}, w_2^{t-1,i}, s_0^i\right\}_{i=1}^N$ *where the resampling probability is given by*

$$q_t^i = \frac{\mathcal{L}\left(\mathbb{Y}_t | w_1^{t,i}, w_2^{t-1,i}, s_0^i; \Upsilon\right)}{\sum_{i=1}^N \mathcal{L}\left(\mathbb{Y}_t | w_1^{t,i}, w_2^{t-1,i}, s_0^i; \Upsilon\right)}.$$

Then $\left\{\widetilde{w}_1^{t,i}, \widetilde{w}_2^{t-1,i}, \widetilde{s}_0^i\right\}_{i=1}^N$ *is a draw from* $p\left(W_1^t, W_2^{t-1}, \mathbb{S}_0 | \mathbb{Y}^t; \Upsilon\right)$.

Proposition 1, a direct application of Bayes' theorem, shows how we can take a draw from $p\left(W_1^t, W_2^{t-1}, \mathbb{S}_0 | \mathbb{Y}^{t-1}; \Upsilon\right)$ to get a draw from $p\left(W_1^t, W_2^{t-1}, \mathbb{S}_0 | \mathbb{Y}^t; \Upsilon\right)$ 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 $\left(W_1^t, W_2^{t-1}, \mathbb{S}_0\right)$. Thanks to SIR, the Monte Carlo method achieves sufficient accuracy in a reasonable amount of time. A naïve Monte Carlo, in comparison, would just draw simultaneously a whole sequence of $\left\{\left\{w_1^{t,i}, w_2^{t-1,i}, s_0^i\right\}_{i=1}^N\right\}_{t=1}^T$ 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 in probability dominates all

the remaining ones in weight. Simple simulations show that the degeneracy appears even after very few steps.

Given $\left\{ \widetilde{w}_1^{t,i}, \widetilde{w}_2^{t-1,i}, \widetilde{s}_0^i \right\}_{i=1}^N$ from $p\left(W_1^t, W_2^{t-1}, \mathbb{S}_0 | \mathbb{Y}^t; \Upsilon \right)$, we can apply the law of motion for states to generate $\left\{ w_1^{t+1,i}, w_2^{t,i}, s_0^i \right\}_{i=1}^N$ from $p\left(W_1^{t+1}, W_2^t, \mathbb{S}_0 | \mathbb{Y}^t; \Upsilon \right)$. 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 $t+1|t$.

4.3. Estimation Algorithms

Our paper emphasizes the likelihood-based estimation of DSGE models. In the interest of space, we will show results for maximum likelihood and comment briefly on how to find results for Bayesian estimation. Obtaining the maximum likelihood point estimate is complicated because the shape of the likelihood function is rugged and multimodal. Moreover, the particle filter generates an approximation to the likelihood that is not differentiable with respect to the parameters, precluding the use of optimization algorithms based on derivatives. To circumvent these problems, our optimization routine is a procedure known as covariance matrix adaptation evolutionary strategy, or CMA-ES (Hansen, Müller, and Koumoutsakos, 2003, and Andreasen, 2007). The CMA-ES is one of the most powerful evolutionary algorithms for real-valued optimization and has been applied successfully to many problems.

The CMA-ES approximates the inverse of the Hessian of the log-likelihood function by simulation. In each step of the algorithm, we simulate m candidate parameter values from the weighted mean and estimated variance-covariance matrix of the best candidate parameter values of the previous step. By selecting the best parameter values in each step and by adapting the variance-covariance matrix to the contour of the likelihood function, we direct the simulation toward the global maximum of our objective function. Thanks to the estimation of the variance-covariance matrix from the simulation, we by-pass the need of computing any derivative. Andreasen (2007) documents the robust performance of CMA-ES and compares it favorably with more common approaches as simulated annealing.

To reduce the "chatter" of the problem, we keep the innovations in the particle filter (that is, the draws from the exogenous shock distributions and the resampling probabilities) constant across different passes of the algorithm. As pointed out by McFadden (1989) and Pakes and Pollard (1989), this is required to achieve stochastic equicontinuity.

The standard errors reported below come from the bootstrapping procedure described by Efron and Tibshirani (1993, chapter 6). The estimated model is used to generate 100 artificial samples of data. These artificial series are used to re-estimate the model 100 times and the standard errors get computed as the standard deviations of the MLE taken across these 100 replications. This bootstrapping procedure accounts for the finite-sample properties of the

MLE and avoids the numerical instabilities that often appear while inverting the matrix of second derivatives of a likelihood function. These instabilities would be even more acute in our case since we are obtaining a non-differentiable approximation of the likelihood function.

With respect to Bayesian inference, the posterior of the model:

$$p\left(\Upsilon|\mathbb{Y}^T\right) \propto \frac{\mathcal{L}\left(\mathbb{Y}^T;\Upsilon\right)p\left(\Upsilon\right)}{\int \mathcal{L}\left(\mathbb{Y}^T;\Upsilon\right)p\left(\Upsilon\right)d\Upsilon},$$

is difficult, if not impossible, to characterize because the likelihood itself is only approximated by simulation. However, once we have an estimate of $\mathcal{L}\left(\mathbb{Y}^T;\Upsilon\right)$ thanks to the particle filter, we can draw from the posterior and build its empirical counterpart by using a Metropolis-Hastings algorithm. As mentioned before, we omit details to keep the paper focused.

5. Data and Main Results

5.1. Data

We take as our sample the period 1953.Q1 to 2008.Q4. Our output and consumption data come from the Bureau of Economic Analysis NIPA. 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 and consumption are defined as the ratio between our nominal output and consumption series and the civilian non-institutional population over 16. For inflation, we use the gross domestic product deflator. The data on bond yields are from CRSP Fama-Bliss discount bond files, which have fully taxable, non-callable, non-flower bonds. Fama and Bliss construct their data by interpolating observations from traded Treasuries. This procedure introduces measurement error, possibly correlated across time and cross-sectionally (although in our estimation, and just to reduce the number of parameters to maximize over, we do not allow for these correlations).

5.2. Summary Statistics

Table 1 reports the summary statistics from our data. Key observations are as follows. First, the volatility of output growth is higher than the volatility of consumption growth. Second, the yield curve is, on average, upward sloping. This points to a positive nominal bond risk premium. Third, the volatilities of bond yields are downward sloping for maturities of one year and longer. These are well-known facts and we will study how the model scores along

	Cons. gr.	Output gr.	Yields					Infl.	Hours
			1Y	2Y	3Y	4Y	5Y		
Mean	2.06%	1.67%	5.56%	5.76%	5.93%	6.06%	6.15%	3.43%	49.99%
St.dev.	1.96%	3.74%	2.91%	2.87%	2.80%	2.76%	2.72%	2.33%	1.12%
25%	0.98%	-0.22%	3.42%	3.63%	3.84%	3.99%	4.03%	1.79%	49.36%
50%	2.11%	1.84%	5.36%	5.45%	5.59%	5.65%	5.71%	2.76%	49.99%
75%	3.25%	3.82%	7.15%	7.31%	7.44%	7.57%	7.67%	4.46%	50.79%

Table 1: The table reports the summary statistics of consumption growth, output growth, bond yields, inflation, and hours worked. All statistics are expressed in annual terms. The sample period is 1953.Q1 to 2008.Q4.

these dimensions. Also, we do not include hours per capita in our observables because our model is not capable of generating enough fluctuations in hours. In any case, we want to put some restrictions on the behavior of the model-based hours. For this reason, we build a series of hours worked per capita using the index of total number of hours worked in the business sector and the civilian non-institutional population between 16 and 65. We normalize hours worked to have mean 0.5 during the sample period (this normalization level is per se irrelevant) and make v a function of the rest of the parameters such that, in steady state, hours worked in our model are always 0.5 for any value of the rest of the parameters.

