# Bond pricing with a time-varying price of risk in an estimated medium-scale Bayesian DSGE model*

Ian Dew-Becker

November 12, 2011

**Abstract**

A New-Keynesian model in which households have Epstein–Zin preferences with time-varying risk aversion and the central bank has a time-varying inflation target can match the dynamics of nominal bond prices in the US economy well. The model generates a steady-state term spread of 152 basis points, compared to the sample average of 207 basis points. The fitting errors for individual bond yields are roughly as large as those obtained from a non-structural three-factor model, and two thirds smaller than in models with constant risk aversion or a constant inflation target. The term premium is estimated to have a standard deviation half as large as that of the term spread.

The model delivers rich variance decompositions for the pricing kernel and the real side of the economy. Shocks to risk aversion account for less than 5 percent of the variation in output, consumption, and investment growth at business-cycle frequencies, but 32 percent of the variation in the term spread. There is little connection between priced risk factors and output, consumption, investment, or hours worked in the short-run.

## 1 Introduction

Nonstructural models are widely used in both macroeconomics and the study of the term structure of interest rates. Recently, Smets and Wouters (2003) have shown that a structural New Keynesian model can match the dynamics of the macroeconomy as well as or better than a VAR. This paper extends that

*I appreciate helpful comments from John Campbell, Emmanuel Farhi, and Andrea Tambalotti. Jason Beeler made important contributions on an earlier version of this paper.

work by showing that a suitably augmented version of their model can also match the dynamics of the term structure of interest rates as well as a standard non-structural model.

Bekaert, Cho, and Moreno (2010) show that a log-linearized macro model naturally also delivers closed-form expressions for bond prices. Their method, however, is not able to describe risk premia, and even if it could, the model assumes that risk premia are constant. This paper builds on their work by using an approximation method that allows for positive and time-varying risk premia. I then estimate the model using Bayesian methods, and show that it fits interest rates with errors that are similar to those generated by a non-structural three-factor model. The errors in fitting annualized yields on bonds with maturities ranging from 1 quarter to 10 years have a standard deviation of 8 basis points.

For the production side of the economy, I take the model described in Justiniano, Primiceri, and Tambalotti (JPT; 2010) and combine it with a preference specification that endogenously generates the essentially affine stochastic discount factor of Duffee (2002). Households are assumed to have Epstein–Zin preferences with time-varying risk aversion as in Dew-Becker (2011a), which induces a time-varying price of risk. I also allow the central bank to have a time-varying inflation target, movements in which shift the entire term structure, inducing a so-called level factor.

The steady-state term spread in the model simply represents the average risk premium on long-term bonds. The steady-state term spread is estimated to be 152 basis points, similar in magnitude to the 207-basis-point average observed in the sample. To understand why that risk premium would be large, we first need to understand what drives the variance of the pricing kernel. When the representative household has Epstein–Zin preferences with a coefficient of relative risk aversion that is substantially larger than the inverse of its EIS (i.e. it has a relatively high EIS), state prices are almost entirely driven by innovations to the household's lifetime utility, i.e. the value placed on its entire future consumption stream. With a high EIS, transitory shocks have a small effect on lifetime utility. Permanent technology shocks, though, will have large effects. Shifts in risk aversion also affect lifetime utility because they affect how much the household penalizes future uncertainty.

Even though there are nine shocks in the economy, only two of them turn out to be relevant for the pricing kernel—labor-neutral technology and risk aversion. Since all of the other shocks (e.g. monetary policy, markups, government spending) are purely transitory, they have little effect on permanent income or welfare (because the household is estimated to have a relatively high EIS of 1.33), and thus they do not have a strong effect on state prices.

Following a positive innovation to the level of technology, interest rates fall. This result turns out to be common to a variety of New-Keynesian models, e.g. JPT, Smets and Wouters (2004), and Christiano, Trabandt, and Walentin (2011). In this paper, the reason is that the central bank's inflation target is estimated to fall following positive technology shocks. Intuitively, a positive supply shock lowers inflationary pressure, and the central bank takes this as an opportunity to drive inflation lower for an extended period. The fact that the negative correlation between technology shocks and interest rates is obtained in numerous other models that assume a constant inflation target suggests that this is in fact a well-identified feature of the data.

Variation in risk aversion also turns out to make an important contribution to the model's ability to the term structure of interest rates, though. Standard statistical tests easily reject a model with constant risk aversion in favor of one with time-varying risk aversion. The pricing errors for bonds are smaller by a factor of three when risk aversion is allowed to vary over time.

While the variance decompositions imply that the pricing kernel is driven entirely by the labor-neutral technology and risk aversion shocks, it turns out that those two shocks have only minor effects on the dynamics of the real economy. Risk aversion explains less than 5 percent of the variance of output, consumption, investment, and hours worked at business cycle frequencies. The variable that is most responsive to the technology shock is hours worked, and the technology shock still explains only 25 percent of its variance. The variance decompositions also differ substantially from the results found by JPT. Whereas JPT find that investment technology shocks are an important driver of the business cycle, I find that they explain little other than investment, and monetary policy and markup shocks play much bigger roles. This finding suggests that including information about bond prices in estimation has important effects on estimation results.

In addition to matching the behavior of the term structure, the estimated parameters imply reasonable behavior for equity prices. The average annualized Hansen–Jagannathan bound is estimated to be 0.53, which is consistent with the observed Sharpe ratio for the stock market in the data sample, even though data on equity returns is not included in the estimation. Furthermore, the estimated degree of variation in risk aversion is similar to (though somewhat higher than) the value used in Dew-Becker (2011a), who calibrates a general-equilibrium model that can match the both the average Sharpe ratio on equities and also empirical stock return forecasting regressions. At business-cycle frequencies, estimated risk aversion displays similar behavior to Cochrane and Piazessi's (2005) tent-shaped bond return forecasting factor (and they are both strongly correlated with the term spread).

This paper is related to a small but growing literature on bond pricing in production economies. Bekaert, Cho, and Moreno (2010) and Doh (2011) estimate New-Keynesian macro models, but they do not focus on the size and volatility of the term premium, whereas that is the feature of the term structure that this paper concentrates on. Rudebusch and Swanson (2011) generate a large and volatile term premium in a calibrated model. This paper moves beyond them by considering a substantially more complex model and showing that it can be estimated through standard Bayesian methods using the Kalman filter. Models of the business cycle have strong implications for the term structure of interest rates, so adding that information can have strong effects on estimation results. For example, I find that when the model is estimated without bond price information, the shock to investment technology is estimated to account for a large fraction of the variance of short-term interest rates and the term spread. But when the term spread is included as part of the information set, the effects of investment technology shocks are much smaller. The implications of the model for wage-setting also change substantially when interest rates are added; I estimate a substantially larger Frisch elasticity than JPT do, coming in closer agreement with micro evidence.

The remainder of the paper is organized as follows. Section 2 describes household preferences and derives the pricing kernel. Section 3 describes the remainder of the economy including the production process, price setting, and monetary and fiscal policy.

Next, section 4 explains how the model is solved. If we used perturbation methods, a third-order approximation would be necessary to capture time-variation in risk premia. The estimation of the model turns out to be sufficiently difficult, however (due to numerous local extrema in the likelihood function, a common feature of models of the term structure), that the use of a nonlinear filter for calculating the model's marginal likelihood is infeasible. I therefore use the essentially affine solution method described in Dew-Becker (2011b). The method approximates the pricing kernel separately from the remainder of the model, allowing it to take the essentially affine form with a time-varying price of risk described in Duffee (2002). The essentially affine method is equivalent to a first-order perturbation local to the non-stochastic steady-state, but it includes corrections for volatility that allow it to substantially outperform perturbation in stochastic simulations. The key feature of the essentially affine method is that risk premia may vary over time and affect real variables, not just asset prices.

Section 5 describes the Bayesian methods used to estimate the model. Sections 6 and 7 examine the implications of the estimates for asset prices and the dynamics of the real economy, respectively. Finally, section 8 concludes.

# 2 Household preferences

## 2.1 Objective function and budget constraint

The household's value function is assumed to be

$$V_t = \left\{ (1 - \beta B_t) U \left( C_t, \bar{C}_{t-1}, N_t, Z_t \right) + \beta B_t \left( E_t V_{t+1}^{1-\alpha_t} \right)^{\frac{1-\rho}{1-\alpha_t}} \right\}^{\frac{1}{1-\rho}} \tag{1}$$

where $C_t$ is consumption, $\bar{C}_t$ is aggregate consumption, $N_t$ is the number of hours worked outside the home, and $E_t$ denotes the expectation operator conditional on information available at date $t$. The term $\bar{C}_{t-1}$ allows the period utility function to potentially include external habit formation. The level of technology, $Z_t$, may also affect household utility in order to ensure balanced growth (as in Rudebusch and Swanson, 2010).

$B_t$ is an exogenous shock to the household's rate of time preference. The choice of exactly how to specify this preference shock is not trivial. The goal is to generate variation in consumption demand conditional on the level of interest rates. However, because $B_t$ enters the value function, it may also affect the level of $V_t$, and hence asset prices. The specification (1) has the feature that if period utility, $U \left( C_t, \bar{C}_{t-1}, N_t, Z_t \right)$, is constant over time, then a change in $B_t$ will have no effect on $V_t$. So in some sense it purely affects the relative preference for consumption today versus in the future, as opposed to also affecting the household's overall level of welfare.[1]

The household's coefficient of relative risk aversion, $\alpha_t$, is allowed to vary over time. Dew-Becker (2011a) motivates variation in $\alpha_t$ by considering adding a time-varying benchmark to the standard Epstein–Zin certainty equivalent, $E_t \left( V_{t+1} - H_t \right)^{1-\alpha}$. When $V_{t+1}$ is close to $H_t$, the household's effective risk aversion over shocks to $V_{t+1}$ rises. The formulation (1) has the advantage that it is log-linear and we do not have to worry about the possibility that $V_{t+1}$ falls below $H_t$. In Dew-Becker (2011a), movements in $\alpha_t$ are connected to movements in the household's welfare. I loosen that constraint here and allow for independent shocks to risk aversion (equivalently, independent shocks to the habit). Melino and Yang (2003) study a

[1] Variance decompositions for the estiamted model reported below confirm that shocks to $B_t$ have essentially no effect on the pricing kernel.

similar specification, but without the emphasis on the habit.

The household's budget constraint is

$$P_t C_t + H_t + D_t = (1 + i_t) H_{t-1} + W_t N_t + D_{t-1} + \Pi_t \tag{2}$$

where $P_t$ is the price of the consumption good, $H_t$ is holdings of one-period nominal bonds, $D_t$ is cash holdings, $i_t$ is the nominally riskless interest rate, $W_t$ is the wage, and $\Pi_t$ represents profits and other lump-sum transfers paid to the household.

For the sake of simplicity, I study the so-called cashless economy described in Woodford (2003). The monetary authority is able to control the interest rate because money enters the household's utility function, but the effect of money on total utility is sufficiently small that we can ignore it when writing $V$ (i.e. we take the limit where the relative importance of money goes to zero). I do not discuss money any further and from now on drop it from the household's budget constraint.

## 2.2 The stochastic discount factor

As usual with Epstein–Zin preferences, it is possible in this setting to obtain an expression for the stochastic discount factor (SDF) involving consumption growth and an asset return. However, the asset whose return enters the SDF is no longer the household's total wealth portfolio: it is now an asset that pays a dividend depending on the period utility function $U$ and the marginal utility of consumption. I am not aware of this result being obtained previously in a model with endogenous labor supply.

The intertemporal marginal rate of substitution of consumption between neighboring dates is

$$\frac{\partial V_t}{\partial C_t} = (1 - \beta B_t) V_t^{\rho} U_{C,t} \tag{3}$$

$$\frac{\partial V_t}{\partial C_{t+1}} = \beta B_t (1 - \beta B_{t+1}) V_t^{\rho} \left(E_t V_{t+1}^{1-\alpha_t}\right)^{\frac{\alpha_t - \rho}{1-\rho}} V_{t+1}^{\rho - \alpha_t} B_{t+1} U_{C,t+1} \tag{4}$$

$$M_{t+1} \equiv \frac{\partial V_t / \partial C_{t+1}}{\partial V_t / \partial C_t} = \beta B_t \frac{1 - \beta B_{t+1}}{1 - \beta B_t} \frac{U_{C,t+1}}{U_{C,t}} \frac{V_{t+1}^{\rho - \alpha_t}}{\left(E_t V_{t+1}^{1-\alpha_t}\right)^{\frac{\rho - \alpha_t}{1-\alpha_t}}} \tag{5}$$

where $U_{C,t} \equiv \partial U_t / \partial C_t$ is the marginal (period) utility of consumption. $M_{t+1}$ denotes the SDF between dates $t$ and $t+1$.

In the case where $U_t = C_t^{1-\rho}$ and $B_t$ is constant, $M_{t+1}$ reduces to the usual formula for the SDF when utility depends only on consumption. If the marginal (period) utility of consumption depends on labor, then the SDF will be distorted in the usual ways through $\frac{U_{C,t+1}}{U_{C,t}}$. Even if $U_C$ only depends on consumption, though (i.e. if period utility is separable between consumption and leisure), variation in labor will still affect the SDF through $V_{t+1}$. That is, with recursive preferences, it is not generally possible to separate labor supply decisions from asset prices, unlike the case where preferences are separable between consumption and labor and over time.

#### 2.2.1 Substituting in an asset return

Now consider an asset that pays $U_t U_{C,t}^{-1}$ as its dividend in each period. In the usual analysis of Epstein–Zin preferences, one substitutes the return on an asset that pays consumption as its dividend into the SDF. In the present case, dividing period utility, $U_t$, by the marginal utility of consumption intuitively converts $U_t$ from utility units into consumption units.

We now derive the price of a claim to $U_t U_{C,t}^{-1}$. Denote the cum-dividend price of this asset as $W_{U,t}$. The appendix confirms the guess that

$$W_{U,t} = V_t^{1-\rho} B_t^{-1} U_{C,t}^{-1} / \left(1 - \beta B_t\right) \tag{6}$$

and that we can substitute the return on this asset into the SDF to obtain

$$M_{t+1} = \beta^{\frac{1-\alpha_t}{1-\rho}} \left( \frac{U_{C,t+1}}{U_{C,t}} B_t \frac{1 - \beta B_{t+1}}{1 - \beta B_t} \right)^{\frac{1-\alpha_t}{1-\rho}} R_{U,t+1}^{\frac{\rho-\alpha_t}{1-\rho}} \tag{7}$$

$$R_{U,t+1} \equiv \frac{W_{U,t+1}}{W_{U,t} - U_t U_{C,t}^{-1}} \tag{8}$$

Again, after substituting in an asset return, the SDF is identical to what we would obtain if utility depended only on market consumption. The only difference is that the return is on a subtly different asset, and the marginal utility of consumption potentially follows a different process from the consumption-only case.

### 2.3 Period utility

The period utility function, $U(C_t, N_t)$ is motivated from a model of household production as in Rudebusch and Swanson (2010). Households have power utility over both market goods and goods produced at home.

$$U_t = \frac{\left(C_t^{\eta}\bar{C}_{t-1}^{1-\eta}\right)^{1-\rho}}{1-\rho} + \varphi_1 \frac{C_{H,t}^{1-\rho}}{1-\rho}$$

where $C_{H,t}$ is consumption of the home good. Households do not derive utility directly from leisure, but rather from what they are able to produce in their non-market-work time (as in Campbell and Ludvigson, 2001). The home production function is $Z_t N_{H,t}^{\alpha_H}$, for hours worked at home $N_{H,t}$ and a coefficient $0 < \alpha_H < 1$. The level of labor-neutral technology in the economy is assumed to be equal (up to a constant of proportionality) in the home and market production sectors.

The period utility function can then be written as

$$U_t \equiv \frac{\left(C_t^{\eta}\bar{C}_{t-1}^{1-\eta}\right)^{1-\rho}}{1-\rho} + Z_t^{1-\rho}\varphi_1 \frac{\left(\bar{H} - N_t\right)^{\alpha_H(1-\rho)}}{1-\rho} \tag{9}$$

where $N_{H,t} = \bar{H} - N_t$. $\bar{H}$ denotes the maximum number of hours that the household can work, either at home or in the market, and $N_t$ is market labor. If sleep is part of home production, then $\bar{H}$ might equal 8760 hours for annual data. More generally, though, $\bar{H}$ might be smaller. As a practical matter, $\bar{H}$ affects the elasticity of utility with respect to market labor and the Frisch elasticity. The parameters $\varphi_1$, $\bar{H}$, and $\alpha_H$ jointly determine the average level of hours worked (conditional on wages and consumption), the Frisch elasticity, and the elasticity of utility with respect to market labor, all near the steady state of the model.

The first term in (9) gives the utility that comes from consumption. The household has power utility over a Cobb–Douglas aggregate of current and (aggregate) past consumption. This formulation differs from the standard recent implementation in the macro literature in that I assume a multiplicative instead of additive habit. Campbell and Cochrane (1999) show that an additive habit can induce time-varying risk aversion, whereas the multiplicative habit will have no affect on risk aversion; that way, variation in risk preferences is driven purely by $\alpha_t$. The key feature of the additive habit is simply that the marginal utility of current consumption is increasing in last period's consumption, which induces consumers to try to smooth consumption growth, as observed in the data. To obtain that result in this setting (assuming

$0 < \eta < 1$), we need $\rho < 1$. The consumption and production-based asset-pricing literatures also often assume $\rho < 1$ as it implies that increases in expected growth or decreases in risk aversion, disaster risk, or volatility lower asset prices (Bansal and Yaron, 2004; Barro, 2006; Gourio, 2010; Dew-Becker, 2011a).

# 3 Aggregate supply

For the supply side of the model, I follow exactly the setup in Justiniano, Primiceri, and Tambalotti (JPT; 2010). JPT is a standard medium-scale New-Keynesian model. It has 7 fundamental shocks—price and wage markups, labor-augmenting technical change, investment-specific productivity, monetary policy, discount rates ($B_t$), and government spending. In JPT's formulation, the monetary authority's inflation target is constant. I allow it to vary to help match the movements in the long end of the yield curve. Other than that and the preference specification, my model is identical to theirs.

The model is also highly similar to Smets and Wouters (SW; 2003). The critical difference between the present setup and SW is that technology is difference-stationary rather than trend-stationary, where the former is standard in the production-based asset pricing literature.[2] The difference-stationarity assumption helps generate large risk premia; when technology is trend-stationary, there is very little overall risk in the economy, so households must have an implausibly high coefficient of relative risk aversion in order to generate realistic asset prices.[3]

Since the model is standard and laid out in JPT and the main contribution of this paper is the preference specification and bond pricing, the remainder of this section gives a relatively short description of the production setup. The reader is referred to JPT for a more detailed analysis. My description follows theirs closely.

[2] A difference-stationary process has first-differences that follow a stationary process, so it is integrated of order one. A trend-stationary process, on the other hand, is a process that has random stationary deviations around a non-stochastic trend (where the trend is generally unmodeled and taken as exogenous).

[3] Below, I estimate average risk aversion to be 18.7 (ignoring the correction from Swanson, 2011). Rudebusch and Swanson (2011), who use stationary technology (with a slightly different preference specification) choose an analogous parameter to be 149.