5.3. Estimation Results

In this section, we report the parameter estimates and assess the extent to which the model can match the properties of the macro and yield data. To fully understand how the parameters are identified in our model, we estimate the model in three steps. First, we estimate the model using all data. Second, we estimate the model excluding inflation. Third, we use only bond yields. By studying which parameters change by changing information sets, we improve our understanding of which moments pin down which parameters.

Before proceeding, we fix a subset of the parameters. We do this because estimating a third-order approximation model, which as we argued before is important for identification, is extremely time consuming. Time constraints make it unfeasible, in practice, to estimate the whole set of parameters. Thus, in addition to the calibrated inflation parameters described above, we set $\lambda = 0.045$, $\zeta = 0.3$, and $\delta = 0.0294$. The value of λ is chosen to match the average growth rate of per capita output that we have in our sample. The values of ζ and δ are quite standard in the literature. Finally, we set the standard deviation of the measurement error shocks such that the model explains 75 percent of the standard deviation observed in the data.

5.3.1. Data set I: Consumption, output, bond yields, and inflation

We report our first findings in table 2. The table displays estimates of the parameters of the model. In the first column, we list our estimated parameters. In the second and third columns, we report the estimates and standard errors if we use consumption growth, output growth, five bond yields, and inflation in estimation. The fourth and fifth column report the results if we exclude inflation from the estimation. The last two columns contain the results if only the five bond yields are used in estimation.

Data	Cons. gr., Output gr., Yields, Inflation		Cons. gr., Output gr., Yields		Yields	
	MLE	Std.Error	MLE	Std. Error	MLE	Std.Error
β	0.994	0.0001	0.994	0.0001	0.994	0.0002
γ	79.34	12.234	88.23	10.157	96.75	20.125
ψ	1.731	0.2124	2.087	0.2348	1.775	0.4614
τ	0.032	0.0061	0.063	0.0071	0.026	0.0125
κ_0	-0.053	0.0088	-0.012	0.0045	-0.055	0.0124
ι	-0.522	0.1018	-0.174	0.0598	-0.175	0.0625
κ_1	-0.046	0.0093	0.235	0.124	0.102	0.0897
σ_ε	0.008	0.0009	0.008	0.0008	0.008	0.0012
σ_ω	0.002	0.0002	0.003	0.0003	0.003	0.0005

Table 2: Point Estimates and Standard Errors

We discuss now the result for the whole data set, which we take as our benchmark case, and explore the other columns in the subsections below. We start with the preference parameters. We estimate the discount factor, β, to be 0.994. This value, a relatively standard result in the literature, allows us to match the nominal yield level (remember that we have both inflation and long-run growth and that both factors affect the nominal yield level). The coefficient that controls risk aversion, γ, is estimated to be around 79, which is rather high.[19]

We estimate the IES to be 1.73. An estimate higher than one resonates with the parameter values picked in the long-run risks literature (for instance, Bansal and Yaron, 2004). Therefore, we find little support for the notion that the IES is around one, an assumption that is commonly used for convenience, as Campbell (1993), Tallarini (2000), and others do. This is not a surprise, because a value of the IES equal to one implies that the consumption-wealth ratio is constant over time. As we mentioned in the introduction, checking for this

---

[19]We need to be careful assessing this number, since our model includes leisure. It happens that, given the Cobb-Douglas specification of our utility aggregator of consumption and leisure, relative risk aversion is equal to γ. This does not need to be the case with other aggregators of consumption and leisure. See Swanson (2009) for a careful investigation.

implication of the model is hard because wealth is not directly observable, since it includes human wealth. However, different attempts at measurement, such as Lettau and Ludvigson (2001) or Lustig, van Nieuwerburgh, and Verdelhan (2007), reject the hypothesis that the ratio of consumption to wealth is constant.

The combination of the estimated values for the parameters controlling risk aversion and IES suggests:

$$\theta \equiv \frac{1 - \gamma}{1 - \frac{1}{\psi}} = \frac{1 - 79.34}{1 - \frac{1}{1.731}} = -185.51$$

indicating very different attitudes toward intertemporal substitution and toward substitution across states of nature. Moreover, since in our point estimate we have that $\gamma \gg \frac{1}{\psi}$, our representative household has a very strong preference for an early resolution of uncertainty. Finally, the high value of the IES, higher than one, generates, according to the formula we derived in section 3.1, a trivial welfare cost of aggregate fluctuactions.

The adjustment cost parameter, τ, is estimated to be 0.032, which indicates substantial adjustment costs. This estimate comes about because our data favor a situation in which capital cannot adjust easily to smooth consumption. When this is the case, the SDF fluctuates more and it is easier to match both premium and volatility of the yield curve. The volatility of the technology process, σ_ε, is 0.00756. This number is similar to many estimates in the literature and allows us to nicely match output and consumption volatility. Since the first-order approximation of our model behaves in the same way as the one from a simple real business cycle model, and this one is also able to match output and consumption properties, this finding is not a surprise.

The parameter controlling the MA component of the inflation process, ι, is well into negative terms, -0.522, and close to the value reported by Stock and Watson (2007), allowing us to capture the negative first-order autocorrelation and the small higher-order autocorrelations of inflation growth observed in the data.[20] Since the nominal yield curve slopes up in the data, κ_0 and κ_1 are estimated such that the correlation between innovations to inflation expectations and innovation to the stochastic discount factor expectations implies that inflation is bad news for consumption growth; that is, such that $(\rho\kappa_0 + \iota\kappa_1)\sigma_\varepsilon$ is negative (see, for a similar reason, Piazzesi and Schneider, 2006). The problem is that observed inflation volatility imposes a constraint on the maximum for the absolute value of $(\rho\kappa_0 + \iota\kappa_1)\sigma_\varepsilon$ and σ_ω (estimated to be 0.00201) and, hence, while we can match inflation volatility the model is barely able to generate an upward-sloping term structure. We will come back to this point

---

[20]Stock and Watson (2007) split their sample into two groups: 1960:Q1 to 1983:Q4 and 1984:Q1 to 2004:Q4. Their estimated values for ι are lower (in absolute value) for the first group and higher for the second. Our sample period 1953.Q1 to 2008.Q4 includes their two groups and, as expected, our estimate is right in the middle of their two estimates.

momentarily.

Table 3 displays means (panel A) and volatilities (panel B) of consumption growth, output growth, five bond yields, and inflation. In each panel, the first row displays the sample moments in the data. The second row corresponds to the estimates of the model for which we use all available data. The third row uses the estimates based on consumption growth, output growth, and five bond yields, but omit inflation data in estimation. The last row uses the estimates that we obtain using only bond yields in estimation. As before, the sample period for the data is 1953.Q1 to 2008.Q4.