### 3.1 Producers of physical goods

Final-good producers are competitive in both input and output markets and have a CES production function,

$$Y_t = \left[ \int_0^1 Y_t (i)^{\frac{1}{1+\lambda_{p,t}}} di \right]^{1+\lambda_{p,t}} \tag{10}$$

where $i$ indexes the types of intermediate goods, $Y_t$ is output of the final good, which can be used for either consumption or investment, $Y_t (i)$ is the use of intermediate of type $i$, and the elasticity of substitution across the intermediates, which determines markups in the intermediate-goods sector, varies over time.

Intermediate-good producers are monopolists for their own goods with production function

$$Y_t (i) = \max \left\{ K_t (i)^{\gamma} Z_t^{1-\gamma} N_t (i)^{1-\gamma} - Z_t \bar{F}, 0 \right\} \tag{11}$$

where $\bar{F}$ is a fixed cost of production that ensures that profits are zero in steady state. $K_t (i)$ and $L_t (i)$ are intermediate-good producer's $i$ purchases of capital and labor services, and $Z_t$ is the level of labor-augmenting technology.

### 3.2 Price setting

We assume Calvo pricing. In every period, a fraction $1 - \xi_p$ of intermediate good producers can change their prices, while the remainder index their prices following the rule,

$$P_t (i) = P_{t-1} (i) \pi_{t-1}^{\iota_p} \pi^{1-\iota_p} \tag{12}$$

where $P_t (i)$ is the price of good $i$ in terms of the numeraire, $\pi_t \equiv P_t / P_{t-1}$ is aggregate inflation, and

$$P_t = \left[ \int_0^1 P_t (i)^{\lambda_{p,t}^{-1}} di \right]^{\lambda_{p,t}} \tag{13}$$

is the aggregate price index (equal to the marginal cost of a unit of the final good). $\pi$ is the steady-state inflation rate, and $\iota_p$ determines the degree of indexation to lagged versus average inflation.

The firms that can choose their prices freely set them to maximize the present discounted value of profits,

$$E_t\left\{\sum_{s=0}^{\infty} M_{t,t+s}\left[P_t\left(i\right)\left(\prod_{k=1}^{s}\pi_{t+k-1}^{\iota_p}\pi^{1-\iota_p}\right)Y_{t+s}\left(i\right)-W_{t+s}N_{t+s}\left(i\right)-r_{t+s}^{k}K_{t+s}\right]\right\} \tag{14}$$

where $M_{t,t+s}\equiv\prod_{j=1}^{s}M_{t+j}$, $W_{t+s}$ is the wage rate, and $r_{t+s}^{k}$ is the rental rate for capital.

### 3.3 Employment agencies and wage setting

Each household is a monopolistic supplier of specialized labor, $N_t\left(j\right)$. Competitive employment agencies aggregate labor supply into homogeneous labor input (just as the final good producers aggregate intermediate goods) with the production function,

$$N_t=\left[\int_0^1 N_t\left(j\right)^{\left(1+\lambda_{w,t}\right)^{-1}}dj\right]^{1+\lambda_{w,t}} \tag{15}$$

where, as with prices, $\lambda_{w,t}$ determines the elasticity of demand and hence markups in the labor market. $\lambda_{w,t}$ acts as a labor-supply shock. Since the employment agencies are competitive, the price of a unit of the homogeneous labor input is

$$W_t=\left[\int_0^1 W_t\left(j\right)^{\lambda_{w,t}^{-1}}dj\right]^{\lambda_{w,t}} \tag{16}$$

The labor demand function is then

$$N_t\left(j\right)=\left(\frac{W_t\left(j\right)}{W_t}\right)^{-\frac{1+\lambda_{w,t}}{\lambda_{w,t}}}N_t \tag{17}$$

As with prices, wages can only be changed intermittently, with probability $\left(1-\xi_w\right)$. If a household cannot change its wage, it indexes according to the rule

$$W_t\left(j\right)=W_{t-1}\left(j\right)\left(\pi_{t-1}\exp\left(z_{t-1}\right)\right)^{\iota_w}\left(\pi\exp\left(\bar{z}\right)\right)^{1-\iota_w} \tag{18}$$

To solve the wage-setting problem, we write out the household's complete optimization problem. With lagrange multipliers $\psi_t$ and $\lambda_t$ on the labor demand and budget constraints, respectively, we can write the household's lagrangian as

$$V_t - E_t \sum_{k=0}^{\infty} \psi_{t+k} \left[ N_{t+k}(j) - \left( \frac{W_{t+k}(j)}{W_{t+k}} \right)^{-\frac{1+\theta_w}{\theta_w}} N_{t+k} \right] \tag{19}$$

$$-E_t \sum_{k=0}^{\infty} \lambda_{t+k} \left[ P_{t+k} C_{t+k} + H_{t+k} - (1 + i_{t+k}) H_{t+k-1} - W_{t+k}(j) N_{t+k}(j) - \Pi_{t+k} \right] \tag{20}$$

(Recall that $H_t$ represents holdings of nominal bonds). The first-order conditions with respect to $C_{t+k}$ and $N_{t+k}(j)$ are

$$C_{t+k} \quad : \quad \frac{dV_t}{dC_{t+k}} = \lambda_{t+k} P_{t+k} \tag{21}$$

$$N_{t+k}(j) \quad : \quad 0 = \frac{dV_t}{dN_{t+k}(j)} - \psi_{t+k} + \lambda_{t+k} W_{t+k}(j) \tag{22}$$

For $W_{t+k}(j)$, the appendix shows that the first-order condition is

$$0 = E_t^* \sum_{k=0}^{\infty} \xi_w^k \frac{dV_t/dC_{t+k}}{dV_t/dC_t} \left[ \frac{dV_t/dN_{t+k}(j)}{dV_t/dC_{t+k}} (1 + \theta_w) + \frac{W_{t+k}(j)}{P_t} \right] N_{t+k}(j) \tag{23}$$

where $E^*$ is the expectation conditional on the household not being able to change its wage rate. Now $\frac{dV_t/dC_{t+k}}{dV_t/dC_t}$ is simply the SDF between dates $t$ and $t+k$. We also have $\frac{dV_t/dN_{t+k}(j)}{dV_t/dC_{t+k}} = \frac{U_{N,t+k}}{U_{C,t+k}}$. That is, it is simply the intratemporal marginal rate of substitution between consumption and leisure, which is identical to what we obtain under time-separable preferences. The first-order condition for the wage choice here is identical to what is obtained under time-separable utility, so we obtain an identical wage Phillips curve through standard manipulations. I am not aware of this result being derived under recursive preferences previously.[4]

### 3.4 Capital and investment

Intermediate goods firms rent capital from the households at rate $r_t^k$. Households own a stock of capital $\bar{K}_t$ and choose a utilization rate $u_t$ so that the effective quantity of capital rented to firms in period $t$ is

$$K_t = u_t \bar{K}_t \tag{24}$$

[4] This result also shows why preferences need to be recursive over both consumption and leisure together. We could not combine recursive preferences for consumption with time-separable preferences over leisure because the marginal rate of substitution between consumption and leisure would then change depending on the time horizon.

The household pays a cost of utilization $a(u_t)$ per unit of capital, with $u = 1$ in steady state, $a(1) = 0$ and $\chi = a''(1)/a'(1)$.[5]

Households accumulate capital according to the rule,

$$\bar{K}_t = (1-\delta)\bar{K}_{t-1} + \mu_t \left(1 - S\left(\frac{I_t}{I_{t-1}}\right)\right) I_t \tag{25}$$

where $\delta$ is the depreciation rate and the function $S$ incorporates adjustment costs in the rate of investment. In steady state, $S = S' = 0$ and $S'' > 0$. $\mu_t$ is a shock to the cost of investment at date $t$.

### 3.5 Government policy

The central bank follows a Taylor rule taking the form

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\rho_R} \left[\pi_t^* \left(\frac{\pi_t}{\pi_t^*}\right)^{\phi_\pi} \left(\frac{X_t}{X_t^*}\right)^{\phi_X}\right]^{1-\rho_R} \left[\frac{X_t/X_{t-1}}{X_t^*/X_{t-1}^*}\right]^{\phi_{dX}} \eta_{mp,t} \tag{26}$$

where $R_t$ is the gross nominal interest rate, $R$ is its steady-state value, $X_t$ is total output, $X_t^*$ is the level of output that would prevail if prices had always been flexible, and $\pi_t^*$ is the inflation target at date $t$. The central bank is allowed to respond to both the level and change in the output gap. This flexibility helps ensure the model can match the dynamics of short-term interest rates, which is obviously critical for capturing the dynamics of the term structure. $\eta_{mp,t}$ is an exogenous monetary policy shock.

$\pi_t^*$ is a time-varying inflation target, which can potentially help match the high inflation and long-term interest rates seen in the early part of the sample. More generally, $\pi_t^*$ induces a level factor in the term structure: see Bekaert, Cho, and Moreno (2010) and Doh (2011).

The government finances public spending by selling single-period bonds. Government expenditures, $G_t$, are a time-varying fraction of total output,

$$G_t = \left(1 - \frac{1}{g_t}\right) Y_t \tag{27}$$

where $g_t$ follows an exogenous process defined below.

[5] As usual, in the log-linear approximation, the conditions on the first and second derivatives in steady state are sufficient to describe the dynamics of the model.

### 3.6 Market clearing

The aggregate resource constraint is

$$C_t + I_t + G_t + a\left(u_t\right)\bar{K}_{t-1} = Y_t \tag{28}$$

### 3.7 Exogenous processes

The price and wage markup shocks follow ARMA(1,1) processes,

$$\log\left(1+\lambda_{p,t}\right) = \left(1-\rho_p\right)\log\left(1+\lambda_p\right) + \rho_p \log\left(1+\lambda_{p,t-1}\right) + \varepsilon_{p,t} - \theta_p \varepsilon_{p,t-1} \tag{29}$$

$$\log\left(1+\lambda_{w,t}\right) = \left(1-\rho_w\right)\log\left(1+\lambda_w\right) + \rho_w \log\left(1+\lambda_{w,t-1}\right) + \varepsilon_{w,t} - \theta_w \varepsilon_{w,t-1} \tag{30}$$

where $\varepsilon_{p,t} \sim N\left(0, \sigma_p^2\right)$ and $\varepsilon_{w,t} \sim N\left(0, \sigma_w^2\right)$. The ARMA(1,1) form potentially helps match both the high and low-frequency features of inflation.

Productivity has a unit root and its growth rate follows an AR(1) process,

$$\Delta z_t = \left(1-\rho_z\right)\bar{z} + \rho_z \Delta z_{t-1} + \varepsilon_{z,t} \tag{31}$$

where $\varepsilon_{z,t} \sim N\left(0, \sigma_z^2\right)$. The AR(1) setup potentially allows the model to incorporate the long-run risks studied by Bansal and Yaron (2004).

The level of investment-specific productivity is assumed to be a stationary AR(1) process,

$$\log \mu_t = \rho_\mu \log \mu_{t-1} + \varepsilon_{\mu,t} \tag{32}$$

where $\varepsilon_{\mu,t} \sim N\left(0, \sigma_\mu^2\right)$. Note that $\mu_t$ simply determines the efficiency of the transformation of the final output good into the investment good, so investment still benefits from the unit-root innovations to $Z_t$.

The government's share of output follows an AR(1)

$$\log g_t = \left(1-\rho_g\right)\log g + \rho_g \log g_{t-1} + \varepsilon_{g,t} \tag{33}$$

The monetary policy shock follows an AR(1),

$$\eta_{mp,t} = \rho_{mp}\eta_{mp,t-1} + \varepsilon_{mp,t} \tag{34}$$

The final shock is the preference shock $B_t$. We assume

$$\log B_t = \rho_b \log B_{t-1} + \varepsilon_{b,t} \tag{35}$$

The two exogenous processes that are added to JPT's original model are the inflation target and risk aversion. As in Dew-Becker (2011a), I allow the innovations to risk aversion to be correlated with $\varepsilon_{z,t}$. Intuitively, this means that risk aversion depends on innovations to the permanent component of consumption. There are also exogenous innovations to risk aversion. We thus have

$$\alpha_t = \rho_\alpha \alpha_{t-1} + (1-\rho_\alpha)\,\bar{\alpha} + \theta_{\alpha,z}\varepsilon_{z,t} + \varepsilon_{\alpha,t} \tag{36}$$

with $\varepsilon_{\alpha,t} \sim N\left(0, \hat{\sigma}_\alpha^2\right)$.

While a number of recent papers have studied models with time-varying inflation targets (e.g. Gurkaynak, Sack, and Swanson, 2005; Doh, 2010), there is little understanding of what actually drives the inflation target. Because the inflation target has a very strong impact on long-term bond prices, the relationship between the inflation target and the other innovations is a key determinant of the prices of long-term bonds. I therefore consider a loose specification where innovations to the inflation target may be correlated with all of the other fundamental shocks (excluding risk aversion). We thus have

$$\begin{aligned} \log \pi_t^* &= (1-\rho_\pi)\log \pi + \rho_\pi \log \pi_{t-1}^* \\ &\quad + \theta_{\pi*,g}\varepsilon_{g,t} + \theta_{\pi*,z}\varepsilon_{z,t} + \theta_{\pi*,p}\varepsilon_{p,t} + \theta_{\pi*,w}\varepsilon_{w,t} + \theta_{\pi*,b}\varepsilon_{b,t} + \theta_{\pi*,\mu}\varepsilon_{\mu,t} + \theta_{\pi*,mp}\varepsilon_{mp,t} + \varepsilon_{\pi*,t} \end{aligned} \tag{37}$$

with $\varepsilon_{\pi*,t} \sim N\left(0, \hat{\sigma}_{\pi*}^2\right)$. All of the shocks $\varepsilon$ are also assumed to be independent.

The $\theta$ parameters in equations (36) and (37) are somewhat difficult to interpret and choose priors for.

I therefore transform these parameters so that they can be interpreted as variance shares. Define

$$\sigma_\alpha^2 \equiv \theta_{\alpha,z}^2\sigma_z^2 + \hat{\sigma}_\alpha^2 \tag{38}$$

$$\sigma_{\pi*}^2 \equiv \theta_{\pi*,g}^2\sigma_g^2 + \theta_{\pi*,z}^2\sigma_z^2 + \theta_{\pi*,p}^2\sigma_p^2 + \theta_{\pi*,w}^2\sigma_w^2 + \theta_{\pi*,b}^2\sigma_b^2 + \theta_{\pi*,\mu}^2\sigma_\mu^2 + \theta_{\pi*,mp}^2\sigma_{mp}^2 + \hat{\sigma}_{\pi*}^2 \tag{39}$$

$\sigma_\alpha^2$ and $\sigma_{\pi*}^2$ are the variances of the total innovations to risk aversion and the inflation target, respectively. Next, define

$$\sigma_{\alpha,z} \equiv sign\left(\theta_{\alpha,z}\right)\frac{\theta_{\alpha,z}^2\sigma_z^2}{\sigma_\alpha^2} \tag{40}$$

$\sigma_{\alpha,z}$ is the share of the total variance of the innovations to risk aversion that is accounted for by labor-neutral technology shocks. The sign of $\sigma_{\alpha,z}$ determines whether the effect of technology shocks on risk aversion is positive or negative. Similarly, for $\pi_t^*$ we can define

$$\sigma_{X,z} = sign\left(\theta_{\pi*,X}\right)\frac{\theta_{\pi*,X}^2\sigma_X^2}{\sigma_{\pi*}^2} \tag{41}$$

for $X \in \{g, z, p, w, b, \mu, mp\}$. The parameters $\sigma_{\alpha,z}$, $\sigma_{X,z}$, $\sigma_\alpha^2$ and $\sigma_{\pi*}^2$ map uniquely into the original parameters $\theta_{\alpha,z}$, $\theta_{\pi*,X}$, $\hat{\sigma}_\alpha^2$, and $\hat{\sigma}_{\pi*}^2$ are are more easily interpreted.

# 4 Model solution

The standard method for approximating models of the form studied here is perturbation. The drawback of perturbation methods for our purposes is that if we want time-variation in risk aversion to have any effect on the dynamics of the model, we need to take a third-order approximation to the model. Since the solution would be non-linear, we would have to use the particle filter or some other nonlinear method in order to calculate the marginal likelihood of the model. I have found, though, that it is in general very difficult to find the peak of the likelihood function for this model, and it would be infeasible with a method as slow as the particle filter. This is a common problem in models of the term structure (e.g. Ang and Piazzesi, 2003; Hamilton and Wu, 2011).

I therefore use the essentially affine approximation method described in Dew-Becker (2011b). The essentially affine method delivers an approximation to the equilibrium dynamics of the model that is linear in the state variables but still allows time-varying risk aversion to affect the behavior of the endogenous variables. Dew-Becker (2011b) describes the method in detail and show that Euler equation errors in simulated models are competitive with third-order perturbations. Local to the non-stochastic steady-state, the essentially affine approximation is as accurate as a first-order perturbation, and hence less accurate than higher-order perturbations. However, in a stochastic setting, he finds it performs well. This section gives a short overview of the method, and the appendix provides further details.

Denote the vector of the variables in the model (including the exogenous processes) as $X_t$ and the vector of fundamental shocks as $\varepsilon_t \equiv [\varepsilon_{mp,t}, \varepsilon_{z,t}, \varepsilon_{b,t}, \varepsilon_{\mu,t}, \varepsilon_{g,t}, \varepsilon_{p,t}, \varepsilon_{w,t}, \varepsilon_{\alpha,t}, \varepsilon_{\pi*,t}]$. The equations determining the equilibrium of the model take the form

$$0 = G\left(X_t, X_{t+1}, \varepsilon_{t+1}\right) \tag{42}$$

where the expectation operator may appear in the function $G$.There is one equation for each variable.

The equations $G$ can be divided into two types: those that do not involve taking expectations over the SDF and those that do.

$$G\left(X_t, X_{t+1}, \varepsilon_{t+1}\right) = \begin{bmatrix} D\left(X_t, X_{t+1}, \varepsilon_{t+1}\right) \\ E_t\left[M\left(X_t, X_{t+1}, \sigma\varepsilon_{t+1}\right) F\left(X_t, X_{t+1}, \varepsilon_{t+1}\right)\right] \end{bmatrix} \tag{43}$$

where $D$ and $F$ are vector-valued functions and $M$ is the (scalar-valued) stochastic discount factor.[6]

For the equations that do not involve the SDF, I use standard perturbation methods and simply take a log-linear approximation. We approximate $D$ as

$$0 = \log\left(D\left(\exp\left(x_t\right), \exp\left(x_{t+1}\right), \varepsilon_{t+1}\right) + 1\right) \tag{44}$$

$$0 \approx d_0 + d_x \hat{x}_t + d_{x'} \hat{x}_{t+1} + d_\varepsilon \varepsilon_{t+1} \tag{45}$$

where the terms $d_0$, $d_x$, $d_{x'}$, and $d_\varepsilon$ are coefficients from a Taylor-series approximation and

$$x_t \equiv \log X_t$$

$$\hat{x}_t \equiv \log X_t - \log \bar{X}$$

[6] Note that this formulation does not actually restrict $F$.Specifically, suppose there were a set of equilibrium conditions $1 = E_t h\left(x_t, x_{t+1}, \sigma\varepsilon_{t+1}\right)$, i.e. that do not involve the SDF. We could simply say that $F\left(x_t, x_{t+1}, \sigma\varepsilon_{t+1}\right) \equiv h\left(x_t, x_{t+1}, \sigma\varepsilon_{t+1}\right) / M\left(x_t, x_{t+1}, \sigma\varepsilon_{t+1}\right)$.

The second set of equations is dynamic and involves expectations. In many economic models, including the present one, equations involving expectations take the form

$$1 = E_t M_{t+1}\left(X_t, X_{t+1}, \varepsilon_{t+1}\right) F\left(X_t, X_{t+1}, \varepsilon_{t+1}\right) \tag{46}$$

where $M\left(X_t, X_{t+1}, \varepsilon_{t+1}\right)$ is the stochastic discount factor coming from the household's intertemporal optimization condition. The key source of non-linearity in the model is the time-variation in risk aversion, which induces heteroskedasticity in the SDF. It is therefore natural to deal with $M$ and $F$ separately to isolate the relevant non-linearity.

I now show that if we log-linearize $F$, we can transform (46) into a linear condition that can be solved alongside the remaining equations. $M_{t+1}$ will not even be log-linear in the state variables, but we will be able to state the equilibrium conditions in as a set of linear expectational difference equations.

First, guess that the equilibrium dynamics of the model take the form

$$\hat{x}_{t+1} = C + \Phi\hat{x}_t + \Psi\varepsilon_{t+1} \tag{47}$$

where $\hat{x}_t \equiv \log\left(X_t\right) - \log\left(\bar{X}\right)$. We confirm in the end that the solution is actually in this form.

The next step then is to take log-linear approximations to $M$ and $F$ separately. Log-linearizing $F$ is straightforward, and we obtain,

$$\log F\left(x_t, x_{t+1}, \varepsilon_{t+1}\right) \approx f_0 + f_x x_t + f_{x'} x_{t+1} \tag{48}$$

For $M$, the appendix shows that is it is possible to derive a first-order accurate expresion of the form

$$m_{t+1}^{(1)} = m_0 + m_x\hat{x}_t + \left(\kappa_0 + \alpha_t\kappa_1\right)\varepsilon_{t+1} - \frac{1}{2}\alpha_t^2\kappa_1\Sigma\kappa_1' \tag{49}$$

where $\Sigma$ is the variance matrix of $\varepsilon_t$. The superscript (1) indicates that $m_{t+1}^{(1)}$ is first-order accurate for the true SDF. (49) is the essentially affine form from Duffee (2002).

Taking the expectation of the approximated Euler equation yields,

$$
\begin{aligned}
0 &= \log E_t \exp\begin{pmatrix} m_0 + m_x \hat{x}_t + (\kappa_0 + \alpha_t \kappa_1)\varepsilon_{t+1} - \frac{1}{2}\alpha_t^2 \kappa_1 \Sigma \kappa_1' \\ +f_0 + f_x x_t + f_{x'} x_{t+1} \end{pmatrix} \\
0 &= m_0 + m_x \hat{x}_t - \frac{1}{2}\alpha_t^2 \kappa_1 \Sigma \kappa_1' + f_0 + f_x x_t + f_{x'}(C + \Phi \hat{x}_t) \\
&\quad + \frac{1}{2} f_{x'} \Psi \Sigma \Psi' f_{x'}' + \alpha_t \kappa_1 \Sigma \Psi' f_{x'}'
\end{aligned}
$$

Since every equation in the system is now linear in the variables of the model, we can solve the system for the parameters $\Phi$ and $\Psi$ from (77). Specifically, we solve the following system,

$$
0 = d_0 + d_x \hat{x}_t + d_{x'} \hat{x}_{t+1} + d_\varepsilon \varepsilon_{t+1} \tag{50}
$$

$$
0 = \log E_t \exp\begin{pmatrix} m_0 + m_x \hat{x}_t + (\kappa_0 + \alpha_t \kappa_1)\sigma \varepsilon_{t+1} - \frac{1}{2}\alpha_t^2 \kappa_1 \Sigma \kappa_1' \\ +f_0 + f_x x_t + f_{x'} x_{t+1} \end{pmatrix} \tag{51}
$$

The reason that the essentially affine SDF is useful is that the expectation in (51) will be linear in the state variables, so we have a simple linear system to solve. This system can be solved through, for example, Sims' (2001) Gensys algorithm.

Dew-Becker (2011b) shows that the transition function for the model obtained through the essentially affine method is first-order accurate for the true transition function and first-order equivalent to a first-order perturbation. Clearly, though, the approximation includes higher-order terms that account for movements in risk aversion. $\alpha_t$ will affect not only asset prices but also the dynamics of real variables. Dew-Becker (2011b) calibrates a simple version of the RBC model with time-varying risk aversion and finds that the essentially affine approximation has accuracy between that of second and third-order perturbations.

Standard results derived in the appendix also derive real and nominal zero-coupon bond prices.

# 5 Empirics

I estimate the model using standard Bayesian methods. The observable data is the same as in JPT, but with bond prices added. Both real variables and bond prices are linear functions of the underlying state variables contained in the vector $x_t$, so we can write the model in state-space form and measure the likelihood using the Kalman filter. I proceed by finding the posterior mode and running a monte carlo chain from that point to sample from full posterior distribution. The appendix describes the details of the estimation.

## 5.1 Data

The sample is 1983q1 to 2004q4. I do not include the financial crisis in the sample because the zero lower bound on nominal interest rates becomes binding, a phenomenon that the model is not designed to capture. The sample is cut off in 1983 in order to ensure that monetary policy is consistent over the estimation period. Future work will reestimate the model in earlier periods.

The observable variables are real GDP, consumption, and investment growth, hours worked per capita, wage and price inflation, and yields on three-month, 1, 2, 3, 5, and 10-year Treasury bonds. The 1 through 5-year yields are obtained from the Fama–Bliss CRSP files, the 10-year yield from Gurkaynak, Sack, and Wright (2006), and the three-month yield is from the Fama risk-free rate CRSP file. The bond yields and inflation rates are always reported in annualized percentage points, unless otherwise noted. The real variables are all obtained from the BEA and the BLS. Consumption is defined as expenditures on non-durables and services, while investment is the sum of residential and non-residential fixed investment and consumer durables expenditures. Real wages are calculated as nominal compensation per hour in the non-farm business sector (from the BLS) divided by the GDP deflator. The change in the log GDP deflator is the measure of inflation. Hours worked per capita in the non-farm business sector are obtained from Francis and Ramey (2009) as updated on Valerie Ramey's website. None of the variables are detrended.

Figure 1 plots the data used in the estimation (with the exception of the intermediate-term bond yields). Output, consumption, and investment growth all look stationary over the sample and relatively homoskedastic. Hours worked per capita has a strong upward trend in this sample. Interest rates decline significantly over the sample, even though inflation only declines marginally. The short-term interest rate is substantially more volatile than the long-term rate, and the term spread is clearly countercyclical.

The model has 9 fundamental shocks, but we have 13 observable variables. I follow JPT in assuming that the 6 macro variables plus the short-term interest rate are observed without error. I also assume that the 10-year bond yield is measured without error, which will help identify the inflation target. For the remaining bonds, I assume that the yields have orthogonal measurement errors with identical standard deviations. The standard deviation of these measurement errors is another parameter that will be estimated. The assumption of zero measurement error for the long and short ends of the yield curve forces the model to focus on matching the term spread, while leaving some flexibility in matching curvature.

### 5.2 Priors

Table 1 lists the parameters and priors. For all of the parameters that I share with JPT, I choose the same priors. The remaining parameters are listed in the bottom section of the table. Many of them have uniform priors since we do not have strong a priori views about, for example, the fraction of the variance of the Federal Reserve's inflation target that is driven by shocks to government spending.

For the volatility of risk aversion, I choose a beta distribution over the ratio of the unconditional standard deviation of risk aversion to its mean. This means that average risk aversion is forced to be at least one standard deviation above zero. This prior could potentially be tightened to enforce a stronger restriction. As a practical matter, the data tends to push for a high volatility for risk aversion, and average risk aversion in the estimation simply rises high enough to accommodate the unconditional standard deviation.

I constrain the persistence of the inflation target to follow nearly a random walk with $\rho_{\pi*} = 0.99$, consistent with the idea that the target is very slow-moving. This also ensures that inflation is stationary so that there is a steady-state around which we can approximate. The priors over the shares of the variances of the inflation target and risk aversion coming from the other shocks are uniform.

### 5.3 Posterior modes

Table 1 lists the posterior modes for the parameters along with the 5th and 95th percentiles of the posterior distribution. Many of the posterior modes are reasonable close to the corresponding prior means. I focus mainly on those parameters that differ from the prior or are unique to this model.

The prior for the variance of the innovations to the inflation target favors a reasonably low standard deviation, but the posterior seems to want a highly volatile target—the standard deviation of the innovations to the annualized inflation target is 1.3 percent. This helps the model capture the observed volatility of the level factor in bond yields, but it is implausibly high. The two shocks with the largest variance shares for $\pi^*$ are the monetary policy shock (39 percent) and the consumption demand shock (28 percent). When the inflation target falls, the model implies that the monetary authority should cut interest rates to create a boom and raise inflation. The positive correlation between the monetary policy and inflation target shocks implies that when its inflation target rises the monetary authority does not cut interest rates as much as we would expect based on the Taylor rule.

The shock to the level of labor-neutral technology also has an important effect on the inflation target, accounting for 16 percent of the variance of its innovations. Following a positive innovation to technology, the central bank is estimated to lower its inflation target, consistent with the idea that following beneficial supply shocks that drive inflation downward, the central bank takes the opportunity to drive inflation lower persistently (e.g. Gurkaynak, Sack, and Swanson, 2005). This mechanism will turn out to be critical to the main results.

The labor-neutral technology shock has a standard deviation of 0.72 and an autocorrelation of 0.22. The permanent component of the technology process (the Beveridge–Nelson trend) thus has a standard deviation of 1.00, which is similar to the values often calibrated in the production-based asset pricing literature (e.g. Gourio, 2010, and Dew-Becker, 2011). The estimated long-run variance of technology growth is far smaller than the value calibrated in the long-run risks literature (e.g. Bansal and Yaron, 2004, and Kaltenbrunner and Lochstoer, 2010), but it is consistent with estimates obtained in JPT and SW and with simple univariate estimates from consumption and output data.

The standard deviation of the investment technology shock is far larger than the prior, which is also consistent with JPT, showing that innovations that are isolated to the investment sector may play a large role in fluctuations. Alternatively, it could mean that the model simply matches investment poorly.

The estimates imply that there is essentially zero correlation between innovations to technology and innovations to risk aversion. This runs against the theory from Dew-Becker (2011a), and implies that the price of risk in bond markets is driven by some factor other than permanent innovations to household wealth. Intuitively, part of the source of this result is that the price of risk is measured to be high in recessions, but over the 1983–2004 period, productivity growth has been only weakly procyclical, and in fact rose substantially in the 2000 recession.

As in JPT, the government spending shock is estimated to follow nearly a unit root, explaining the trend in the consumption-output ratio over the sample. The wage markup shock also follows nearly a unit root, which helps capture the strong trend in hours worked per capita seen in figure 1.

In general, my parameter estimates are reassuringly similar to the values from JPT reported in the far-right column of table 1, even though I use a different sample period (post-1983 versus post-war) and extra data on bond yields. The main place where my estimates seem to differ from JPT is in price and wage determination. My estimates imply that wages are strongly indexed to inflation, whereas JPT estimate little indexation. The wage-markup shock is also substantially more volatile under my estimates than JPT. Interestingly, when bond prices are dropped from the estimation, I obtain values for wage indexation and the volatility of the wage markup shock that are much closer to JPT. This suggests that wage dynamics are important for matching the term structure. I also estimate a smaller inverse Frisch elasticity and a lower average price markup. In both cases, my estimates bring the model in closer to the priors and micro estimates.

# 6 Asset pricing

This section studies the asset-pricing implications of the model. I first analyze the fit of the model to the term structure and show that it is competitive with a non-structural model. Next, I decompose the variance of the SDF to understand the source of the positive term premium in the model. I then analyze the prices of other assets, including the aggregate capital stock and a claim to aggregate consumption.

## 6.1 Bond prices

### 6.1.1 Fitted yields

Figure 2 plots the deviations of the fitted yields from their actual values for the five yields that are assumed to be measured with error (reported in annualized basis points). The estimated standard deviation of these fitting errors is 8 basis points, which is economically small compared to the overall variation of the yields that is on the order of hundreds of basis points. The errors are all centered around zero, meaning that the model can capture the shape of the term structure on average. The volatility of the errors looks somewhat higher for the 1 and 5-year yields and in the earlier part of the sample. There is clearly some autocorrelation in the errors; the fitted value for the 3-year yield is consistently too high in the first half of the sample, and the 4-year fitted yield is consistently too low in the second half, for example. And there is also some cross-correlation in the errors; the first principal component explains 37 percent of the total variance of the errors (twice what it would if the errors were orthogonal). These are thus clearly not classical (i.i.d.) measurement errors, but their small mean and volatility shows that the model does a reasonable job of fitting the data.

While there nine unobservable shock processes that can help us match the data, the model is asked to fit 13 data series, so obtaining a good fit for the bond yields is not trivial. Loosely, we have 6 macro variables that identify 6 shock processes, plus three extra processes (the monetary policy shock, the inflation target, and risk aversion) that can be used to fit the bond yields. The degrees of freedom here are thus comparable to a non-structural bond-pricing model with three unobservable factors, but we also have numerous constraints on dynamics and risk prices. Table 3 lists the standard deviation of the yield errors in basis points obtained from regressing the bond yields on their first three principal components. The standard deviations are all between 4 and 8 basis points. I force the structural model to match the 1-quarter and 10-year yields exactly, and the remaining yields have errors with standard deviations of 8 basis points. The fit of the model to the yields is thus comparable to a non-structural model with three unobservable factors that have completely unrestricted dynamics.

The third and fourth rows of table 3 report the measurement errors in constrained models that assume constant relative risk aversion and a constant inflation target (the other parameters are reesetimated). In both cases, the measurement errors have standard deviations roughly three times larger than the benchmark model, giving another measure of the improvement in fit generated by the benchmark model.

Figure 2 and table 3 show that the model is able to provide a very close fit to bond yields in the data. The quality of the fit is essentially identical to the that of a purely non-structural model.

### 6.1.2 Steady-state yields

Another way to evaluate the fit of the model is to ask whether the steady state of the model matches the average term structure in the data. Looking at the steady state keeps the Kalman filter from using large deviations in the unobservable state variables to fit the term structure. This is an especially important test when we have three very highly autocorrelated exogenous processes—the inflation target, the wage markup, and the level of government spending. Figure 3 plots the average term structure in the sample along with its model-implied steady state. The solid black line gives the steady-state term structure in the model, renormalized so that the ten-year yield matches the empirical ten-year yield.[7] To capture the uncertainty in the empirical term structure, the grey area gives the 95 percent confidence intervals for the means of the empirical yields relative to the ten-year yield (i.e. the confidence intervals for the spreads; the intervals are calculated using the Newey–West method with lag a 6-quarter lag window). What figure

[7] I use this normalization because the estimated inflation target is above zero through most of the sample. The unconditional variance of the inflation target is sufficiently high that its average level is not well identified.

3 shows is that the model matches the spread between the 10 and 2-year yields, but it does not match the curvature of the term structure below two years. However, all of the model-implied yields are within the 95 confidence intervals. One potential explanation at the very short end of the yield curve is that there is a small liquidity premium that the model is not incorporating.

We will see that two features of the model are critical for generating the large steady-state term premium: first, following a positive shock to technology, the Fed's inflation target falls; second, variation in risk aversion raises the premia on risky assets. To see the prima facie evidence that these two effects are key, figure 3 includes two lines giving the steady-state term structure in constrained models. The first line assumes that innovations to the inflation target are uncorrelated with the permanent technology shock, while the second line assumes that risk aversion is fixed. Neither line reestimates the other parameters, so they simply isolate the effects of those two features of the model.

The line exiting the top of the chart is for the model when shocks to technology are assumed to have no impact on the inflation target. We then obtain the usual result that the term structure is downward-sloping, and the steady-state term spread is -2.61 percentage points. Time-varying risk aversion also turns out to be important, though. When risk aversion is fixed, the term structure is still upward-sloping, but the spread is quantitatively small—only 0.63 percentage points in steady-state, compared to 2.07 in the data and 1.52 in the benchmark model.

#### 6.1.3 Variation in interest rates

To show how the bond yields respond to the various shocks, figure 4 plots responses of a level, slope, and curvature factor to the 9 fundamental shocks. Following Bekaert, Cho, and Moreno (2010), the level factor is defined as the average of the 1-quarter and 5 and 10-year yields; the slope factor is the 10-year/1-quarter term spread; and curvature is the sum of the 5 and 1-year yields minus twice the 3-year yield. The shocks are orthogonalized in the sense that the interactions between the inflation target and risk aversion and the other shocks are switched off. So figure 4 shows, for example, the effect of a pure positive monetary policy shock holding the inflation target fixed. The shocks are all unit standard deviations.

For the level factor, a number of shocks, the monetary policy and time preference shocks in particular, have important effects at high frequencies. The low-frequency movements, as we would expect, are mainly driven by shifts in the inflation target, while risk aversion also plays a role. Somewhat surprisingly, positive monetary policy shocks, which raise the short-term interest rate above its Taylor-rule value, are actually associated with declines in the level factor. The reason is that these shocks drive down expected inflation. So a positive monetary policy shock drives the real interest rate up, but nominal interest rates actually fall. The response to the time-preference shock is more intuitive: an increase in $B_t$ is analogous to an increase in patience, so interest rates fall.

The determinants of the slope factor are similar to those for the level factor: monetary policy and time-preference matter at high frequencies, while risk aversion determines the dynamics at lower frequencies. An increase in risk aversion increases the term spread. This result fits with results on bond return forecasting (Campbell and Shiller, 1988) and the fact that the term spread forecasts high equity returns (Fama and French, 1989).

#### 6.1.4 Term premia

The size of the steady-state term spread in the model can be interpreted as the average term premium—it is the excess return (in logs) that an investor earns in expectation by buying a long-term bond and holding it to maturity instead of buying short-term bonds and rolling them over for the same amount of time. An important feature of this model is that risk aversion varies over time, which should make the term premium also vary over time.

The top panel of figure 5 plots the expected annualized excess return on holding a ten-year nominal bond (over a one-quarter bond) from the benchmark model against the expected excess return from a regression of bond returns on the Cochrane–Piazzesi (CP) factor. Cochrane and Piazzesi (2005) argue that a tent-shaped factor in forward yields summarizes the price of risk in the term structure, so their factor can be viewed as a simple non-structural benchmark for return forecasting.

Because the CP forecast here is based on a linear regression, its errors are, by construction, zero on average.The structural forecast is highly correlated (34 percent) with the fitted value using the CP factor, and its standard deviation is roughly 20 percent larger. The two series rise by similar amounts in the two recessions in the sample, but the benchmark model also implies that the term premium rose in 1988 and 1999, whereas the CP factor is stable in those episodes.