Table 3 tells us that the model that uses all the data does a fair job at matching the mean of consumption growth and the average level of the yields. However, it has a bit more problems with output growth and with the average slope of the yields. The difference between the 5-year and 1-year yields amounts to 59 basis points in the data, whereas our model produces an average yield spread of only 17 basis points. Beeler and Campbell (2009) and Koijen *et al.* (2010) show that it is also a challenge to generate realistic nominal bond risk premia in standard long-run risks models. Furthermore, our estimated model does reasonably well with inflation volatility, but underestimates the volatility of bond yields by about a factor of two. Hence, our model has a difficult time jointly reproducing the salient features of the term structure of nominal interest rates and inflation.

Table 4 has a structure similar to table 3, but reports the autocorrelation of consumption growth (panel A), the 1-year bond yield (panel B), and inflation (panel C) for lag lengths varying from one quarter to ten quarters. The model is able to generate the autocorrelation patterns remarkably well. This is an advantage of a likelihood-based method, which tries to match the whole set of moments of the data, including the autocorrelations, instead of focusing on a limited set of moments, as the GMM.

5.3.2. Data set II: Consumption, output, and bond yields

To gain further insight into why the model does not generate a substantial bond risk premium and volatility of bond yields, we re-estimate our model using only parts of the data. We first omit the observations on inflation. In column 4 of table 2, we see how omitting inflation leads to an increase in the risk aversion coefficient and changes the estimates of the inflation parameters. This is because these parameters are no longer disciplined by observed inflation. In particular, κ_0 and κ_1 are estimated such that the absolute value of $(\rho\kappa_0 + \iota\kappa_1)\sigma_\varepsilon$ and σ_ω are twice as big. Table 3 shows that this leads to a dramatic improvement in terms of the bond risk premium and volatilities of bond yields. The model now replicates the observed bond risk premium and bond yield volatilities, at least for shorter maturities. This success of the model is accomplished at a cost. We now overestimate the volatility of inflation; it

Panel A: Means

	Cons. gr.	Output gr.	Yields					Inflation
			1Y	2Y	3Y	4Y	5Y	
Observed Data	2.06%	1.67%	5.56%	5.76%	5.93%	6.06%	6.15%	3.43%
All data	2.12%	2.11%	5.92%	5.98%	6.05%	6.09%	6.09%	3.67%
All data, but no inflation	2.12%	2.11%	5.63%	5.78%	5.92%	6.04%	6.13%	3.65%
Yields	2.12%	2.10%	5.54%	5.74%	5.90%	5.99%	6.11%	3.68%

Panel B: Volatilities

	Cons. gr.	Output gr.	Yields					Inflation
			1Y	2Y	3Y	4Y	5Y	
Data	1.96%	3.74%	2.91%	2.87%	2.80%	2.76%	2.72%	2.33%
All data	2.40%	2.91%	1.79%	1.64%	1.50%	1.38%	1.28%	2.32%
All data, but no inflation	2.54%	2.95%	3.28%	3.00%	2.75%	2.53%	2.33%	3.75%
Yields	2.32%	2.85%	3.31%	3.03%	2.78%	2.56%	2.46%	3.79%

Table 3: Means (Panel A) and volatilities (Panel B) of consumption growth, output growth, five bond yields, and inflation.

is 2.33% in the data, and the estimates imply a volatility of inflation of 3.75%. Omitting inflation data is inconsequential for matching the autocorrelation patterns in consumption growth, the 1-year bond yield, and inflation in table 4.

This exercise illustrates the importance of a joint estimation of inflation and structural parameters. Without the constraint of having to jointly match inflation and the yield curve, the model is sufficiently flexible to capture selected aspects of the data. This is an excellent example of how simple calibration exercises, by focusing on a set of moments selected by the researcher without tight discipline, are fraught with peril.

5.3.3. Data set III: Bond yields

As a last exercise, we estimate the model parameters using only information contained in bond yields. Perhaps surprisingly, the model estimates and their implications for the term structure are roughly unaffected if we omit consumption growth and output growth in estimation. The risk aversion parameter increases even further, to 96.75, and the adjustment cost falls to 0.026. The slope of the average nominal yield curve increases slightly We read this result as indicating that yield data (slopes and volatilities) carry a large amount of information about structural parameters of the economy, including the discount factor, risk aversion and the IES. This emphasizes the potentiality of incorporating finance data into the standard estimation

Panel A: Consumption growth

Lag length	1Q	2Q	3Q	4Q	5Q	6Q	7Q	8Q	9Q	10Q
Data	0.32	0.17	0.18	0.09	-0.03	0.06	0.01	-0.151	-0.04	0.01
All data	0.37	0.04	-0.01	-0.07	-0.05	-0.07	-0.03	-0.05	-0.06	0.02
All data, but no inflation	0.38	0.03	-0.01	-0.07	-0.06	-0.07	-0.03	-0.05	-0.06	0.02
Yields	0.48	0.07	-0.01	-0.08	-0.06	-0.07	-0.05	-0.06	-0.06	0.01

Panel B: 1-year bond yield

Lag length	1Q	2Q	3Q	4Q	5Q	6Q	7Q	8Q	9Q	10Q
Data	0.95	0.91	0.87	0.81	0.76	0.71	0.67	0.64	0.61	0.58
All data	0.96	0.92	0.88	0.83	0.79	0.75	0.7	0.66	0.62	0.58
All data, but no inflation	0.97	0.93	0.89	0.83	0.8	0.75	0.71	0.67	0.63	0.6
Yields	0.95	0.91	0.87	0.82	0.78	0.74	0.69	0.65	0.61	0.57

Panel C: Inflation

Lag length	1Q	2Q	3Q	4Q	5Q	6Q	7Q	8Q	9Q	10Q
Data	0.88	0.83	0.8	0.77	0.71	0.67	0.61	0.59	0.55	0.54
All data	0.80	0.77	0.74	0.69	0.66	0.63	0.58	0.54	0.50	0.47
All data, but no inflation	0.94	0.90	0.86	0.81	0.77	0.73	0.68	0.64	0.60	0.56
Yields	0.94	0.90	0.86	0.81	0.77	0.73	0.68	0.64	0.60	0.56

Table 4: Autocorrelation of consumption growth (Panel A), the 1-year bond yield (Panel B), and inflation (Panel C).

of DSGE models as a key additional source of information. Also, this result confirms Hall's (1988) intuition that a cross-section of asset yields is highly informative about the values of preference parameters.

6. Extensions

Despite some empirical shortcomings, our previous estimation has shown that a rich DSGE model with production and recursive preferences can be successfully taken to the data. Thus, we have opened the door to a large number of potential extensions. We discuss several that can be solved using our estimation procedure and that we believe might improve the fit of the model to the data. We leave them, though, for future work, since they will complicate the current paper, already a lengthy piece with much new content to digest.

Predictable technology growth   We assume technology growth is i.i.d., which might be too restrictive. We can extend the model to feature a predictable component in technology growth. Such a model is analyzed, for instance, in Croce (2006) and relates to the long-run risk literature (Bansal and Yaron, 2004, Hansen, Heaton, and Li, 2008, and Kaltenbrunner and Lochstoer, 2008).

Internal habit formation   In our specification of recursive preferences, the period utility is of the CRRA type. We can enrich the model to allow for habit formation in the period utility. Habit formation preferences have been successfully applied in asset pricing by for instance Constantinides (1990) and Campbell and Cochrane (1999).