The bottom panel of figure 5 plots the term premium against the term spread. The term premium is defined as the spread between the 10-year yield and the average of the expected 1-quarter yields over the life of a 10-year bond. The variance of the term premium is non-trivial in comparison to the term spread. In the two recessions in the sample, the increases in the term spread are substantially larger than the movements in the term premium, but the term premium does rise in both episodes. Interestingly, the movements in the term spread outside of the two recessions seem almost entirely driven by movements in the term premium. In particular, the rises in the term spread in 1984, 1985, 1987, 1996, and 1999 are all associated with increases in the term premium of equal magnitudes. On the other hand, the inversions of the yield curve in 1989 and 2000, both just prior to recessions, are associated with only minor declines in risk aversion, and the subsequent rises in the term spread with similarly small rises in risk aversion.

This section shows that the model is able to generate an average term premium of similar magnitude to what is observed in the data, and the implied term premium is as volatile as standard regression-based measures. When compared with the term spread, the term premium seems to be the dominant factor in driving the slope of the term structure during expansions, while it plays a smaller role in recessions themselves.

### 6.2 Model comparison

The primary difference between the model studied here and JPT is the addition of time-varying risk aversion and the time-varying inflation target. An important question, then, is the extent to which those two factors improve the fit of the model to the data. Clearly, allowing a time-varying inflation target will help increase the volatility of long-term bond yields, and time-varying risk aversion will induce time-varying risk premia. But it is possible that the data can be well explained with a constant risk premium, or perhaps there is no need to have a time-varying inflation target.

Table 2 considers two alternatives to the benchmark model: a version where the inflation target is constant, and a version where the coefficient of relative risk aversion is constant. It lists two statistics for each model. First, it gives the standard likelihood ratio used in MLE, i.e. based on $\log f(y|\theta, M)$, where $y$ represents the data and $\theta$ the parameter vector. $M$ denotes the model choice, i.e. the full model or one of the two restricted versions. $\log f(y|\theta, M)$ is closely related to the one-step-ahead forecast error of the model. The likelihood ratio test favors the benchmark over each alternative by a wide margin. One way to see the source of this rejection is to note that in table 3, the "measurement errors" for bond yields, essentially a residual variance, are three times higher in the alternative models than the benchmark. This difference alone is more than sufficient to explain the magnitude of the likelihood ratios.

The second statistic is the Bayes factor, which is based on the marginal likelihood conditional only on the model, $\log f(y|M)$ (in the economics literature, see Fernandez-Villaverde and Rubio-Ramirez, 2004). I calculate $\log f(y|M)$ using the monte carlo chain as in Fernandez-Villaverde and Rubio-Ramirez (2004).[8]

The difference in the Bayes factors listed in the bottom row of table 2 is similar to the values obtained under the usual likelihood ratio test. In order to accept the model with constant risk aversion or a constant inflation target, the ratio of the prior probability of either of those models to that of the benchmark model would have to be greater than $\exp(140)$. In a statistical sense, then, both the time-varying inflation target and time-varying risk aversion substantially improve the fit of the model.

## 6.3 Determinants of asset prices

### 6.3.1 The variance of the SDF

An asset's expected excess return over the real riskless interest rate is determined by its covariance with the stochastic discount factor. One of the more interesting outputs of a model as rich as this one is the variance decomposition for the SDF. Table 4 reports a variance decomposition for the SDF at the one-quarter horizon. The variance of the SDF is essentially entirely driven by the neutral technology and risk aversion shocks. The bottom panel of table 4 is a bar chart decomposing the variance of the SDF into components coming from the neutral technology shock, the risk aversion shock, and the remaining shocks combined. The lines at the top of each bar give the 2.5 and 97.5 percentiles of the posterior distribution. The 97.5 percentile for the variance share in the SDF for non-technology and non-risk aversion shocks is less than 2 percent

On first glance this result might be somewhat surprising, but it is in fact a deep characteristic of models with Epstein–Zin preferences with a high EIS and high risk aversion. One way to see the source of this finding is to simply look at the household's SDF,

$$M_{t+1} = \beta B_t \frac{1-\beta B_{t+1}}{1-\beta B_t} \frac{U_{C,t+1}}{U_{C,t}} \frac{V_{t+1}^{\rho-\alpha_t}}{\left(E_t V_{t+1}^{1-\alpha_t}\right)^{\frac{\rho-\alpha_t}{1-\alpha_t}}} \tag{52}$$

For a household with a large EIS, the variance of $U_{C,t+1}/U_{C,t}$ is generally small (at least with standard preferences). In the case where the household does not have a habit ($\eta = 0$), this term is equal to $(C_{t+1}/C_t)^{-\rho}$. If the household has an EIS greater than 1, then $\rho$ is less than 1 and the variance of $U_{C,t+1}/U_{C,t}$ will be less than the variance of log consumption growth. A one percent permanent decline in consumption will raise this term by the factor $1.01^{\rho}$.

[8] See also Gelfand and Dey (1994) and Geweke (1999)

The majority of the variance of the SDF is driven by the term $\frac{V_{t+1}^{\rho-\alpha_t}}{\left(E_t V_{t+1}^{1-\alpha_t}\right)^{\frac{\rho-\alpha_t}{1-\alpha_t}}}$. Here, a 1 percent permanent decline in consumption will make this term (approximately) equal to $1.01^{\alpha_t-\rho}$. When $\alpha_t \gg \rho$, as estimated here and in most models with Epstein–Zin preferences, it is therefore the variation in $V_{t+1}$ that determines the behavior of the SDF.

So what determines movements in $V_t$? One way to think about it is to split movements in consumption into permanent and temporary components. A purely temporary shock to consumption will have a relatively small impact on $V_t$ because the household is not very averse to shifting consumption over time. A permanent shock to consumption, on the other hand, will tend to shift $V_t$ by an equal amount. Risk aversion shocks also affect $V_t$. The reason is simply that increases in risk aversion directly drive $V_t$ down because households are more averse to the future uncertainty that they face. Since it is the shocks to $V_{t+1}$ that determine the movements in $M_{t+1}$, we should thus not be surprised that it is mainly the neutral technology and risk aversion shocks that drive the variance of the SDF.

#### 6.3.2 Impulse responses

Since the technology and risk aversion shocks are the key to understanding the pricing kernel, it is natural to ask how they affect the economy. Figure 6 plots impulse responses to an increase in labor-neutral technology and a decrease in risk aversion. Dotted lines give 95 percent credible intervals (the range between the 2.5 and 97.5 percentiles in the posterior distribution). These impulse responses are different from those in figure 4 because they do not turn off the interactions between the inflation target and risk aversion and the other shocks. The idea is that we want to see what happens on average following these two shocks, since that behavior is what is relevant for understanding the correlations with the SDF (above, the goal was to isolate innovations to the exogenous processes).

Following the technology shock, inflation falls, while output and real interest rates rise: a standard positive supply shock. The declines in inflation and nominal interest rates are especially pronounced and persistent. This behavior is the result of the fact that the estimates imply that the inflation target falls following an increase in technology. However, note that inflation falls more than the inflation target does, so the effect is not entirely driven by the inflation target.

A similar result is obtained in JPT and SW, even though they do not have time-varying inflation targets. In all of these models, because prices are sticky, when there is a positive supply shock, rather than cutting prices, firms simply produce the same quantity as previously, thereby reducing employment and hence demand. Positive technology shocks are thus associated with small increases in output and large declines in the output gap (defined as the difference between output and the level it would take if prices were flexible). There is also a small empirical literature that provides more reduced-form evidence on effects of this sort using direct measures of technology (e.g. Basu, Fernald, and Kimball, 2006).

The slow response of output to a technology shock is particularly notable. The impulse response function in figure 6 suggests that output and perhaps also consumption growth should be predictable and positively autocorrelated, though other shocks could obscure those relationships. Figure 7 therefore plots the empirical and model-predicted autocorrelation functions for consumption and output growth. For output, the model implies that the 1-quarter autocorrelation should be 0.50, while it is only 0.33 in the data. But 0.50 is well within the 95 percent confidence interval for the empirical value. In fact, nearly the entire autocorrelation function for output in the model is captured within the 95 percent confidence interval in the data.

The model also implies strong one-quarter autocorrelation in consumption growth, and the autocorrelation is in fact higher than the upper end of the 95 percent confidence interval. However, this autocorrelation dies out very rapidly, and at lags longer than one quarter the autocorrelation of consumption growth in the model matches the behavior in the data well. The model implies little or no serial correlation in consumption and output growth at horizons longer than two quarters.

Figure 6 also reports impulse responses for a decline in risk aversion. Inflation, interest rates, and the output gap all rise. This shock therefore takes the form of a classic demand shock. The effects are far smaller than those for the technology shock. Risk aversion has two main channels through which it affects the real economy. First, to the extent that physical investment is risky, a decline in risk aversion makes households more willing to purchase physical capital. Second, a decline in precautionary saving demand makes households want to consume more for any given level of interest rates. For this model, the increase in consumption demand dominates, which is why the shock is expansionary. However, the overall effect on the real economy is small.

### 6.3.3 Other asset prices

After the SDF, table 4 reports variance decompositions for the returns on a number of assets. The bottom panel of table 4 reports the fraction of the variance of the one-quarter innovation to each return coming from the neutral technology shock, the risk aversion shock, and all other shocks combined.

Column 2 reports the variance decomposition for the return on the utility portfolio. 87 percent of its variance comes from the time-preference shock. The reason is simply that the utility claim has a relatively long duration, like that of a consol with a coupon that grows at the average rate of the economy, so shifts in real interest rates have a large effect on its price. The time-preference shock mainly affects real interest rates, so it drives the variance of the utility claim. Row 11 shows that the correlation of the utility return with the SDF is 0.31.

More interestingly, the third and fourth columns of table 4 report variance decompositions for a claim on capital and the same claim levered two to one. The capital claim is simply a claim to the marginal product of capital in each period, with the dividends declining at the rate of depreciation on capital. The price of this claim is the usual definition of Tobin's q. Once again, little of the variance of the return is driven by the technology or risk aversion shocks. While the neutral technology shock does play a role, it is relatively small. The reason is that the marginal product of capital is stationary. So even though the technology shock permanently raises productivity, the dividend accruing to any given unit of capital is only raised temporarily due to the subsequent growth in the capital stock. The return on the capital claim is thus driven by the temporary shocks that drive the business cycle. Furthermore, the capital claim is not very strongly correlated with the SDF (-0.31), so it does not have an appreciable excess return on average.

Adding leverage does not change this result. I model a leveraged capital claim as a claim to a unit of capital for which 50 percent of the initial cost of the capital is financed with a 10-year zero-coupon nominal bond. The return of the levered capital claim is twice as volatile as that of the unlevered claim, but it is still only weakly correlated with the SDF, and the annualized excess return is only 10 basis points on average, well below any estimate of the equity premium (perhaps 700 basis points). The bottom panel of table 4 shows that this basic result is robust.

Columns 5 and 6 give variance decompositions for a levered and unlevered consumption claim. The unlevered claim looks similar to the utility claim, and adding leverage does little to change the results. The Sharpe ratios for the consumption claims are far smaller than the Hansen–Jagannathan bound.

In terms of magnitudes, none of the assets studied in table 4 is able to generate an equity premium as large as we observe in the data. The capital claim is nowhere near sufficiently volatile or correlated with the SDF.

#### 6.3.4 The Hansen–Jagannathan bound

Table 4 shows that the standard deviation of the SDF near the steady-state is high—0.43, which implies a Hansen–Jagannathan bound of roughly the same magnitude. In the empirical sample, the average Sharpe ratio for the CRSP value-weighted equity portfolio (its average annual return divided by the standard deviation of annual returns) is 0.52. So even without including stock prices in the estimation, the model generates an empirically reasonable price of risk.

Perhaps more importantly, though, the price of risk in this model is highly volatile. The estimated standard deviation of the Hansen–Jagannathan bound is 0.25. The level of variability here is similar to that used in Dew-Becker (2011a) to match the degree of predictability observed for aggregate stock returns in the post-war sample.

Lettau and Ludvigson (2010) and Cochrane (2011) argue that the variation in the equity premium may be as large as its mean. That is, they claim the price of risk may be twice as volatile as what I find here. The prior for the volatility of risk aversion constrains the standard deviation of risk aversion to be no more than its mean. To match the results from Lettau and Ludvigson (2010) and Cochrane (2011), I would need to allow risk aversion to be perhaps twice as large as it is in the current benchmark model.

## 7 The real economy

Up to now, the analysis has focused mainly on asset pricing. But the model gives a rich description of the real side of the economy. While I leave a deeper analysis of New Keynesian models to papers focused on those models for their own sake, the interaction of the real side of the economy with asset prices is important to this paper.

Figure 8 gives a variance decomposition for the variables used in the estimation. The decomposition gives the fraction of the variance of each of the variables driven by the various shocks at frequencies of 6 to 32 quarters.[9] Except for investment growth, for which the investment-specific shock is completely dominant, none of the other variables examined in figure 8 are dominated by any particular shock. Notably, the shock to risk aversion has almost no effect on the variance of any of the real variables at business cycle frequencies. Its largest effect is on consumption growth, for which the variance share is still only 4 percent.

[9] The variance decomposition is calculated usinga spectral decomposition of the state-space form of the model. I calculate the spectral density at 500 increments between wavelengths of 6 and 32 quarters.

The far-right bar, though, shows that risk aversion has a large effect on the term spread, as we saw in figure 4; it explains roughly 1/3 of the variance of the term spread at business-cycle frequencies.

The variance decomposition reported in the top panel of figure 8 is rather different from that reported by JPT. They found that the investment shock was an important determinant of not only investment, but also output and consumption growth at business-cycle frequencies. Their model differs from mine in three ways: risk aversion and the inflation target are constant, the preference specification is slightly different (log utility and additive habits), and the data sample covers the entire post-war period and only includes the one-quarter nominal interest rate. The bottom panel of figure 8 removes a number of those differences. It drops bond yields (except the short rate) from the estimation and assumes the inflation target and coefficient of relative risk aversion are constant. The model is then reestimated.

While I do not replicate JPT's results exactly, I do also find that the investment technology shock is important for more variables than just investment itself. It accounts for 30 percent of the variation in output growth, 20 percent of consumption growth. Moreover, it accounts for 41 and 48 percent of the variance of the one-quarter nominal interest rate and the term spread, respectively. It is this latter result that explains the divergence between the two models. Since long-term bond prices are included in the estimation of the benchmark model, that model is forced to match the relationship between investment and the term spread. The difference in the variance decompositions suggests that JPT gets this relationship wrong. What is interesting, though, is that when the model is forced to match long-term bond yields, that also changes the decomposition for other variables.

Table 4 gives one-quarter-ahead variance decompositions for output, consumption, and investment growth. This decomposition is useful for understanding whether any of these variables would be powerful asset pricing factors. Specifically, in a world where consumption followed a random walk and households had Epstein–Zin preferences with constant relative risk aversion, consumption growth would be perfectly correlated with the SDF, so it would price assets in the economy.

What table 4 shows is that consumption, output, and investment growth are all only weakly correlated with the SDF at the one-quarter horizon—they all have correlations less than 16 percent. So asset pricing with only consumption growth will not work well in this economy. Many asset-pricing studies with Epstein–Zin preferences include both consumption growth and the return on the stock market as pricing factors. If we believe that the stock market is a claim on aggregate capital, then table 3 shows that it will do little to help with asset pricing as it is also only weakly correlated with the SDF.

## 8 Conclusion

This paper studies bond pricing in a medium-scale New-Keynesian model with a time-varying price of risk. I show that the model can generate a large and volatile term premium. The term premium is driven by the combination of two factors—a negative response of interest rates to positive technology shocks and variation in risk aversion. Removing either of these effects eliminates the model's ability to match the magnitude of the term premium.

While shocks to risk aversion and technology determine average asset returns, they have only weak effects on real variables at business-cycle frequencies. The covariance of asset returns with real variables over the business cycle is therefore unimportant for determining average returns. It is true that the Federal Reserve tends to cut interest rates in recessions, but the model shows that most recessions are not high-marginal-utility states of the world. So the usual intuition that the Taylor rule should lead to a downward-sloping yield curve is inaccurate since it does not take into the difference between the shocks that have high risk prices and the shocks that drive the business cycle.

Furthermore, while risk aversion is estimated to be highly volatile and to be an important determinant of the dynamics of the term spread, it has almost no effects on the real economy. This model thus suggests that there is a separation between the price of risk in financial markets and the real economy.

## References

**Ang, Andrew and Monika Piazzesi**, "No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," *Journal of Monetary Economics*, 2003, *50*, 745–787.

**Bansal, Ravi and Amir Yaron**, "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 2004, *59* (4), 1481–1509.

**Barro, Robert J.**, "Rare Disasters and Asset Markets in the Twentieth Century," *Quarterly Journal of Economics*, 2006, *121(3)*, 823–866.

**Basu, Susanto, John G. Fernald, and Miles S. Kimball**, "Are Technology Improvements Contractionary?," *American Economic Review*, 2006, *96(5)*, 1418–1448.

**Bekaert, Geert, Seonghoon Cho, and Antonio Moreno**, "new Keynesian Macroeconomics and the Term Structure," *Journal of Money, Credit ans Banking*, 2010, *42(1)*, 33–62.

**Campbell, John Y. and Robert J. Shiller**, "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies*, 1991, *58(3)*, 495–514.

**Christiano, Lawrence J., Mathias Trabandt, and Karl Walentin**, "DSGE Models for Monetary Policy Analysi," in "Handbook of Monetary Economics," North-Holland, 2011.

**Cochrane, John H.**, "Discount Rates," *Journal of Finance*, 2011, *66(4)*, 1047–1108. AFA Presidential Address.

___ **and Monika Piazzesi**, "Bond Risk Premia," *American Economic Review*, 2005, *95* (1), 138–160.

**Dew-Becker, Ian**, "A Model of Time-Varying Risk Premia with Habits and Production," 2011. Working paper.

___, "Essentially affine approximations for economic models." Working paper.

**Doh, Taeyoung**, "Yield Curve in an Estimated Nonlinear Macro Model," *Journal of Economic Dynamics & Control*, 2011.

**Duffee, Gregory R.**, "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance*, 2002, *57(1)*, 405–443.

**Fama, Eugene F. and Kenneth R. French**, "Business Conditions and Expected Returns on Stocks and Bonds," *Journal of Financial Economics*, 1989, *25(1)*, 23–49.

**Fernandez-Villaverde, Jesus and Juan Rubio-Ramirez**, "Comparing dynamic equilibrium models to data: a Bayesian approach," *Journal of Econometrics*, 2004, *123(1)*, 153–187.

**Francis, Neville and Valerie A. Ramey**, "Measures of Per Capita Hours and Their Implications for the Technology-Hours Debate," *Journal of Money, Credit, and Banking*, 2009, *41(6)*, 1071–1097.

**Gelfand, A. E. and D. K. Dey**, "Bayesian Model Choice: Asymptotics and Exact Calculations," *Journal of the Royal Statistical Society. Series B*, 1994, *56(3)*, 501–514.

**Gelman, A., G.O. Roberts, and W.R. Gilks**, "Efficient Metropolis Jumping Rules," in J.M. Bernardo, J.O. Berger, A.P. Dawid, and A.F.M. Smith, eds., *Bayesian Statistics*, Vol. 5, Oxford University Press, 1996.

**Geweke, John**, "Using simulation methods for bayesian econometric models: inference, development,and communication," *Econometric Reviews*, 1999, *18(1)*, 1–73.

**Gourio, Francois**, "Disaster Risk and Business Cycles," 2010. Working Paper.

**Gurkaynak, Refet S., Brian Sack, and Eric Swanson**, "The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models," *American Economic Review*, 2005, *95(1)*, 425–436.

**Haario, Heikki, Eero Saksman, and Johanna Tamminen**, "An Adaptive Metropolis Algorithm," *Bernoulli*, 2001, *7(2)*, 223–242.

**Hamilton, James D. and Cynthia Wu**, "Identification and Estimation of Affine-Term-Structure Models." Working Paper.

**Joslin, Scott, Kenneth J. Singleton, and Haoxiang Zhu**, "A New Perspective on Gaussian Dynamic Term Structure Models," *Review of Financial Studies*, 2011, *24(3)*, 926–970.

**Judd, Kenneth L.**, *Numerical Methods for Economists*, MIT Press, Cambridge, MA, 1999.

**Justiniano, Alejandro, Giorgio Primiceri, and Andrea Tambalotti**, "Investment Shocks and Business Cycles," *Journal of Monetary Economics*, 2010, *57(2)*, 132–145.

**Lettau, Martin and Sydney Ludvigson**, "Measuring and Modeling Variation in the Risk-Return Tradeoff," in Yacine Ait-Sahalia and Lars P. Hansen, eds., *Handbook of Financial Econometrics*, Vol. 1, Elsevier Science B.V., North Holland, Amsterdam, 2010, pp. 617–690.

**Melino, Angelo and Alan X. Yang**, "State-Dependent Preferences can Explain the Equity Premium Puzzle," *Review of Economic Dynamics*, 2003, *6(4)*, 806–830.

**Rudebusch, Glenn D. and Eric T. Swanson**, "The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks," *American Economic Journal: Macroeconomics*, Forthcoming. Federal Reserve Bank of San Francisco Working Paper 2008-31.

**Swanson, Eric T.**, "Risk Aversion and the Labor Margin in Dynamic Equilibrium Models," *American Economic Review*, 2011.

**Woodford, Michael**, *Interest and Prices*, Princeton University Press, Princeton, NJ, 2003.

# A Results from the text

## A.1 Price of a utility claim and the SDF

The utility claim pays $U_t U_{C,t}^{-1}$ as its dividend. We confirm that its cum-dividend price is $W_{U,t} = V_t^{1-\rho} B_t^{-1} U_{C,t}^{-1} / (1-\beta)$ by simply inserting this guess into the Euler equation,

$$
\begin{aligned}
W_{U,t} &= U_t U_{C,t}^{-1} + E_t \left[ W_{U,t+1} M_{t+1} \right] &(53)\\
V_t^{1-\rho} B_t^{-1} U_{C,t}^{-1} / (1-\beta) &= U_t U_{C,t}^{-1} + E_t \left[ \frac{V_{t+1}^{1-\rho} B_{t+1}^{-1} U_{C,t+1}^{-1}}{(1-\beta)} \beta \frac{B_{t+1}}{B_t} \frac{U_{C,t+1}}{U_{C,t}} \frac{V_{t+1}^{\rho-\alpha_t}}{\left(E_t V_{t+1}^{1-\alpha_t}\right)^{\frac{\rho-\alpha_t}{1-\alpha_t}}} \right] &(54)\\
V_t^{1-\rho} &= (1-\beta) B_t U_t + \beta E_t \left[ V_{t+1}^{1-\rho} \frac{V_{t+1}^{\rho-\alpha_t}}{\left(E_t V_{t+1}^{1-\alpha_t}\right)^{\frac{\rho-\alpha_t}{1-\alpha_t}}} \right] &(55)\\
V_t^{1-\rho} &= (1-\beta) B_t U_t + \beta E_t \left[ V_{t+1}^{1-\alpha_t} \right]^{\frac{1-\rho}{1-\alpha_t}} &(56)
\end{aligned}
$$

The last line shows that the guess for the price of the utility claim was in fact correct, since the Euler equation is guaranteed to hold.

Next, we consider the return on the utility claim. We have

$$
\begin{aligned}
R_{U,t+1} &= \frac{V_{t+1}^{1-\rho} B_{t+1}^{-1} U_{C,t+1}^{-1} / (1-\beta)}{V_t^{1-\rho} B_t^{-1} U_{C,t}^{-1} / (1-\beta) - U_t U_{C,t}^{-1}} &(57)\\
&= \frac{V_{t-1}^{1-\rho}}{V_t^{1-\rho} - (1-\beta) B_t U_t} \left( \frac{B_{t+1} U_{C,t+1}}{B_t U_{C,t}} \right)^{-1} &(58)\\
&= \frac{V_{t+1}^{1-\rho}}{\beta \left(E_t V_{t+1}^{1-\alpha_t}\right)^{\frac{1-\rho}{1-\alpha}}} \left( \frac{B_{t+1} U_{C,t+1}}{B_t U_{C,t}} \right)^{-1} &(59)\\
\left( \frac{V_{t+1}^{1-\rho}}{\left(E_t V_{t+1}^{1-\alpha_t}\right)^{\frac{1-\rho}{1-\alpha_t}}} \right)^{\frac{\rho-\alpha_t}{1-\rho}} &= R_{U,t+1}^{\frac{\rho-\alpha_t}{1-\rho}} \beta^{\frac{\rho-\alpha_t}{1-\rho}} \left( \frac{U_{C,t+1}}{U_{C,t}} \right)^{\frac{\rho-\alpha_t}{1-\rho}} \left( \frac{B_{t+1}}{B_t} \right)^{\frac{\rho-\alpha_t}{1-\rho}} &(60)
\end{aligned}
$$

Now substitute the return into the SDF:

$$M_{t+1} = \beta \frac{B_{t+1}}{B_t} \frac{U_{C,t+1}}{U_{C,t}} R_{U,t+1}^{\frac{\rho-\alpha_t}{1-\rho}} \beta^{\frac{\rho-\alpha_t}{1-\rho}} \left( \frac{U_{C,t+1}}{U_{C,t}} \right)^{\frac{\rho-\alpha_t}{1-\rho}} \left( \frac{B_{t+1}}{B_t} \right)^{\frac{\rho-\alpha_t}{1-\rho}} \tag{61}$$

$$= \beta^{\frac{1-\alpha_t}{1-\rho}} \left( \frac{U_{C,t+1} B_{t+1}}{U_{C,t} B_t} \right)^{\frac{1-\alpha_t}{1-\rho}} R_{U,t+1}^{\frac{\rho-\alpha_t}{1-\rho}} \tag{62}$$

This is the formula from the text.

### A.2 First-order condition for wage setting

$$0 = E_t^* \sum_{k=0}^{\infty} \xi_w^k \left[ \psi_{t+k} \frac{1+\theta_w}{\theta_w} \left( \frac{W_{t+k}(j)}{W_{t+k}} \right)^{-\frac{1+\theta_w}{\theta_w}} \frac{N_{t+k}}{W_{t+j}(j)} - \lambda_{t+k} N_{t+k}(j) \right] \tag{63}$$

$$= E_t^* \sum_{k=0}^{\infty} \xi_w^k \left[ \left( \frac{dV_t}{dN_{t+k}(j)} + \frac{dV_t}{dC_{t+k}} \frac{W_{t+k}(j)}{P_t} \right) (1+\theta_w) - \theta_w \frac{dV_t}{dC_{t+k}} \frac{W_{t+k}(j)}{P_t} \right] N_{t+k}(j) \tag{64}$$

$$= E_t^* \sum_{k=0}^{\infty} \xi_w^k \left[ \frac{dV_t}{dN_{t+k}(j)} (1+\theta_w) + \frac{dV_t}{dC_{t+k}} \frac{W_{t+k}(j)}{P_t} \right] N_{t+k}(j) \tag{65}$$

$$= E_t^* \sum_{k=0}^{\infty} \xi_w^k \frac{dV_t/dC_{t+k}}{dV_t/dC_t} \left[ \frac{dV_t/dN_{t+k}(j)}{dV_t/dC_{t+k}} (1+\theta_w) + \frac{W_{t+k}(j)}{P_t} \right] N_{t+k}(j) \tag{66}$$

## B Approximation method

This section proceeds as follows. First, I fix notation for a general set of equilibrium conditions. Next, I describe the specifics of how to solve for the model's equilibrium dynamics. Third, I show that the essentially affine approximation method delivers bond-pricing formulas in the essentially affine class of Duffee (2002).

### B.1 Equilibrium conditions

Denote the vector of the variables in the model (including the exogenous processes) as $X_t$. In addition to the various variables described above that track the state of the economy and the shocks, $X_t$ will include the price/dividend ratio of the utility portfolio, the return on the utility portfolio, and the marginal utility of consumption. The vector of fundamental shocks in the model is denoted $\varepsilon_t \equiv [\varepsilon_{mp,t}, \varepsilon_{z,t}, \varepsilon_{b,t}, \varepsilon_{\mu,t}, \varepsilon_{g,t}, \varepsilon_{p,t}, \varepsilon_{w,t}, \varepsilon_{\alpha,t}, \varepsilon_{\pi*,t}]$.

The equations determining the equilibrium of the model take the form

$$0 = G\left(X_t, X_{t+1}, \sigma\varepsilon_{t+1}\right) \tag{67}$$

where the expectation operator may appear in the function $G$.There is one equation for each variable in the model. The new parameter $\sigma$ is the usual parameter used in perturbation approximations that controls the variance of the shocks. In the true model, $\sigma = 1$, but we will consider an approximation around the point $\sigma = 0$.

We assume that there is a solution $H$ to the system $G$ taking the form

$$X_{t+1} = H\left(X_t, \sigma\varepsilon_{t+1}, \sigma\right) \tag{68}$$

such that

$$0 = G\left(X_t, H\left(X_t, \sigma\varepsilon_{t+1}, \sigma\right), \sigma\varepsilon_{t+1}\right) \tag{69}$$

The goal is then to find an approximation to $H$.

Define the non-stochastic steady-state as the point $\bar{X}$ such that

$$0 = G\left(\bar{X}, \bar{X}, 0\right) \tag{70}$$

For the sake of clarity about the approximations, we begin with a definition:

**Definition 1** *(Judd, 1999) A function $g^j(x)$ is called "$j$-th order accurate for $g(x)$" around the point $x_0$ if*

$$\lim_{x \to x_0} \frac{\left|\left|g^j(x) - g(x)\right|\right|}{\left|\left|x - x_0\right|\right|^j} = 0 \tag{71}$$

*Equivalently, $\frac{d^k g(x_0)}{dx^k} = \frac{d^k g^j(x_0)}{dx^k}$ for all $k \leq j$.*

$g^j$ is sometimes called a $j$-th order approximation for $g$, but I use the terminology in definition (1) because it emphasizes the fact that $g^j$ need not be a Taylor series and may involve terms of order higher than $j$ (i.e. $x^{j+1}$). Dew-Becker (2011b) derives the following useful result:

**Theorem 2** *(Dew-Becker, 2011b) Suppose* $G^j(X_t, X_{t+1}, \sigma\varepsilon_{t+1})$ *is* $j$*-th order accurate for* $G(X_t, X_{t+1}, \sigma\varepsilon_{t+1})$ *around the non-stochastic steady-state of the model,* $\{X_t, X_{t+1}, \sigma\} = \{\bar{X}, \bar{X}, 0\}$*. Furthermore, suppose we have a function* $H^j$ *that solves the approximate equilibrium conditions* $G^j$*, i.e.*

$$0 = G^j\left(X_t, H^j(X_t, \sigma\varepsilon_{t+1}, \sigma), \sigma\varepsilon_{t+1}\right) \tag{72}$$

*then* $H^j$ *is* $j$*-th order accurate for* $H$ *around the non-stochastic steady state.*

What this result means is that to find a first-order approximation to $H$, it is sufficient to take a first-order accurate approximation to $G$, denoted $G^1$, and then find a function $H^1$ such that

$$0 = G^1\left(X_t, H^1(X_t, \sigma\varepsilon_{t+1}, \sigma), \sigma\varepsilon_{t+1}\right) \tag{73}$$

holds exactly. $H^1$ will be first-order accurate for $H$. To be clear, the fact that $H^1$ is first-order accurate does not mean that it must be fully linear, simply that its first derivatives at the non-stochastic steady state are identical to those for $H$.

The equations $G$ can be divided into two types: those that do not involve taking expectations over the SDF and those that do.

$$G(X_t, X_{t+1}, \varepsilon_{t+1}) = \begin{bmatrix} D(X_t, X_{t+1}, \sigma\varepsilon_{t+1}) \\ E_t\left[M(X_t, X_{t+1}, \sigma\varepsilon_{t+1}) F(X_t, X_{t+1}, \sigma\varepsilon_{t+1})\right] \end{bmatrix} \tag{74}$$

where $D$ and $F$ are vector-valued functions and $M$ is the (scalar-valued) stochastic discount factor. Note that this formulation does not actually restrict $F$. Specifically, suppose there were a set of equilibrium conditions $1 = E_t h(x_t, x_{t+1}, \sigma\varepsilon_{t+1})$, i.e. that do not involve the SDF. We could simply say that $F(x_t, x_{t+1}, \sigma\varepsilon_{t+1}) \equiv h(x_t, x_{t+1}, \sigma\varepsilon_{t+1}) / M(x_t, x_{t+1}, \sigma\varepsilon_{t+1})$. Note, though, that in this model, all expectational equations involve the SDF.

For the equations that do not involve the SDF, I use standard perturbation methods and simply take a log-linear approximation. We approximate $D$ as

$$0 = \log\left(D(\exp(x_t), \exp(x_{t+1}), \sigma\varepsilon_{t+1}) + 1\right) \tag{75}$$

$$0 \approx d_0 + d_x \hat{x}_t + d_{x'} \hat{x}_{t+1} + d_\varepsilon \sigma\varepsilon_{t+1} \tag{76}$$

where the terms $d_0$, $d_x$, $d_{x'}$, and $d_\varepsilon$ are coefficients from a Taylor-series approximation and

$$\begin{aligned} x_t &\equiv \log X_t \\ \hat{x}_t &\equiv \log X_t - \log \bar{X} \end{aligned}$$

### B.2 Linearizing the Euler equations

I now show that if we log-linearize the function $F$, we can transform (46) into a linear condition that can be solved alongside the remaining equations. $M_{t+1}$ will not even be log-linear in the state variables, but we will be able to state the equilibrium conditions as a set of linear expectational difference equations.

First, guess that the equilibrium dynamics of the model (i.e. $H^j$) take the form

$$\hat{x}_{t+1} = C + \Phi \hat{x}_t + \Psi \sigma \varepsilon_{t+1} \tag{77}$$

where $\hat{x}_t \equiv \log(X_t) - \log(\bar{X})$. We confirm in the end that the solution is actually in this form. Next, define the matrices $\Gamma_{UC}$, $\Gamma_b$, and $\Gamma_r$ as matrices that select individual elements of $x_t$ such that

$$\hat{u}_{C,t} = \Gamma_{UC} \hat{x}_t, \ \hat{b}_t = \Gamma_b \hat{x}_t, \ \hat{r}_{U,t} = \Gamma_r \hat{x}_t \tag{78}$$

where lower-case letters with circumflexes denote log deviations from non-stochastic steady-state values. That is, $\hat{u}_{C,t} \equiv \log U_{C,t} - \log \bar{U}_C$, etc.For the sake of arithmetical convenience, also define an auxiliary variable $\zeta_t \equiv \frac{1-\alpha_t}{1-\rho}$.

#### B.2.1 The essentially affine SDF

The SDF is

$$M_{t+1} = \beta^{\frac{1-\alpha_t}{1-\rho}} \left( \frac{U_{C,t+1}}{U_{C,t}} B_t \frac{1 - \beta B_{t+1}}{1 - \beta B_t} \right)^{\frac{1-\alpha_t}{1-\rho}} R_{U,t+1}^{\frac{\rho-\alpha_t}{1-\rho}} \tag{79}$$

$M_{t+1}$ is completely log-linear in the endogenous variables $U_{C,t}$ and $R_{U,t+1}$, but it is not log-linear in $B_t$ and $B_{t+1}$. If not for these terms, we could use the exact formula for $M_{t+1}$ in what follows. Since those terms are non-linear, I use the approximations,

$$\log \frac{1-\beta \exp\left(b_{t+1}\right)}{\exp\left(-b_t\right)-\beta} \approx \frac{1}{1-\beta} b_t-\frac{\beta}{1-\beta} b_{t+1} \tag{80}$$

$$m_{t+1}^{(1)} \equiv \bar{m}+\zeta_t\left(\Delta \hat{u}_{C, t+1}+\frac{1}{1-\beta} b_t-\frac{\beta}{1-\beta} b_{t+1}\right)+\zeta_t \hat{r}_{U, t+1} \tag{81}$$

$m_{t+1}^{(1)}$ is first-order accurate for $m_{t+1}$, hence the superscript. In the continuous time limit, this formula becomes exact. It is straightforward to show that for any first-order approximation to a function $F$, denoted $F^j$, $E_t\left[\exp\left(m_{t+1}^{(1)}\right) F_{t+1}^j\right]$ will be first-order accurate for the true Euler equation, $E_t\left[\exp\left(m_{t+1}\right) F_{t+1}\right]$, around the non-stochastic steady-state, and so theorem (2) will apply (see Dew-Becker and Baqaee, 2011).

The Euler equation for the return on the utility portfolio is

$$1=E_t \exp\left(\zeta_t\left(\Delta \hat{u}_{C, t+1}+\frac{1}{1-\beta} b_t-\frac{\beta}{1-\beta} b_{t+1}\right)+\zeta_t \hat{r}_{U, t+1}\right) \tag{82}$$

Taking logs of both sides and taking advantage of log-normality gives

$$\begin{aligned} 0 = & E_t \hat{u}_{C, t+1}-\hat{u}_{C, t}+\frac{1}{1-\beta} b_t-\frac{\beta}{1-\beta} E_t b_{t+1}+E_t \hat{r}_{U, t+1} \\ & +\frac{1}{2} \zeta_t \sigma^2\left(\Gamma_{UC}-\frac{\beta}{1-\beta} \Gamma_b+\Gamma_r\right) \Psi \Psi^{\prime}\left(\Gamma_{UC}-\frac{\beta}{1-\beta} \Gamma_b+\Gamma_r\right)^{\prime} \end{aligned} \tag{83}$$

Moreover, this implies that the approximated SDF can be written as

$$\begin{aligned} m_{t+1}^{(1)} = & \zeta_t\left(\Gamma_{UC}-\frac{\beta}{1-\beta} \Gamma_b+\Gamma_r\right) \Psi \sigma \varepsilon_{t+1}-\Gamma_r\left(\Phi x_t+\Psi \sigma \varepsilon_{t+1}\right)-\bar{r} \\ & -\frac{1}{2} \zeta_t^2 \sigma^2\left(\Gamma_{UC}-\frac{\beta}{1-\beta} \Gamma_b+\Gamma_r\right) \Psi \Psi^{\prime}\left(\Gamma_{UC}-\frac{\beta}{1-\beta} \Gamma_b+\Gamma_r\right)^{\prime} \end{aligned} \tag{84}$$

which is the essentially affine form from the text. The reason that this form is useful is that any time the SDF is essentially affine, we can obtain an exact expression for $E_t \exp\left(m_{t+1}+f_0+f_x x_t+f_{x^{\prime}} x_{t+1}\right)$. It also means that we can price any asset whose payoffs are linear in the endogenous variables, including real and nominal bonds.