Taylor rule   We assumed an exogenous process for inflation. A natural extension of our model is to endogenize the price process using, for instance, a Taylor rule as in Ang, Dong, and Piazzesi (2007). In this case, we not only embed the Taylor rule in an arbitrage-free term structure model, but also impose all equilibrium restrictions implied by an otherwise standard DSGE model.

Variable rare disasters   Gabaix (2009) shows that variable rare disasters might be a fruitful way to think about asset pricing in a production economy. Gabaix constructs a model in which the real business cycle properties of the model are unaffected relative to a standard model without rare disaster, but the asset pricing properties are improved substantially. We can enrich our model and estimate such models as well. This extension, however, would depend on our ability to have a perturbation method that can properly capture the effect of large, yet rare shocks.

7. Conclusions

We have studied the term structure of interest rates in a DSGE model in which the representative agent household has Epstein and Zin preferences. We have estimated the model by maximum likelihood using a solution method that perturbs the value function. Our estimation procedure, thus, imposes all economic restrictions implied by the equilibrium model.

Our paper has methodological and substantive contributions. Methodologically, we have shown how such a rich model can be accurately computed and estimated, thanks to the combination of perturbation methods and the particle filter. This leads the way for a large set of future applications. Our substantive findings are that the data indicate large levels of risk aversion, high levels of the IES, and high adjustment costs. The cross-equation restrictions imposed by the equilibrium of the model, in particular by the presence of endogenous physical capital accumulation, limits the ability of the model to jointly account for the slope of the nominal yield curve and the associated volatilities. However, we have pointed out a number of potential avenues of improvement that may solve this problem. All those can be explored for the first time in the context of a likelihood-estimated DSGE models that can move toward the integration of macro and finance observations with the tools that we have provided in this paper.

8. Appendices

In the next four appendices, we offer some further technical details about several parts of the paper. First, we show how to derive the SDF of the model. Second, we show that the value function representing the social planner's problem formulation of our model is homogeneous of degree ν. With these two results, in the third appendix, we write a stationary representation of the model. The last appendix explains how we maximize the resulting loglikelihood function.

8.1. Derivation of the SDF

First, to economize on notation, we rewrite the household's preferences as

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

where $\rho = 1/\psi$.

In the optimum, the household holds any asset with price p_t and payoff x_{t+1} such that:

$$
\frac{\partial}{\partial \xi} U_t \left(c_t - \xi p_t, c_{t+1} + \xi x_{t+1} \right) \bigg|_{\xi=0} = 0.
$$

Thus:

$$
\frac{\partial}{\partial c_t} U_t \left(c_t - \xi p_t, c_{t+1} + \xi x_{t+1} \right) p_t = \beta U_t^\rho \left(\mathbb{E}_t \left[U_{t+1}^{1-\gamma} \right] \right)_t^{\frac{1-\rho}{1-\gamma} - 1} \mathbb{E} \left[U_{t+1}^{-\gamma} \frac{\partial U_{t+1}}{\partial c_{t+1}} x_{t+1} \right].
$$

For the left-hand side, we have:

$$
\frac{\partial}{\partial c_t} U_t \left(c_t - \xi p_t, c_{t+1} + \xi x_{t+1} \right) =
$$

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

$$
= U_t^\rho \left(c_t^\nu (1 - l_t)^{1-\nu} \right)^{1-\rho} \nu c_t^{-1}.
$$

The optimality condition therefore implies (and switching from U_t to V_t to respect our notational convention):

$$
\left(c_t^\nu (1 - l_t)^{1-\nu} \right)^{1-\rho} \nu c_t^{-1} p_t
$$

$$
= \beta \left(\mathbb{E}_t \left[V_{t+1}^{1-\gamma} \right] \right)_t^{\frac{\gamma-\rho}{1-\gamma}} \mathbb{E} \left[V_{t+1}^{-\gamma} \left(V_{t+1}^\rho \left(c_{t+1}^\nu (1 - l_{t+1})^{1-\nu} \right)^{1-\rho} \nu c_{t+1}^{-1} \right) x_{t+1} \right]
$$

and hence:

$$p_t = \frac{\beta \left(\mathbb{E}_t \left[V_{t+1}^{1-\gamma}\right]\right)_t^{\frac{\gamma-\rho}{1-\gamma}} \mathbb{E}_t \left[V_{t+1}^{\rho-\gamma} \left(c_{t+1}^{\nu} \left(1-l_{t+1}\right)^{1-\nu}\right)^{1-\rho} \nu c_{t+1}^{-1} x_{t+1}\right]}{\left(c_t^{\nu} \left(1-l_t\right)^{1-\nu}\right)^{1-\rho} \nu c_t^{-1}}$$

$$= \mathbb{E}_t \left[\beta \left(\frac{c_{t+1}^{\nu} \left(1-l_{t+1}\right)^{1-\nu}}{c_t^{\nu} \left(1-l_t\right)^{1-\nu}}\right)^{1-\rho} \frac{c_{t+1}^{-1}}{c_t^{-1}} \left(\frac{V_{t+1}}{\mathbb{E}_t \left[V_{t+1}^{1-\gamma}\right]^{\frac{1}{1-\gamma}}}\right)^{\rho-\gamma} x_{t+1}\right].$$

The SDF is therefore given by:

$$M_{t+1} = \beta \left(\frac{c_{t+1}^{\nu} \left(1-l_{t+1}\right)^{1-\nu}}{c_t^{\nu} \left(1-l_t\right)^{1-\nu}}\right)^{1-\rho} \frac{c_t}{c_{t+1}} \left(\frac{V_{t+1}}{\mathbb{E}_t \left[V_{t+1}^{1-\gamma}\right]^{\frac{1}{1-\gamma}}}\right)^{\rho-\gamma}$$

$$= \beta \left(\frac{c_{t+1}^{\nu} \left(1-l_{t+1}\right)^{1-\nu}}{c_t^{\nu} \left(1-l_t\right)^{1-\nu}}\right)^{\frac{1-\gamma}{\theta}} \frac{c_t}{c_{t+1}} \left(\frac{V_{t+1}^{1-\gamma}}{\mathbb{E}_t \left[V_{t+1}^{1-\gamma}\right]^{\frac{1}{1-\gamma}}}\right)^{1-\frac{1}{\theta}}$$

which is the formula that we use in the main body of the paper.

8.2. Homotheticity

Taking advantage of the fact that the welfare theorems hold in our model, its equilibrium can be characterized by the solution of the social planner's problem:

$$V\left(k_t, z_t; \chi\right) = \max_{c_t, l_t} \left[\left(c_t^{\nu} \left(1-l_t\right)^{1-\nu}\right)^{\frac{1-\gamma}{\theta}} + \left(\mathbb{E}_t \left[V^{1-\gamma}\left(k_{t+1}, z_{t+1}; \chi\right)\right]\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}} \quad (12)$$

subject to:

$$c_t + k_{t+1} = k_t^{\zeta} \left(z_t l_t\right)^{1-\zeta} + \left(1-\delta\right) k_t + G\left(\frac{i_t}{k_t}\right) k_t.$$

As mentioned before, nothing of substance in the following argument depends on working with the social planner's problem, rather than with the competitive equilibrium. It just avoid having to deal with heavier notation (as we would need to do, for example, in models with nominal rigidities).