#### B.2.2 Approximation to $F$

Next, we take a first-order Taylor approximation to $\log F$ such that $\log F(x_t, x_{t+1}, \varepsilon_{t+1}) \approx f_0 + f_x x_t + f_{x'} x_{t+1}$, giving

$$1 = E_t \exp\left(\hat{m}_{t+1}^{(1)} + f_x x_t + f_{x'} x_{t+1}\right) \tag{85}$$

Taking logs and evaluating the expectation yields

$$\begin{aligned} 0 &= -E_t \hat{r}_{U,t+1} + f_x x_t + f_{x'} E_t x_{t+1} + \frac{1}{2}\sigma^2 (-\Gamma_r + f_{x'}) \Psi\Psi' (-\Gamma_r + f_{x'})' \\ &\quad + \zeta_t \sigma^2 \left(\Gamma_{UC} - \frac{\beta}{1-\beta}\Gamma_b + \Gamma_r\right) \Psi\Psi' (-\Gamma_r + f_{x'})' \end{aligned} \tag{86}$$

(86) is the equation that we ultimately place into the system to be solved. It is completely linear in both $x_t$ and $\zeta_t$ (equivalently, $\alpha_t$).

#### B.2.3 Solution

Since every equation in the system is now linear in the variables of the model, we can solve the system for the parameters $\Phi$ and $\Psi$ from (77). Specifically, we solve the following system,

$$\begin{aligned} 0 &= -E_t \hat{r}_{U,t+1} + f_x x_t + f_{x'} E_t x_{t+1} + \frac{1}{2}\sigma^2 (-\Gamma_r + f_{x'}) \Psi\Psi' (-\Gamma_r + f_{x'})' \\ &\quad + \zeta_t \sigma^2 \left(\Gamma_{UC} - \frac{\beta}{1-\beta}\Gamma_b + \Gamma_r\right) \Psi\Psi' (-\Gamma_r + f_{x'})' \end{aligned} \tag{87}$$

$$0 = d_0 + d_x x_t + d_{x'} x_{t+1} + d_\varepsilon \sigma \varepsilon_{t+1} \tag{88}$$

where $\sigma = 1$ in the stochastic equilibrium that we approximate. This system can be solved using, for example, Sims' (2001) Gensys algorithm.

The last wrinkle here is that we cannot simply insert (86) into the set of equations to be solved since it involves the matrix $\Psi$, which is one of the unknown structures we are solving for. I deal with this with a simple fixed-point iteration: I begin with the equations that we obtain from perturbation,

$$\begin{aligned} 0 &= -\Gamma_r E_t x_{t+1} + f_x x_t + f_{x'} E_t x_{t+1} \\ 0 &= d_0 + d_x \hat{x}_t + d_{x'} \hat{x}_{t+1} + d_\varepsilon \sigma \varepsilon_{t+1} \end{aligned}$$

which will deliver an initial value of $\Psi$, denoted $\Psi^{(1)}$. We then use $\Psi^{(1)}$ to change the equilibrium condition to take the form

$$\begin{aligned} 0 &= -\Gamma_r \Phi x_t - \bar{r} + f_x x_t + f_{x'} \Phi x_t + \frac{1}{2}\sigma^2 \left(-\Gamma_r + f_{x'}\right) \Psi^{(1)} \Psi^{(1)\prime} \left(-\Gamma_r + f_{x'}\right)' \\ &\quad + \zeta_t \sigma^2 \left(\Gamma_{UC} - \frac{\beta}{1-\beta}\Gamma_b + \Gamma_r\right) \Psi^{(1)} \Psi^{(1)\prime} \left(-\Gamma_r + f_{x'}\right)' \qquad (89) \\ 0 &= d_0 + d_x \hat{x}_t + d_{x'} \hat{x}_{t+1} + d_\varepsilon \sigma \varepsilon_{t+1} \qquad (90) \end{aligned}$$

which delivers a value $\Psi^{(2)}$. Then simply iterate to convergence. I treat parameter sets for which the iteration diverges as inadmissible, setting the marginal likelihood to zero.

### B.3 Discussion

So why does the essentially affine approximation differ from a first-order perturbation? Perturbation involves a first-order approximation in $\{x_t, \sigma\}$, so the terms involving $\zeta_t \sigma^2$ are viewed as third-order. But for the purposes of solving the system of equilibrium equations, we take $\sigma^2$ as a constant. That is why (86) is linear for our purposes.

Now note that when $\sigma = 0$, the equilibrium condition (86) can be shown to take exactly the form that it would under perturbation. That is, the variance terms drop out, and we only have

$$0 = -E_t \hat{r}_{U,t+1} + f_x x_t + f_{x'} E_t x_{t+1} \qquad (91)$$

or, equivalently,

$$0 = E_t \Delta \hat{u}_{C,t+1} + \frac{1}{1-\beta} b_t - \frac{\beta}{1-\beta} b_{t+1} + f_x x_t + f_{x'} E_t x_{t+1} \qquad (92)$$

which may be familiar as the usual perturbation approximation. Note here that $\zeta_t$ no longer appears, confirming the usual result that risk aversion does not appear in the first-order perturbation.

As noted above, both perturbation and the essentially affine approximation are first-order accurate for the true transition function for $x_t$ around the non-stochastic steady state. The only way to compare the approximations is to compare the Euler equation errors that they generate on the domain that the stochastic model reaches. For a simple version of the RBC model, Dew-Becker (2011b) shows that the Euler equation errors for the essentially affine approximation are on average roughly two orders of magnitude smaller than those for a first-order perturbation, and are competitive with a third-order perturbation.

### B.4 Bond pricing

To solve for bond prices, we guess that bond prices are log-linear in the vector of state variables, so that

$$p_{n,t} = A_n + B_n x_t \tag{93}$$

where $p_{n,t}$ is the price of a zero-coupon bond that matures on date $t+n$ and pays 1 unit of consumption. We can also write the price of a nominal bond as $p^{\$}_{n,t}$. Using the formula for the SDF from above, we have

$$\begin{aligned}
\exp\left(p^{\$}_{n,t}\right) &= E_t \exp\left(\begin{array}{c} \zeta_t\left(\Gamma_{UC} - \frac{\beta}{1-\beta}\Gamma_b + \Gamma_r\right)\Psi\varepsilon_{t+1} - \Gamma_r\left(\Phi x_t + \Psi\varepsilon_{t+1}\right) - \bar{r} + A_{n-1} + \left(B_{n-1} - \Gamma_\pi\right)\left(\Phi x_t + \Psi\varepsilon_{t+1}\right) \\ -\frac{1}{2}\zeta_t^2\sigma^2\left(\Gamma_{UC} - \frac{\beta}{1-\beta}\Gamma_b + \Gamma_r\right)\Psi\Psi'\left(\Gamma_{UC} + \Gamma_b + \Gamma_r\right)' \end{array}\right) \\
&= -\bar{r} - \Gamma_r\Phi x_t + A_{n-1} + \left(B_{n-1} - \Gamma_\pi\right)\Phi x_t + \frac{1}{2}\sigma^2\left(B_{n-1} - \Gamma_\pi - \Gamma_r\right)\Psi\Psi'\left(B_{n-1} - \Gamma_\pi - \Gamma_r\right)' \\
&\quad +\sigma^2\left(\Gamma_{UC} - \frac{\beta}{1-\beta}\Gamma_b + \Gamma_r\right)\Psi\Psi'\left(B_{n-1} - \Gamma_\pi - \Gamma_r\right)'\Gamma_\zeta x_t
\end{aligned}$$

Matching coefficients gives

$$A_n = -\bar{r} + A_{n-1} + \frac{1}{2}\sigma^2\left(B_{n-1} - \Gamma_\pi - \Gamma_r\right)\Psi\Psi'\left(B_{n-1} - \Gamma_\pi - \Gamma_r\right)' \tag{96}$$

$$B_n = -\Gamma_r + \left(B_{n-1} - \Gamma_\pi\right)\Phi + \left(\Gamma_{UC} - \frac{\beta}{1-\beta}\Gamma_b + \Gamma_r\right)\Psi\Psi'\left(B_{n-1} - \Gamma_\pi - \Gamma_r\right)'\Gamma_\zeta \tag{97}$$

## C Estimation

Much of the analysis discusses the behavior of the model around the posterior mode (i.e. the peak of the posterior distribution; it is also a maximum-likelihood estimate penalized by the prior distribution). I also start the Metropolis–Hastings chain from that point. To search for the posterior mode, I begin by running a genetic algorithm on a population of 60 points drawn from the prior distribution. The genetic algorithm searches the parameter space by mixing parameter sets in the population and also allowing random mutations. After 30 iterations of the genetic algorithm, I take the point in the population with the highest posterior density and use it as the starting point for Chris Sims' CSMINWEL algorithm, which is a derivative-based hill-climbing algorithm that is designed for DSGE models. When CSMINWEL gets stuck, I also try the standard simplex algorithm.

I ran this combined search 2500 times (each search takes roughly an hour, so access to a large computing cluster was essential). The point that I am calling the posterior mode was found to be the peak in fewer than 100 of the searches. In other words, it is extremely difficult to find the peak of the posterior likelihood. In general, I found that it was far easier to find the posterior mode when risk aversion or the inflation target was held fixed, and easier still when bond prices were also dropped from the estimation. Even though the priors help to add curvature to the posterior likelihood surface, I still find many local maxima, a problem that plagues the bond-pricing literature. Furthermore, since the model is so highly constrained there is no straightforward way to use the more recent estimation algorithms for ffine term structure models proposed by, for example, Hamilton and Wu (2011) and Joslin, Singleton, and Zhu (2010).

I simulate the posterior distribution using the adaptive random-walk Metropolis–Hastings algorithm of Haario, Saksman, and Tamminen (2001). I initialize the chain at the posterior mode. For the proposal distribution, I begin with a normal distribution whose variance matrix is equal to that of the prior, multiplied by $(2.38^2)/d$, where $d$ is the dimension of the parameter vector (49), which is the optimal scaling factor of Gelman, Roberts, and Gilks (1996). After 10,000 iterations of the algorithm, I update the variance matrix of the proposal distribution to be equal to the observed variance matrix for the first 10,000 iterations of the chain. Subsequently, the variance is updated each on each iteration using the sample variance of the chain up to the current iteration.[10] I achieve relatively rapid mixing this way. The full chain has 500,000 draws, but it mixes well even by 100,000 draws.

[10] A more common method in the DSGE literature is to use the hessian of the posterior around the posterior mode to determine the variance of the proposal distribution. I have difficulty calculating the hessian due to numerical instability.

**Table 1. Priors and posterior modes**

| | | Description | Distribution | Priors Mean | Priors Std. Dev. | Posterior Mode | Posterior 5% | Posterior 95% | Estimates from JPT |
|---|---|---|---|---|---|---|---|---|---|
| 1 | $\alpha$ | Capital share | Normal | 0.3 | 0.05 | 0.13 | 0.10 | 0.15 | 0.17 |
| 2 | $\iota_p$ | Price indexation | Beta | 0.5 | 0.15 | 0.39 | 0.21 | 0.56 | 0.24 |
| 3 | $\iota_w$ | Wage indexation | Beta | 0.5 | 0.15 | 0.77 | 0.66 | 0.88 | 0.11 |
| 4 | $100\gamma$ | Mean technology growth | Normal | 0.5 | 0.25 | 0.48 | 0.43 | 0.52 | 0.48 |
| 5 | $\eta$ | Habit parameter | Beta | 0.5 | 0.1 | 0.52 | 0.32 | 0.71 | 0.78 |
| 6 | $\lambda_p$ | Mean price markup | Normal | 0.15 | 0.05 | 0.10 | 0.02 | 0.18 | 0.23 |
| 7 | $\lambda_w$ | Mean wage markup | Normal | 0.15 | 0.05 | 0.15 | 0.07 | 0.24 | 0.15 |
| 8 | $L_{SS}$ | Mean log hours per capita | Normal | 6.7 | 0.2 | 6.75 | 6.71 | 6.79 | N/A |
| 9 | $\pi_{SS}$ | Mean quarterly inflation | Normal | 0.5 | 0.1 | -0.20 | -0.90 | 0.74 | 0.71 |
| 10 | $100(\beta^{-1}-1)$ | Discount factor | Gamma | 0.25 | 0.1 | 0.28 | 0.21 | 0.47 | 0.13 |
| 11 | $\nu$ | Inverse Frisch elasticity | Gamma | 2 | 0.75 | 1.83 | 1.16 | 2.87 | 3.79 |
| 12 | $\xi_p$ | Price adjustment frequency | Beta | 0.66 | 0.1 | 0.67 | 0.60 | 0.71 | 0.84 |
| 13 | $\xi_w$ | Wage adjustment frequency | Beta | 0.66 | 0.1 | 0.67 | 0.57 | 0.72 | 0.7 |
| 14 | $\chi$ | Capital utilization costs | Gamma | 5 | 1 | 5.08 | 3.23 | 7.19 | 5.3 |
| 15 | $S$ | Investment adjustment costs | Gamma | 4 | 1 | 4.96 | 3.16 | 6.71 | 2.85 |
| 16 | $f_\pi$ | Taylor rule inflation | Normal | 1.7 | 0.3 | 1.89 | 1.51 | 2.51 | 2.09 |
| 17 | $f_y$ | Taylor rule output gap | Gamma | 0.125 | 0.04 | 0.08 | 0.04 | 0.16 | 0.07 |
| 18 | $f_{\Delta y}$ | Taylor rule output gap growth | Normal | 0.125 | 0.05 | 0.25 | 0.21 | 0.30 | 0.24 |
| 19 | $\varrho_R$ | Interest rate smoothing | Beta | 0.6 | 0.2 | 0.96 | 0.92 | 0.97 | 0.82 |
| 20 | $\varrho_z$ | Technology shock AR | Beta | 0.6 | 0.2 | 0.20 | 0.09 | 0.31 | 0.23 |
| 21 | $\varrho_g$ | Government spending AR | Beta | 0.6 | 0.2 | 0.99 | 0.99 | 0.99 | 0.99 |
| 22 | $\varrho_\mu$ | Investment technology AR | Beta | 0.6 | 0.2 | 0.50 | 0.34 | 0.67 | 0.72 |
| 23 | $\varrho_{\lambda p}$ | Price markup AR | Beta | 0.6 | 0.2 | 0.95 | 0.92 | 0.97 | 0.94 |
| 24 | $\varrho_{\lambda w}$ | Wage markup AR | Beta | 0.6 | 0.2 | 0.99 | 0.99 | 0.99 | 0.97 |
| 25 | $\varrho_b$ | Consumption demand shock AR | Beta | 0.6 | 0.2 | 0.77 | 0.70 | 0.80 | 0.67 |
| 26 | $\varrho_{mp}$ | Monetary policy AR | Beta | 0.4 | 0.2 | 0.19 | 0.09 | 0.30 | 0.14 |
| 27 | $\theta_p$ | Price markup MA | Beta | 0.5 | 0.2 | 0.16 | 0.02 | 0.39 | 0.77 |
| 28 | $\theta_w$ | Wage markup MA | Beta | 0.5 | 0.2 | 0.98 | 0.97 | 0.98 | 0.91 |
| 29 | $\sigma_R$ | MP shock vol. | IG(1) | 0.1 | 1 | 0.14 | 0.12 | 0.17 | 0.22 |
| 30 | $\sigma_z$ | Neutral tech. shock vol. | IG(1) | 0.5 | 1 | 0.80 | 0.69 | 0.97 | 0.88 |
| 31 | $\sigma_g$ | Gov't spending vol. | IG(1) | 0.5 | 1 | 0.29 | 0.25 | 0.35 | 0.35 |
| 32 | $\sigma_\mu$ | Investment tech. vol. | IG(1) | 0.5 | 1 | 6.21 | 3.76 | 9.39 | 6.03 |
| 33 | $\sigma_{\lambda p}$ | Price markup vol. | IG(1) | 0.1 | 1 | 0.12 | 0.10 | 0.18 | 0.14 |
| 34 | $\sigma_{\lambda w}$ | Wage markup vol. | IG(1) | 0.1 | 1 | 0.35 | 0.30 | 0.43 | 0.2 |
| 35 | $\sigma_b$ | Demand shock vol. | IG(1) | 1 | 1 | 0.41 | 0.29 | 0.63 | 0.04 |
| 36 | $\sigma_{\pi^*}$ | Inflation target vol. | IG(1) | 0.1 | 0.1 | 0.33 | 0.28 | 0.41 | N/A |
| 37 | $\sigma_{\pi^*,mp}$ | MP var. shr. in pi* | U[-1,1] | 0 | 0.58 | 0.39 | 0.19 | 0.48 | N/A |
| 38 | $\sigma_{\pi^*,z}$ | z var. shr. in pi* | U[-1,1] | 0 | 0.58 | -0.16 | -0.26 | -0.12 | N/A |
| 39 | $\sigma_{\pi^*,g}$ | g var. shr. in pi* | U[-1,1] | 0 | 0.58 | 0.00 | 0.00 | 0.01 | N/A |
| 40 | $\sigma_{\pi^*,\mu}$ | mu var. shr. in pi* | U[-1,1] | 0 | 0.58 | 0.01 | 0.00 | 0.04 | N/A |
| 41 | $\sigma_{\pi^*,\lambda p}$ | Price shock var. shr. in pi* | U[-1,1] | 0 | 0.58 | -0.04 | -0.10 | -0.01 | N/A |
| 42 | $\sigma_{\pi^*,\lambda w}$ | Wage shock var. shr. in pi* | U[-1,1] | 0 | 0.58 | -0.06 | -0.12 | -0.01 | N/A |
| 43 | $\sigma_{\pi^*,b}$ | b var. shr. in pi* | U[-1,1] | 0 | 0.58 | 0.28 | 0.18 | 0.40 | N/A |
| 44 | $\sigma_\alpha/\alpha_{SS}$ | RRA volatility/RRA mean | Beta | 0.5 | 0.2 | 0.95 | 0.83 | 0.99 | N/A |
| 45 | $\sigma_{\alpha,z}$ | z var. shr. in RRA | U[-1,1] | 0 | 0.58 | 0.00 | -0.03 | 0.05 | N/A |
| 46 | $\varrho_\alpha$ | RRA persistence | U[-1,1] | 0.5 | 0.29 | 0.77 | 0.72 | 0.83 | N/A |
| 47 | $\varrho$ | Inverse EIS | U[-1,1] | 0.5 | 0.29 | 0.76 | 0.51 | 0.90 | 1 |
| 48 | $\alpha_{SS}$ | Mean risk aversion | Normal | 15 | 5 | 18.70 | 13.42 | 26.57 | 1 |
| 49 | $\sigma_{yields}$ | Bond measurement errors (bp) | IG(1) | 13 | 33 | 8.20 | 7.66 | 8.96 | N/A |

Note: Priors, posterior mode, and percentiles of the posterior distribution from the benchmark model. The far-right column reports the parameters from JPT, where applicable.