We now show that the value function is homothetic of degree ν, a result that we use in the main text to rewrite our problem in a stationary form. We build on the argument by Epstein and Zin (1989), Dolmas (1996), and specially Backus, Routledge, and Zin (2007).

First, note that for any $a > 0$, we can rework the resource constraint as:

$$\frac{c_t}{a} + \frac{k_{t+1}}{a} = \left(\frac{k_t}{a}\right)^\zeta \left(\frac{z_t}{a}l_t\right)^{1-\zeta} + (1-\delta)\frac{k_t}{a} + G\left(\frac{i_t}{k_t}\right)\frac{k_t}{a}.$$

For our purposes, it suffices to show that a function that is homogeneous of degree ν satisfies (12). If this is the case, by uniqueness of the solution to the Bellman equation, we know that V needs to be homogeneous of degree ν. Hence, consider a value function homogeneous of degree ν of the form:

$$\widetilde{V}(k_{t+1}, z_{t+1}; \chi) = a^\nu \widetilde{V}(k_{t+1}/a, z_{t+1}/a; \chi),$$

We plug this function into the Bellman equation:

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

$$= a^\nu \max_{c_t/a, l_t} \left[\left(\left(\frac{c_t}{a}\right)^\nu (1-l_t)^{1-\nu}\right)^{\frac{1-\gamma}{\theta}} + \left(\mathbb{E}_t \left[\widetilde{V}^{1-\gamma}(k_{t+1}/a, z_{t+1}/a; \chi)\right]\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}}$$

where the term after a^ν is the value function $\widetilde{V}(k_{t+1}/a, z_{t+1}/a; \chi)$ subject to:

$$\frac{c_t}{a} + \frac{k_{t+1}}{a} = \left(\frac{k_t}{a}\right)^\zeta \left(\frac{z_t}{a}l_t\right)^{1-\zeta} + (1-\delta)\frac{k_t}{a} + G\left(\frac{i_t}{k_t}\right)\frac{k_t}{a}.$$

This in turn implies:

$$V(k_t, z_t; \chi) = a^\nu V(k_t/a, z_t/a; \chi).$$

In other words, if we divide capital and productivity by a, we can divide consumption by a and still satisfy the Bellman equation. This shows that the value function is homogeneous of degree ν.

The homotheticity result can be strengthened to show that

$$V(k_t, z_t; \chi) = z_t^\nu V\left(\tilde{k}_t, 1; \chi\right) = z_t^\nu \widetilde{V}\left(\tilde{k}_t; \chi\right). \tag{13}$$

This stronger result can be useful in some investigations where we want to characterize the structure of the value function. This formulation would not make much difference in our perturbation approach because, in any case, we would need to take derivatives of (13) with respect to z_t to find the coefficients of the decision rule associated with z_t.

8.3. Stationary Recursive Form of the Model

Taking advantage of the result in appendices 8.1 and 8.2, we can easily make the model stationary by defining $\widetilde{var}_t = \frac{var_t}{z_{t-1}}$ for any nonstationary variable var_t. Note that, given the law of motion of productivity, we have that:

$$\widetilde{z}_t = \frac{z_t}{z_{t-1}} = \exp\left(\lambda + \chi\sigma_z\varepsilon_{zt}\right).$$

We can go equation by equation. First, the value function:

$$z_{t-1}^v V_t = \max\left[\left(z_{t-1}^v \widetilde{c}_t^v (1-l_t)^{1-v}\right)^{\frac{1-\gamma}{\theta}} + \beta\left(z_t^{v(1-\gamma)}\mathbb{E}_t V_{t+1}^{1-\gamma}\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}} \Rightarrow$$

$$V_t = \max\left[\left(\widetilde{c}_t^v (1-l_t)^{1-v}\right)^{\frac{1-\gamma}{\theta}} + \beta\widetilde{z}_t^{v\frac{1-\gamma}{\theta}}\left(\mathbb{E}_t V_{t+1}^{1-\gamma}\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}}.$$

Second, the optimality condition between leisure and consumption:

$$\frac{1-v}{v}\frac{z_{t-1}\widetilde{c}_t}{1-l_t} = z_{t-1}\widetilde{w}_t \Rightarrow \frac{1-v}{v}\frac{\widetilde{c}_t}{1-l_t} = \widetilde{w}_t$$

Third, the stochastic discount factor:

$$\frac{z_t}{z_{t-1}}\frac{\widetilde{\xi}_{t+1}}{\widetilde{\xi}_t} = \beta\left(\frac{z_t^v\widetilde{c}_{t+1}^v (1-l_{t+1})^{1-v}}{z_{t-1}^v\widetilde{c}_t^v (1-l_t)^{1-v}}\right)^{\frac{1-\gamma}{\theta}}\frac{z_{t-1}\widetilde{c}_t}{z_t\widetilde{c}_{t+1}}\left(\frac{z_t^{v(1-\gamma)}V_{t+1}^{1-\gamma}}{z_t^{v(1-\gamma)}\mathbb{E}_t V_{t+1}^{1-\gamma}}\right)^{1-\frac{1}{\theta}} \Rightarrow$$

$$\widetilde{z}_t\frac{\widetilde{\xi}_{t+1}}{\widetilde{\xi}_t} = \beta\widetilde{z}_t^{v\frac{1-\gamma}{\theta}-1}\left(\frac{\widetilde{c}_{t+1}^v (1-l_{t+1})^{1-v}}{\widetilde{c}_t^v (1-l_t)^{1-v}}\right)^{\frac{1-\gamma}{\theta}}\frac{\widetilde{c}_t}{\widetilde{c}_{t+1}}\left(\frac{V_{t+1}^{1-\gamma}}{\mathbb{E}_t V_{t+1}^{1-\gamma}}\right)^{1-\frac{1}{\theta}}$$

Fourth, the Euler equation for capital:

$$\left(\frac{z_{t-1}\widetilde{i}_t}{z_{t-1}\widetilde{k}_t}\right)^{\frac{1}{\tau}} = \mathbb{E}_t\left[\widetilde{z}_t\frac{\widetilde{\xi}_{t+1}}{\widetilde{\xi}_t}\left(a_2 r_{t+1} + \left(\frac{z_t\widetilde{i}_{t+1}}{z_t\widetilde{k}_{t+1}}\right)^{\frac{1}{\tau}}\left(1-\delta+a_1+\frac{a_2}{\tau-1}\left(\frac{z_t\widetilde{i}_{t+1}}{z_t\widetilde{k}_{t+1}}\right)^{1-\frac{1}{\tau}}\right)\right)\right] \Rightarrow$$

$$\left(\frac{\widetilde{i}_t}{\widetilde{k}_t}\right)^{\frac{1}{\tau}} = \mathbb{E}_t\left[\widetilde{z}_t\frac{\widetilde{\xi}_{t+1}}{\widetilde{\xi}_t}\left(a_2 r_{t+1} + \left(\frac{\widetilde{i}_{t+1}}{\widetilde{k}_{t+1}}\right)^{\frac{1}{\tau}}\left(1-\delta+a_1+\frac{a_2}{\tau-1}\left(\frac{\widetilde{i}_{t+1}}{\widetilde{k}_{t+1}}\right)^{1-\frac{1}{\tau}}\right)\right)\right]$$