**Table 2. Model comparison statistics**

| | Constant RRA | Constant π* |
|---|---|---|
| Likelihood ratio | 185.6 | 183.8 |
| p-value | 5.0E-41 | 1.6E-35 |
| Bayes factor | 141.50 | 142.50 |

Note: Row 1 gives the marginal likelihood of the data given the model and parameters. Row two gives the p-value for the frequentist LR test. Row 3 is the Bayes factor, the marginal probability of the data conditional on the model (i.e. integrated over the parameter space).

**Table 3. Fitting errors**

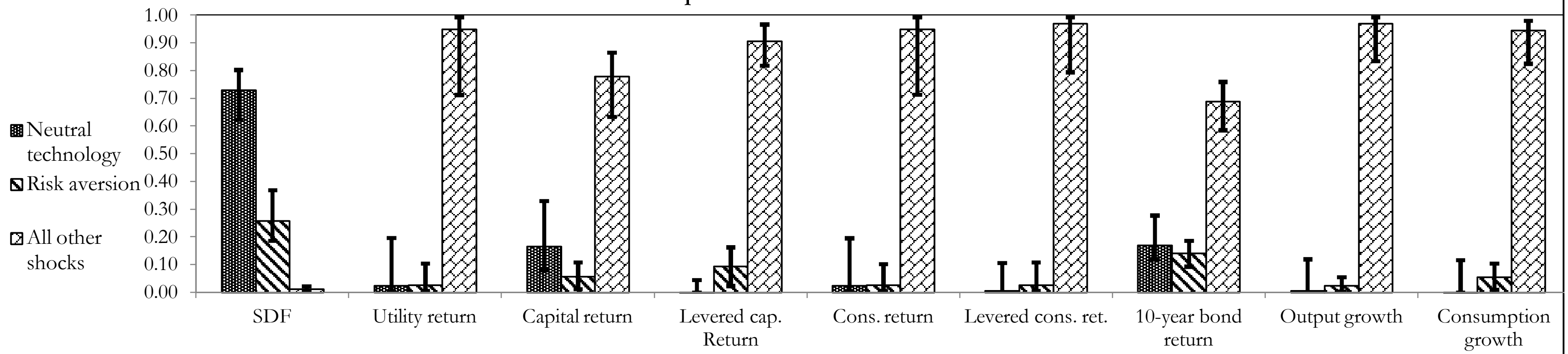
| | 1-quarter | 1-year | 2-year | 3-year | 4-year | 5-year | 10-year |
|---|---|---|---|---|---|---|---|
| PCA | 4.54 | 7.32 | 6.29 | 4.26 | 6.20 | 7.98 | 6.53 |
| Benchmark model | 0 | 8.12 | 8.12 | 8.12 | 8.12 | 8.12 | 0 |
| Constant RRA | 0 | 23.30 | 23.30 | 23.30 | 23.30 | 23.30 | 0 |
| Constant π* | 0 | 23.87 | 23.87 | 23.87 | 23.87 | 23.87 | 0 |

Note: Fitting errors measured in annualized basis points. The model-based estimates use the posterior modal estimate for the standard deviation. The 1-quarter and 10-year errors are constrained to equal zero in the structural model. The errors from PCA are the standard deviations of the residuals from regressions on the bond yields on their first three principal components.

**Table 4. One-quarter ahead variance decompositions**

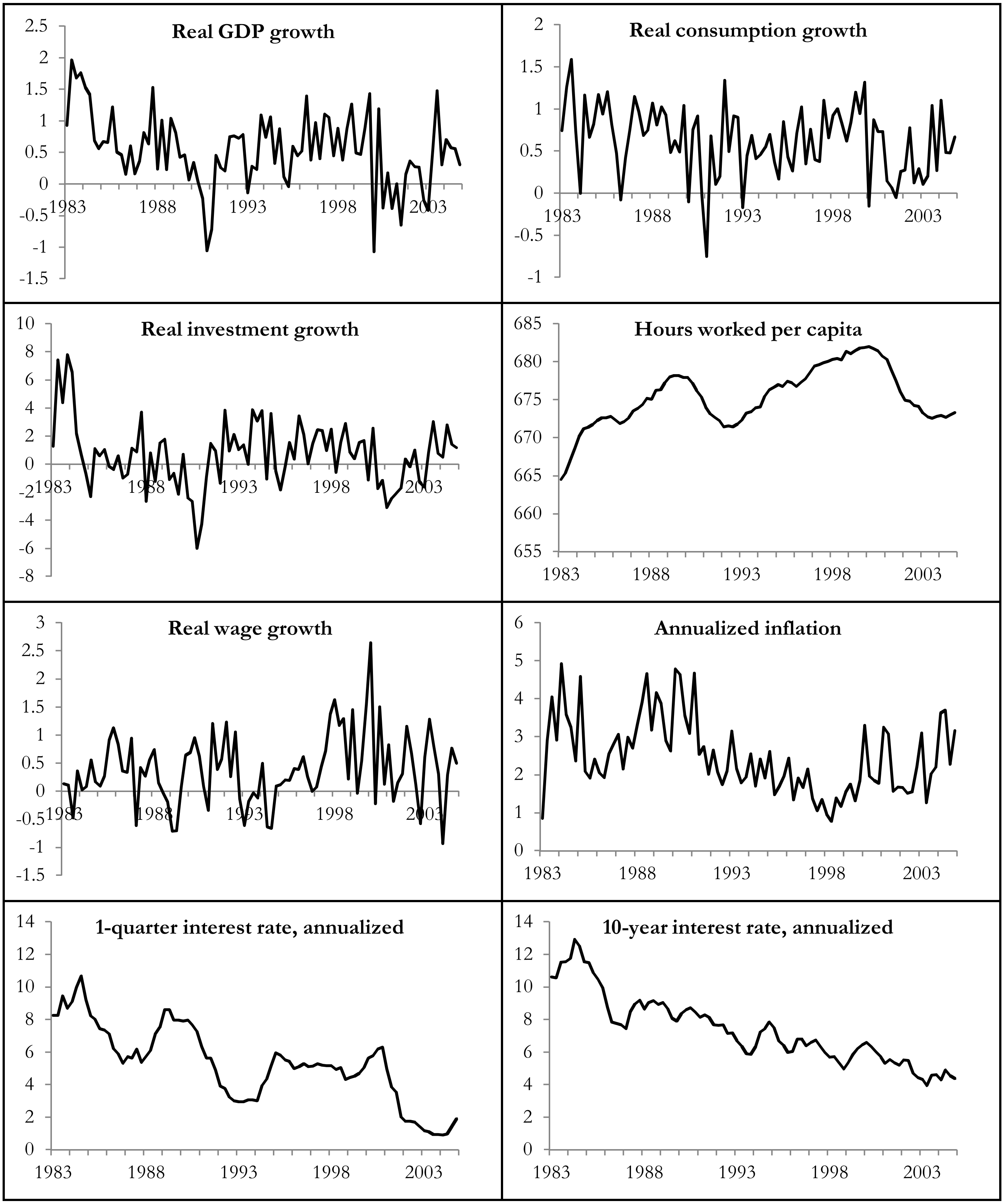
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| | SDF | Utility return | Capital return | Levered cap. Return | Cons. return | Levered cons. ret. | 10-year bond return | Output growth | Consumption growth |
| 1 Monetary policy | 0.00 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.30 | 0.03 | 0.05 |
| 2 Neutral tech. | 0.73 | 0.02 | 0.17 | 0.00 | 0.02 | 0.01 | 0.17 | 0.01 | 0.00 |
| 3 Gov't spending | 0.00 | 0.01 | 0.01 | 0.02 | 0.01 | 0.01 | 0.00 | 0.07 | 0.08 |
| 4 Investment tech. | 0.00 | 0.00 | 0.19 | 0.16 | 0.00 | 0.00 | 0.01 | 0.36 | 0.01 |
| 5 Price markup | 0.00 | 0.00 | 0.50 | 0.29 | 0.00 | 0.00 | 0.06 | 0.01 | 0.00 |
| 6 Wage markup | 0.00 | 0.00 | 0.03 | 0.00 | 0.00 | 0.01 | 0.03 | 0.01 | 0.02 |
| 7 Time preference | 0.01 | 0.87 | 0.02 | 0.24 | 0.87 | 0.87 | 0.22 | 0.00 | 0.00 |
| 8 Inflation target | 0.00 | 0.07 | 0.03 | 0.18 | 0.07 | 0.08 | 0.07 | 0.49 | 0.78 |
| 9 Risk aversion | 0.26 | 0.03 | 0.06 | 0.09 | 0.03 | 0.03 | 0.14 | 0.02 | 0.05 |
| *Moments:* | | | | | | | | | |
| 10 Standard deviation | 0.43 | 3.25 | 0.74 | 1.26 | 3.25 | 6.80 | 19.74 | 1.22 | 1.10 |
| 11 Correl. w/ SDF | 1.00 | -0.31 | -0.25 | 0.12 | -0.31 | -0.24 | -0.49 | -0.15 | -0.16 |
| 12 Expected return | N/A | 0.44 | 0.08 | -0.06 | 0.44 | 0.71 | 4.15 | N/A | N/A |

**Variance decompositions and 95% credible intervals**



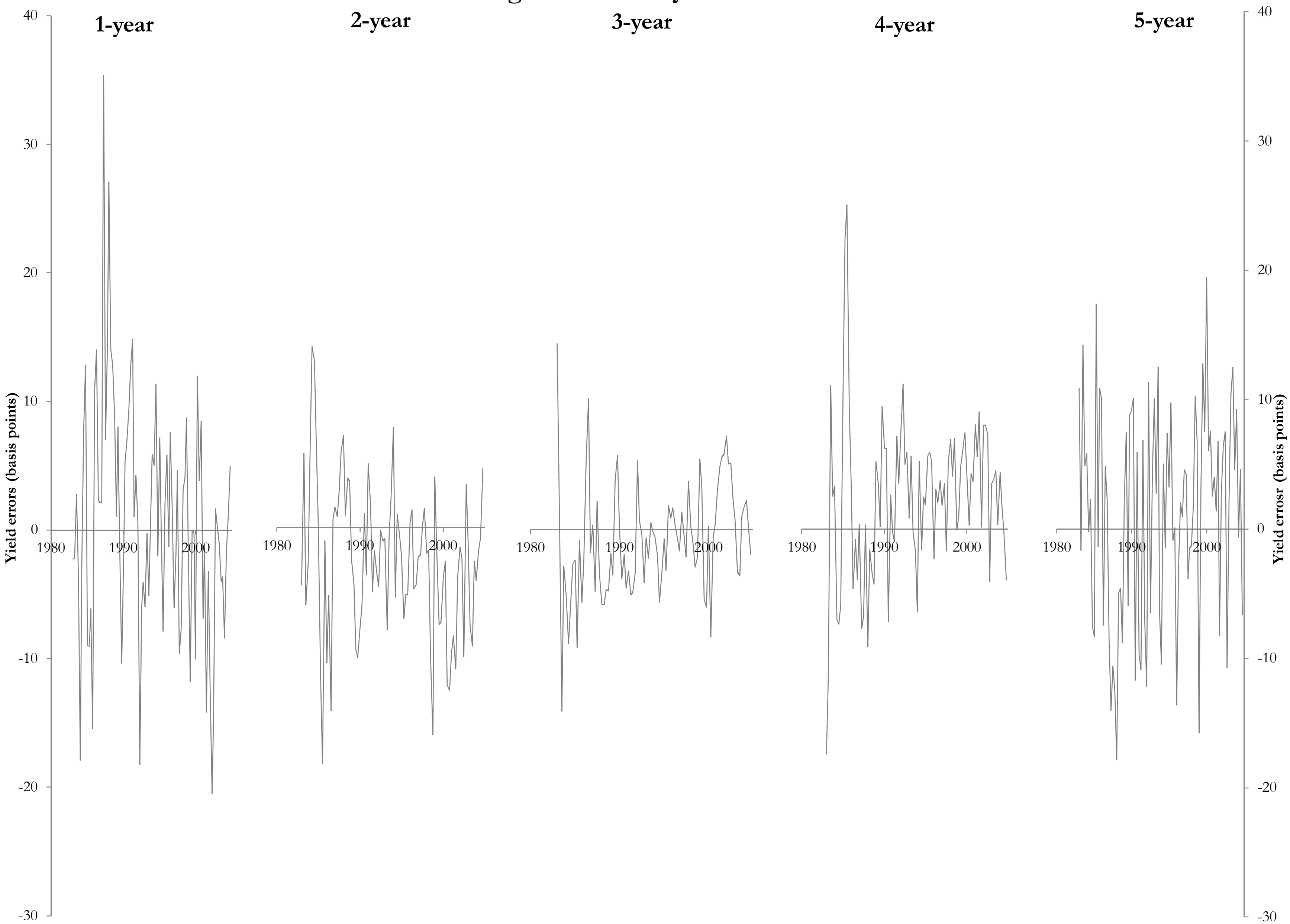

Note: decompositions of one-quarter ahead forecast error. Levered returns for capital and consumption claims assume that the investor finances half the purchase price of the given claim with a 10-year nominally riskless bond. The 10-year return is the one-quarter return from holding a 10-year nominally riskless bond. The moments in rows 10–12 are annualized. The black bars in the figure give the 95 percent credible region based on random draws from the posterior density.

# Figure 1. Data series for estimation



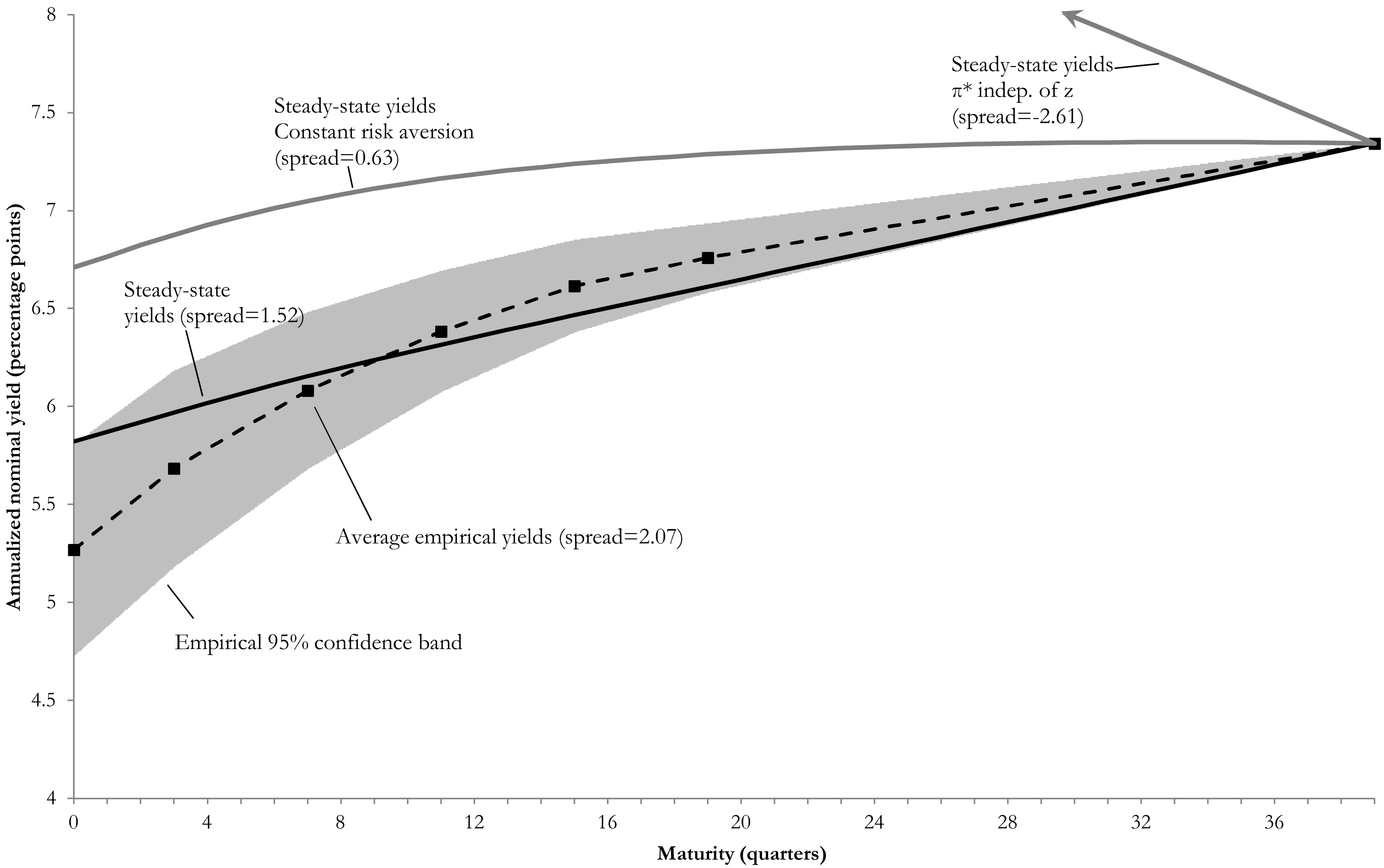

Note: No variables are detrended. GDP, consumption, and investment are obtained from the BEA. Compensation per hour, and inflation are obtained from the BLS. Hours worked is obtained from Valerie Ramey's website. The one-quarter yield is the Fama risk-free rate. The ten-year yield is from Gurkaynak, Sack, and Wright (2006).

## Figure 2. Bond yield errors



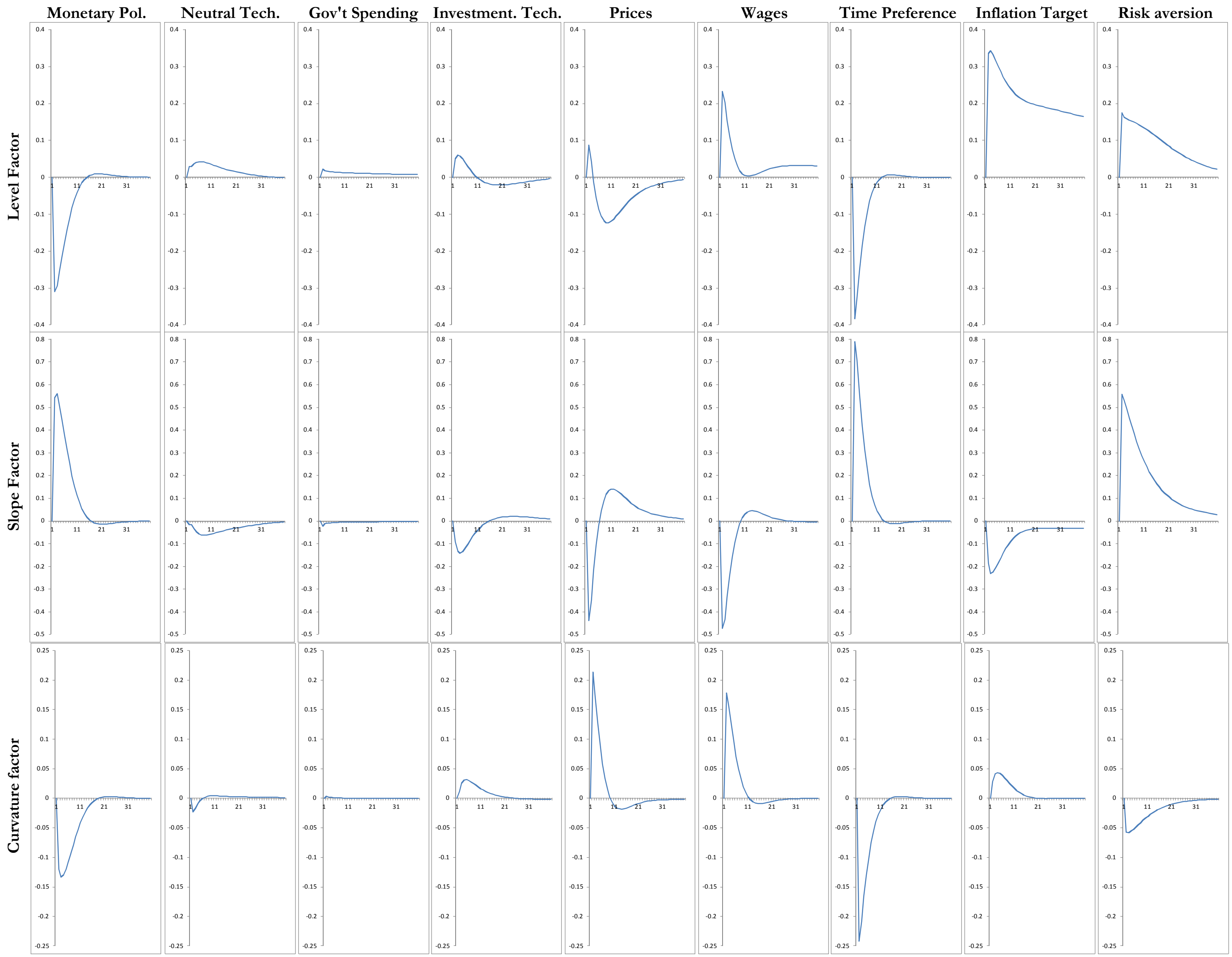

Note: each axis plots the measurement errors in basis points for one of the bond yields. Errors are measured from the Kalman-filtered estimates at the posterior mode.