Fifth, the Euler equation for nominal bonds:

$$R_t\mathbb{E}_t\widetilde{z}_t\frac{\widetilde{\xi}_{t+1}}{\widetilde{\xi}_t}\frac{1}{\pi_{t+1}} = 1$$

Sixth, output:

$$z_{t-1}\widetilde{y}_t = z_{t-1}^{\zeta}\widetilde{k}_t^{\zeta}\left(z_t l_t\right)^{1-\zeta} \Rightarrow \widetilde{y}_t = \frac{z_t^{1-\zeta}}{z_{t-1}^{1-\zeta}}\widetilde{k}_t^{\zeta}\left(l_t\right)^{1-\zeta} \Rightarrow \widetilde{y}_t = \widetilde{k}_t^{\zeta}\left(\widetilde{z}_t l_t\right)^{1-\zeta}$$

Seventh, the resource constraint:

$$z_{t-1}\widetilde{c}_t + z_{t-1}\widetilde{i}_t = z_{t-1}\widetilde{y}_t \Rightarrow \widetilde{c}_t + \widetilde{i}_t = \widetilde{y}_t$$

Eight, law of motion for capital:

$$z_t\widetilde{k}_{t+1} = (1-\delta)\, z_{t-1}\widetilde{k}_t + \left(a_1 + \frac{a_2}{1-\frac{1}{\tau}}\left(\frac{z_{t-1}\widetilde{i}_t}{z_{t-1}\widetilde{k}_t}\right)^{1-\frac{1}{\tau}}\right) z_{t-1}\widetilde{k}_t \Rightarrow$$

$$\widetilde{z}_t\widetilde{k}_{t+1} = (1-\delta)\,\widetilde{k}_t + \left(a_1 + \frac{a_2}{1-\frac{1}{\tau}}\left(\frac{\widetilde{i}_t}{\widetilde{k}_t}\right)^{1-\frac{1}{\tau}}\right)\widetilde{k}_t$$

Ninth, and finally, input prices:

$$z_{t-1}\widetilde{w}_t = (1-\zeta)\,\frac{z_{t-1}\widetilde{y}_t}{l_t} \Rightarrow \widetilde{w}_t = (1-\zeta)\,\frac{\widetilde{y}_t}{l_t}$$

$$r_t = \zeta\frac{z_{t-1}\widetilde{y}_t}{z_{t-1}\widetilde{k}_t} \Rightarrow r_t = \zeta\frac{\widetilde{y}_t}{\widetilde{k}_t}$$

Collecting all terms:

$$V_t = \max\left[\left(\widetilde{c}_t^{\upsilon}\left(1-l_t\right)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}} + \beta\widetilde{z}_t^{\upsilon\frac{1-\gamma}{\theta}}\left(\mathbb{E}_t V_{t+1}^{1-\gamma}\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}},$$

$$\frac{1-\upsilon}{\upsilon}\frac{\widetilde{c}_t}{1-l_t} = \widetilde{w}_t,$$

$$\widetilde{z}_t\frac{\widetilde{\xi}_{t+1}}{\widetilde{\xi}_t} = \beta\widetilde{z}_t^{\upsilon\frac{1-\gamma}{\theta}-1}\left(\frac{\widetilde{c}_{t+1}^{\upsilon}\left(1-l_{t+1}\right)^{1-\upsilon}}{\widetilde{c}_t^{\upsilon}\left(1-l_t\right)^{1-\upsilon}}\right)^{\frac{1-\gamma}{\theta}}\frac{\widetilde{c}_t}{\widetilde{c}_{t+1}}\left(\frac{V_{t+1}^{1-\gamma}}{\mathbb{E}_t V_{t+1}^{1-\gamma}}\right)^{1-\frac{1}{\theta}},$$

$$\left(\frac{\widetilde{i}_t}{\widetilde{k}_t}\right)^{\frac{1}{\tau}} = \mathbb{E}_t\left[\widetilde{z}_t\frac{\widetilde{\xi}_{t+1}}{\widetilde{\xi}_t}\left(a_2 r_{t+1} + \left(\frac{\widetilde{i}_{t+1}}{\widetilde{k}_{t+1}}\right)^{\frac{1}{\tau}}\left(1-\delta+a_1+\frac{a_2}{\tau-1}\left(\frac{\widetilde{i}_{t+1}}{\widetilde{k}_{t+1}}\right)^{1-\frac{1}{\tau}}\right)\right)\right],$$

$$R_t\mathbb{E}_t\widetilde{z}_t\frac{\widetilde{\xi}_{t+1}}{\widetilde{\xi}_t}\frac{1}{\pi_{t+1}} = 1,$$

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

$$\widetilde{c}_t + \widetilde{i}_t = \widetilde{y}_t,$$

$$\widetilde{z}_t \widetilde{k}_{t+1} = (1 - \delta) \widetilde{k}_t + \left(a_1 + \frac{a_2}{1 - \frac{1}{\tau}} \left(\frac{\widetilde{i}_t}{\widetilde{k}_t} \right)^{1 - \frac{1}{\tau}} \right) \widetilde{k}_t,$$

$$\widetilde{w}_t = (1 - \zeta) \frac{\widetilde{y}_t}{l_t},$$

$$r_t = \zeta \frac{\widetilde{y}_t}{\widetilde{k}_t},$$

and

$$\widetilde{z}_t = \exp\left(\lambda + \chi \sigma_z \varepsilon_{zt} \right)$$

Now, we can write:

$$\frac{1}{R_t} \left(\mathbb{E}_t V_{t+1}^{1-\gamma} \right)^{\frac{1}{\theta} - 1} = \mathbb{E}_t \beta \widetilde{z}_t^{v \frac{1-\gamma}{\theta} - 1} \left(\frac{\widetilde{c}_{t+1}^v (1 - l_{t+1})^{1-v}}{\widetilde{c}_t^v (1 - l_t)^{1-v}} \right)^{\frac{1-\gamma}{\theta}} \frac{\widetilde{c}_t}{\widetilde{c}_{t+1}} \left(V_{t+1}^{1-\gamma} \right)^{1 - \frac{1}{\theta}} \frac{1}{\pi_{t+1}} = 1.$$

8.4. Optimizing the Loglikelihood

The estimation of the model was done with mixed-programming as follows. `Mathematica` computed the analytical derivatives of the value function and decision rules and generated `Fortran 95` code that included those 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 a Markov chain Monte Carlo, as a random-walk Metropolis-Hastings. The `Fortran 95` code was compiled in `Intel Visual Fortran 10.3` to run on Windows-based machines. We used a Xeon Processor 5160 EMT64 at 3.00 GHz with 16 GB of RAM.