# Figure 3. Steady-state nominal bond yields



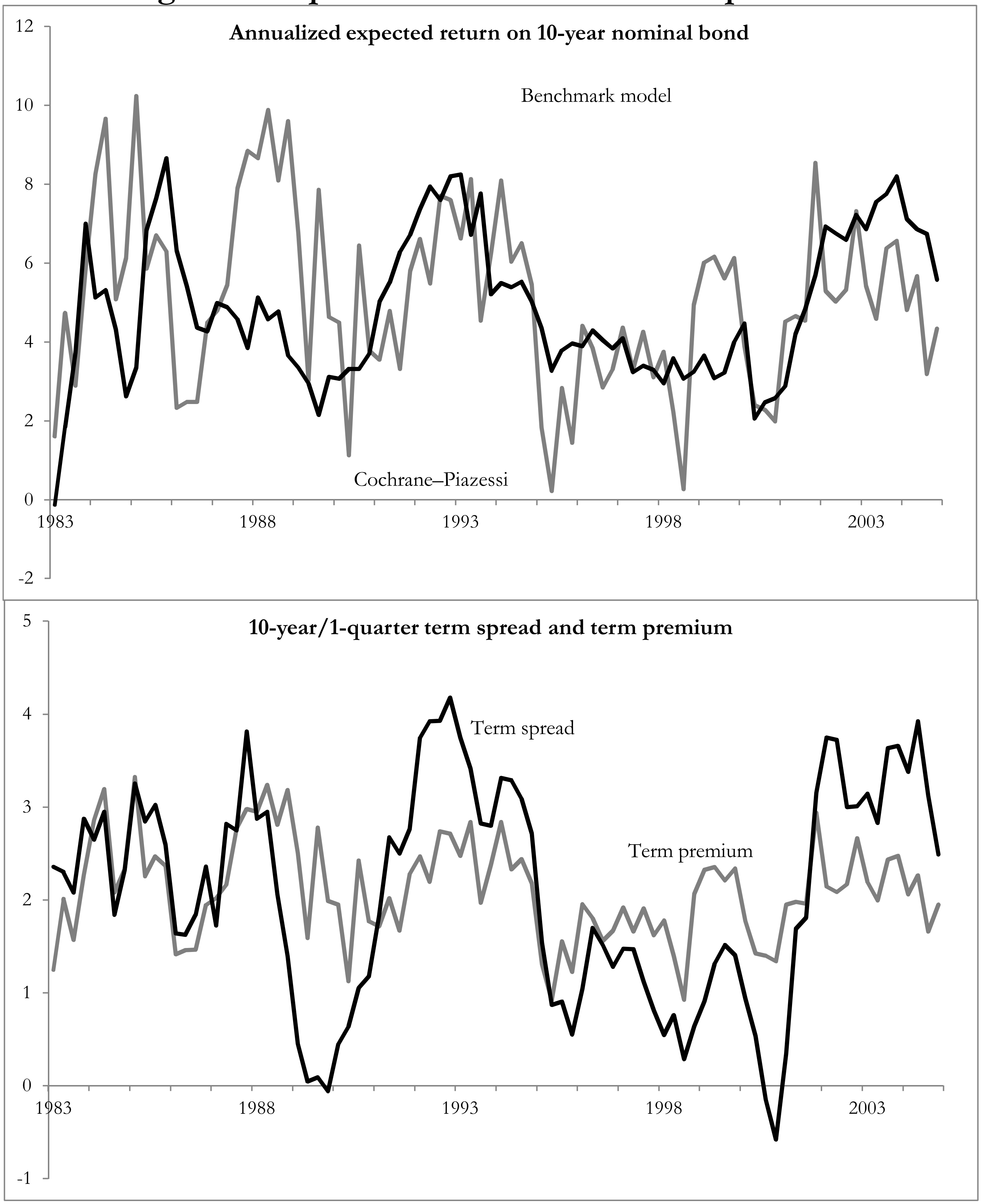

Note: The solid black line gives the yield curve at the model's steady state; the grey lines are for the model with constant risk aversion and where the inflation target is unaffected by shocks to labor-neutral technology. The other parameters are not reestimated. All the linear are normalized to match the 10-year yield exactly, so the plot measures steady-state spreads. Boxes are average sample yields. The grey area is the 95% confidence band for the average yields relative to the 10-year yield, calculated using the Newey–West method with 6 lags. The solid black line gives the yield curve at the model's steady state, normalized to match the 10-year yield exactly.

## Figure 4. Responses of term structure factors to orthogonalized shocks



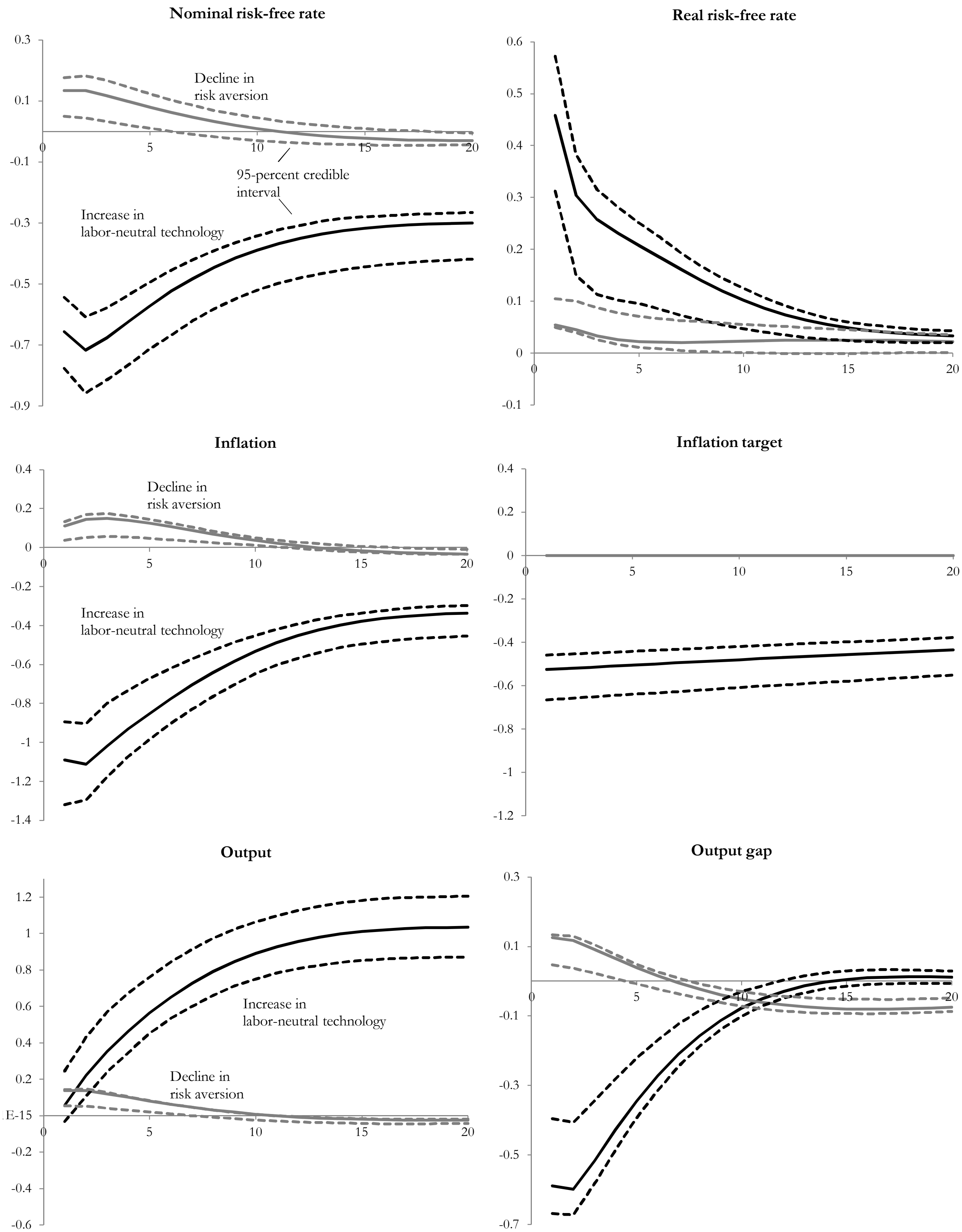

Note: responses of each of the term structure factors to the orthogonal structural shocks. Specifically, risk aversion and the inflation target are only affected by their own shocks, not the shocks to the other exogenous processes. The level factor is the average of the 1-quarter, 5-year, and 10-year yields. The slope factor is the gap between the 10-year and 1-quarter yields. Curvature is the sum of the 5-year and 1-year yields minus twice the 3-year yield. The shocks are all unit standard deviations. All scales in each row are identical and are measured in annualized percentage points.

# Figure 5. Expected returns and the term premium



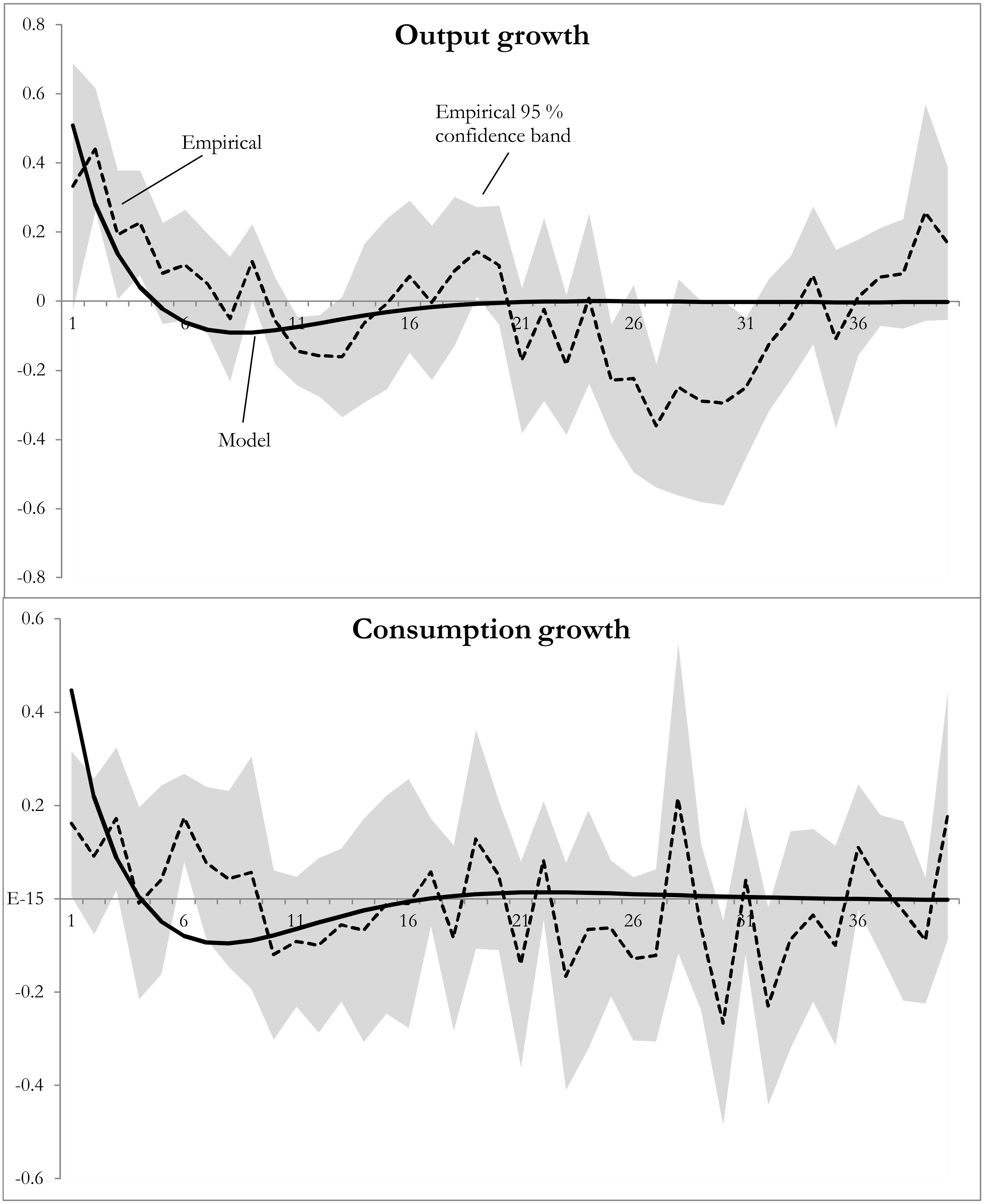

Note: Top panel gives expected excess returns on a 10-year bond over the following quarter, annualized. Values for the Cochrane–Piazzesi are from a linear regression. The term premium in the bottom panel is defined as the gap between the 10-year nominal yield and the mean of expected 1-quarter yields over the following 10 years.

# Figure 6. Responses to technology and risk-aversion shocks

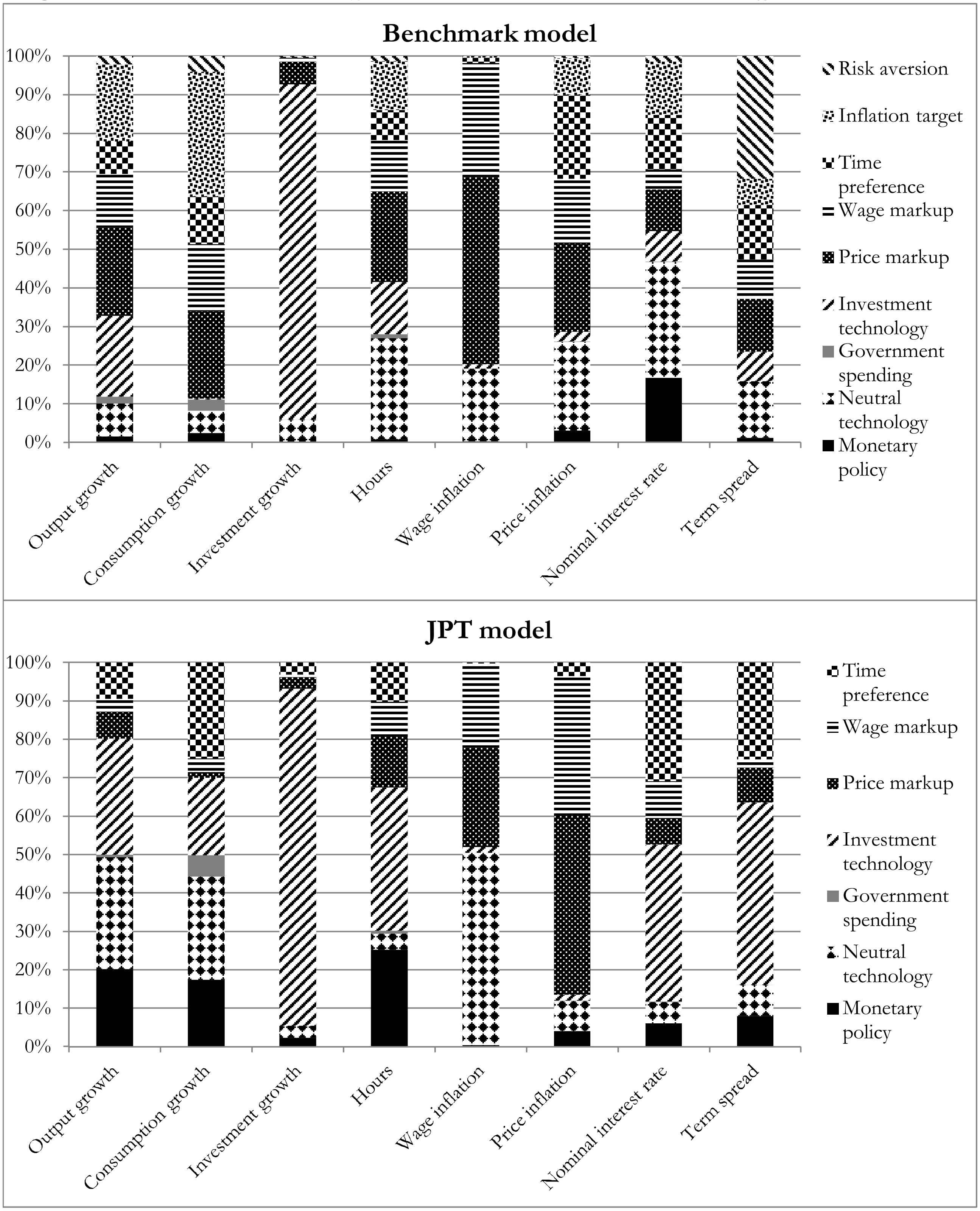



Note: responses in percentage points to a unit standard deviation positive shock to labor-neutral technology and a negative shock to risk aversion. Interest rate, inflation, and inflation target are annualized. Dotted lines give 2.5 and 97.5 percentiles from the posterior distribution

# Figure 7. Empirical and model-implied autocorrelations



## Figure 8. Variance decompositions at business-cycle frequencies




Note: each section of a bar represents a fraction of the variance of one of the series at frequencies between 6 and 32 quarters. The top panel gives results for the benchmark model. The bottom panel report results where the inflation target and risk aversion are fixed and bond prices are dropped from the estimation. That latter model is identical to JPT except that it uses post-1983 data and uses Epstein–Zin preferences.

# Figure 9. Estimated annualized inflation target and level factor


Note: Kalman-filtered estimates of the Federal Reserve's target for annualized inflation.