As we pointed out in the main text, the CMA-ES is an evolutionary algorithm that approximates the inverse of the Hessian of the loglikelihood function by simulation. We iterate in the procedure until the change in the objective function is lower than some tolerance level. In each step g of the routine, m candidate $n-$parameter values are proposed from a normal distribution:

$$\Upsilon_i^g \sim \mathcal{N} \left(\Upsilon_\mu^{g-1}, (\sigma^g)^2 C_\mu^{g-1} \right), \text{ for } i = 1, ..., m$$

where $\Upsilon_\mu^{g-1} \in R^n$ and $C_\mu^{g-1} \in R^n \times R^n$ are the mean and variance-covariance matrix of the μ best candidates for optimal parameter values in step $g - 1$ and σ^g is a scaling parameter. The normal distribution can be truncated to have support only on that part of the parameter space where the parameters take admissible values (for example, positive discount factors). To save on notation, we re-order the draws in decreasing relation to the value they attain in

the likelihood of the model:

$$\mathcal{L}\left(\mathbb{Y}^{T}; \Upsilon_{i}\right) \geq \mathcal{L}\left(\mathbb{Y}^{T}; \Upsilon_{i+1}\right)$$

The mean of step g is defined as:

$$\Upsilon_{\lambda}^{g} = \sum_{i=1}^{\mu} w_{i} \Upsilon_{i}^{g}$$

where the weights w_{i} are defined between 0 and 1 (and its sum normalizes to 1) and μ is smaller than m.

The variance-covariance of step g fits the search distribution to the contour lines of the likelihood function. To do so, we set:

$$
\begin{aligned}
C_{\lambda}^{g} = & \underbrace{\left(1 - c_{cov}\right) C_{\lambda}^{g-1}}_{\text{term 1}} + \\
& + \underbrace{\frac{c_{cov}}{\mu_{cov}} \left(P_{c}^{g} \left(P_{c}^{g}\right)' + c_{c} \left(2 - c_{c}\right) \left(1 - H^{g}\right)\right)}_{\text{term 2}} \\
& + \underbrace{c_{cov} \left(1 - \frac{1}{\mu_{cov}}\right) \sum_{i=1}^{\mu} \frac{w_{i}}{\left(\sigma^{g}\right)^{2}} \left(\Upsilon_{1:\mu}^{g} - \Upsilon_{\mu}^{g-1}\right) \left(\Upsilon_{1:\mu}^{g} - \Upsilon_{\mu}^{g-1}\right)'}_{\text{term 3}}
\end{aligned}
\tag{14}
$$

where:

$$P_{c}^{g} = \left(1 - c_{c}\right) P_{c}^{g-1} + H^{g} \sqrt{c_{c} \left(2 - c_{c}\right) \mu_{eff}} \left(\frac{\Upsilon_{\lambda}^{g} - \Upsilon_{\lambda}^{g-1}}{\sigma^{g}}\right)$$

$$H^{g} = \begin{cases} 1 \text{ if } \frac{\|P_{c}^{g}\|}{\sqrt{1-(1-c_{g})^{2g}}} < \left(1.5 + \frac{1}{n-0.5}\right) \mathbb{E}\left(\|\mathcal{N}\left(\mathbf{0}, \mathbf{I}\right)\|\right) \\ 0 \text{ otherwise} \end{cases}$$

and $c_{c}, c_{cov}, \mu_{eff}$, and μ_{cov} are constants. Term 1 of (14) is a standard persistence term from previous variance-covariance matrix to dampen changes in C_{λ}^{g}. Term 2 captures the correlation across steps of the algorithm through the evolution of P_{c}^{g}. Term 3 controls for a large number of points in the simulation. The term

$$\mathbb{E}\left(\|\mathcal{N}\left(\mathbf{0}, \mathbf{I}\right)\|\right) = \sqrt{2}\Gamma\left(\frac{n+1}{2}\right) \Big/ \Gamma\left(\frac{n}{2}\right) \approx \sqrt{n} + \mathcal{O}\left(\frac{1}{n}\right)$$

is the expectation of the euclidean norm of a random normal vector.

Finally, the scaling parameter evolves according to:

$$\sigma^g = \sigma^{g-1} \exp\left[\frac{c_\sigma}{d_\sigma}\left(\frac{\|P_\sigma^g\|}{\mathbb{E}\left(\|\mathcal{N}\left(\mathbf{0},\mathbf{I}\right)\|\right)} - 1\right)\right]$$

$$P_\sigma^g = (1 - c_\sigma) P_\sigma^{g-1} + \sqrt{c_\sigma\left(2 - c_\sigma\right)} B^{g-1} \left(D^{g-1} - 1\right)^{-1} \left(B^{g-1}\right)' \frac{\sqrt{\mu_{eff}}}{\sigma^g} \left(\Upsilon_\mu^g - \Upsilon_\mu^{g-1}\right)$$

where c_σ and d_σ are constants and B^{g-1} is an orthogonal matrix and D^{g-1} a diagonal matrix such that $C_\lambda^g = B^g \left(D^g\right)^2 \left(B^g\right)'$.

Standard values for the constants of the algorithm are:

$$m = 4 + \lfloor 3 \ln n \rfloor$$

$$\mu = \lfloor m/2 \rfloor$$

where $\lfloor \cdot \rfloor$ is the integer floor of a real number, and

$$w_i = \frac{\ln\left(\mu + 1\right) - \ln i}{\sum_{j=1}^n \left(\ln\left(\mu + 1\right) - \ln j\right)}$$

$$\mu_{eff} = \mu_{cov} = \left(\sum_{i=1}^\mu w_i^2\right)^{-1}$$

$$c_\sigma = \frac{\mu_{eff} + 2}{n + \mu_{eff} + 3}$$

$$c_c = \frac{4}{n + 4}$$

$$c_{cov} = \frac{1}{\mu_{cov}} \frac{2}{\left(n + \sqrt{2}\right)^2} + \left(1 - \frac{1}{\mu_{cov}}\right) \min\left(1, \frac{2\mu_{eff} - 1}{\left(n + 2\right)^2 + \mu_{eff}}\right)$$

$$d_\sigma = 1 + 2\max\left(0, \sqrt{\frac{\mu_{eff} - 2}{n + 1}} - 1\right) \max\left(0.3, 1 - \frac{n}{\min\left(g^{\max}, \frac{\mathcal{L}_{eval}^{\max}}{m}\right)}\right) + c_\sigma$$

where $g^{\max}$ is the maximum number of steps and $L_{eval}^{\max}$ is the maximum number of likelihood evaluations. Finally, we can initialize the algorithm by setting Υ_λ^0 to some standard calibrated parameters, σ^0 to 1, $P_c^g = 0$, and Σ_μ^0 to an identity matrix. Of all these constants, Hansen and Kern (2004) recommend only to change m to adapt the algorithm to particular problems. As one could have guessed, Hansen and Kern show that by increasing the number of simulations m, the global properties of the search procedure improve.

References

[1] Álvarez, F. and U.J. Jermann (2005). "Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth." *Econometrica* 73, 1977-2016.

[2] Andreasen, M. (2007). "How to Maximize the Likelihood of a DSGE Model." *Mimeo*, Bank of England.

[3] Andreasen, M. and P. Zabczyk (2010). "An Efficient Method of Computing Higher Order Bond Price Perturbation Approximations." *Mimeo*, Bank of England.

[4] Ang, A., G. Bekaert, and M. Wei (2008). "The Term Structure of Real Rates and Expected Inflation." *Journal of Finance*, 63, 797-849.

[5] Ang, A., S. Dong, and M. Piazzesi (2007). "No-Arbitrage Taylor Rules." *Mimeo*, Columbia University.

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

[7] Backus, D.K., B.R. Routledge, and S.E. Zin (2007). "Asset Pricing in Business Cycle Analysis." *Mimeo*, New York University.

[8] Bansal, R., Dittman, and D. Kiku (2007). "Cointegration and Consumption Risks in Asset Returns." *Review of Financial Studies* 22, 1343 - 1375.

[9] Bansal, R., R. Gallant, and G. Tauchen (2008). "Rational Pessimism, Rational Exuberance, and Asset Pricing Models." *Review of Economic Studies* 74, 1005-1033.

[10] Bansal, R., D. Kiku, and A. Yaron (2007). "Risks for the Long Run: Estimation and Inference." *Mimeo*, Duke University.

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

[12] Beeler, J. and J.Y. Campbell (2009). "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment." *NBER Working Paper* 14788.

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

[14] Binsbergen, J. van (2009). "Deep Habits and the Cross Section of Expected Returns." *Mimeo*, Stanford University.

[15] Binsbergen, J. van and R.S.J. Koijen (2010). "Likelihood-based Estimation of Exactly-Solved Present-Value Models." *Mimeo*, University of Chicago and Stanford University.

[16] Caldara, D., J. Fernández-Villaverde, J.F. Rubio-Ramírez, and W. Yao (2010). "Computing Models with Recursive Preferences." *Mimeo*, University of Pennsylvania.

[17] Campanale, C., R. Castro, and G. Clementi (2010). "Asset Pricing in a Production Economy with Chew-Dekel Risk Preferences." *Review of Economic Dynamics* 13, 379-402.

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

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

[20] Campbell, J.Y. and J. Cochrane (1999). "Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy* 107, 205-251.

[21] Campbell, J.Y. and J. Cochrane (2000). "Explaining the Poor Performance of Consumption-Based Asset Pricing Models." *Journal of Political Economy* 55, 2863-78.

[22] Campbell, J.Y. and L. Viceira (2001). "Who Should Buy Long-Term Bonds?" *American Economic Review* 91, 99-127.

[23] Chen, X., J. Favilukis and S. Ludvigson (2007). "An Estimation of Economic Models with Recursive Preferences." *Mimeo*, New York University.

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

[25] Christiano, L., M. Eichenbaum, and C.L. Evans (2005). "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113, 1-45.

[26] Cochrane, J.H. and M. Piazzesi (2005). "Bond Risk Premia." *American Economic Review* 95, 138-160.

[27] Cochrane, J.H. and M. Piazzesi (2008). "Decomposing the Yield Curve." *Mimeo*, University of Chicago.

[28] Constantinides, G.M. (1990). "Habit Formation: A Resolution of the Equity Premium Puzzle." *Journal of Political Economy* 98 , 519-543.

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

[30] Dai, Q. and K. Singleton (2000). "Specification Analysis of Affine Term Structure Models." *Journal of Finance* 55, 1943-1978.

[31] Dai, Q. and K. Singleton (2002). "Expectations Puzzle, Time-varying Risk Premia, and Affine Models of the Term Structure"" *Journal of Financial Economics* 63, 415-441.

[32] 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 Communications, American Mathematical Society.

[33] De Paoli, B., A. Scott, and O. Weeken (2007). "Asset Pricing Implications of a New Keynesian Model." *Bank of England* Working Paper 326.

[34] DeJong, D.N., H. Dharmarajan, R. Liesenfeld, and J.F. Richard (2007). "An Efficient Filtering Approach to Likelihood Approximation for State-Space Representations." *Mimeo*, University of Pittsburgh.

[35] Doh, T. (2009). "Yield Curve in an Estimated Nonlinear Macro Model." *Mimeo*, Federal Reserve Bank of Kansas City.

[36] Dolmas, J. (1996). "Balanced-growth-consistent Recursive Utility." *Journal of Economic Dynamics and Control* 204, 657-680.

[37] Duffee, G.R. (2002). "Term Premia and Interest Rate Forecasts in Affine Models." *Journal of Finance* 57, 405-443

[38] Efron. B. and R.J. Tibshirani (1993). *An Introduction to the Bootstrap.* Chapman & Hall.

[39] 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.

[40] 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.

[41] Fermanian, J.D. and B. Salanié (2004). "A Nonparametric Simulated Maximum Likelihood Estimation Method." *Econometric Theory* 20, 701-734.

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

[43] Fiorentini, G., E. Sentana, and N. Shephard (2004). "Likelihood Estimation of Latent Generalised ARCH Structures." *Econometrica* 72, 1481-1517.

[44] Gabaix, X. (2009). "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance." *Mimeo*, NYU.

[45] Gomes, F. and A. Michealides (2005). "Optimal Life-Cycle Asset Allocation: Understanding the Empirical Evidence." *Journal of Finance* 60, 869-904.

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

[47] Hansen, L.P., J. Heaton, and N. Li (2008). "Consumption Strikes Back? Measuring Long Run Risk." *Journal of Political Economy*, 116, 260-302.

[48] Hansen, L.P., J. Heaton, J. Lee, and N. Roussanov (2008). "Intertemporal Substitution and Risk Aversion." In J.J. Heckman and E. Leamer (eds), *Handbook of Econometrics*, volume 6, 3967-4056. North-Holland.

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

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

[51] Hansen, N., and S. Kern (2004). "Evaluating the CMA Evolution Strategy on Multimodel Test Functions." *Springer-Verlag: Eighth International Conference on Parallel Problems from Nature,* 282-291.

[52] Hansen, N., S.D. Müller, and P. Koumoutsakos (2003). "Reducing the Time Complexity of the Derandomized Evolution Strategy with Covariance Matrix Adaptation (CMA-ES)." *Evolutionary Computation* 11, 1-18.

[53] Jermann, U.J. (1998). "Asset Pricing in Production Economies." *Journal of Monetary Economics* 41, 257-275.

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

[55] Kaltenbrunner, G. and L. Lochstoer (2008). "Long-Run Risk through Consumption Smoothing." *Review of Financial Studies,* forthcoming.

[56] Kim, S., N. Shephard, and S. Chib (1998). "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models." *Review of Economic Studies* 65, 361-93.

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

[58] Koijen, R.S.J., H. Lustig, S. Van Nieuwerburgh, and A. Verdelhan (2010). "Long-Run Risk, the Wealth-Consumption Ratio, and the Temporal Pricing of Risk." *American Economic Review,* forthcoming.

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

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

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

[62] Lettau, M. and H. Uhlig (2002). "Sharpe Ratios and Preferences: An Analytical Approach." *Macroeconomic Dynamics* 6, 242-265.

[63] Levin, A., D. López-Salido, and T. Yun (2007). "Risk-Sensitive Monetary Policy." *Mimeo,* Federal Reserve Board.

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

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

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

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

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

[69] Rouwenhost, K.G. (1995). "Asset Pricing Implications of Equilibrium Business Cycle Models." In Cooley, T. F. (Eds.), *Frontiers of Business Cycle Research*. Princeton University Press, Princeton, 293-330.

[70] 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.

[71] Rudebusch, G. and E. Swanson (2008). "The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks." *Federal Reserve Bank of San Francisco Working Paper* 2008-31.

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

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

[74] Smets, F. and R. Wouters (2007). "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 97, 586-606.

[75] 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.

[76] Swanson, E. (2009). "Risk Aversion, the Labor Margin, and Asset Pricing in DSGE Models." *Federal Reserve Bank of San Francisco* Working Paper 2009-26.

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

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

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

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