The New Keynesian Model and Bond Yields

Martin M. Andreasen*

November 27, 2024

Abstract

This paper presents a New Keynesian model to capture the linkages between macro fundamentals and the nominal yield curve. The model explains bond yields with a low level of news in expected inflation and plausible term premia. This implies that the slope of the yield curve predicts future bond yields, and that risk-adjusted historical bond yields satisfy the expectations hypothesis. The model also explains the spanning puzzle, matches key moments for real bond yields, captures the evolution of the price-dividend ratio, and implies that the slope of the yield curve and the price-dividend ratio forecast excess equity returns.

Keywords: Inflation variance ratios, Robust structural estimation, Term premia, The expectations hypothesis, Unspanned macro variation.

JEL: E44, G12.

*Aarhus University and the Danish Finance Institute. Fuglesangs Allé 4, 8210 Aarhus V, Denmark, email: mandreasen@econ.au.dk, telephone +45 87165982. I thank an anonymous referee, Kasper Jørgensen, Anders Bredahl Kock, Dennis Kristensen, Giovanni Pellegrino, Morten Ravn, and Daniel Wilhelm for useful comments and discussions. I also appriciate comments from participants at the Cemmap seminar at University College London. I finally acknowledge funding from the Independent Research Fund Denmark, project number 7024-00020B, and support from the Danish Finance Institute (DFI).

I Introduction

The New Keynesian model is one of the most prominent models in macroeconomics to understand the dynamics of economic activity, inflation, and the monetary policy rate. The model has also implications for prices on financial assets, although these prices often are ignored. The market for nominal government bonds is of particular interest, because bond yields contain information about expected future policy rates and term premia, where the latter is determined by consumption and inflation risk. But, it is well-known that the model struggles to explain bond yields, mainly because term premia are too low and too stable (see, e.g., Rudebusch and Swanson (2008)). This is obviously a concern, because prices are the key device that ensure equilibrium in the model, as emphasized by Cochrane (2007).

Much work has therefore been devoted to improve the New Keynesian model along this dimension.[1] The inclusion of recursive preferences by Rudebusch and Swanson (2012) is crucial in this context, because these preferences allow an otherwise standard New Keynesian model to generate plausible term premia without distorting macro fundamentals. In this model news about expected inflation is a key driver of bond yields. However, Duffee (2018) shows that this implication is not consistent with data for the United States (U.S.), and this finding therefore challenges the current specification of bond yields and inflation dynamics in the New Keynesian model.

This paper modifies the New Keynesian model to explain macro fundamentals *and* bond yields with a low level of news in expected inflation. The proposed model generates realistic term premia that have the same level and variability as in reduced-form dynamic term structure models (DTSMs). In addition, the model satisfies the two requirements in Dai and Singleton (2002) for a correct specification of term premia, which serve as useful diagnostic tests given that term premia are unobserved. The first requirement is that the difference between long- and short-term bond yields (i.e., the yield spread) predicts future bond yields as in Campbell and Shiller (1991). This property of bond yields is evaluated using a long simulated sample and is

[1] See for instance Hordahl, Tristani, and Vestin (2008), Graeve, Emiris, and Wouters (2009), Bekaert, Cho, and Moreno (2010), Dew-Becker (2014), Kung (2015), Bretscher, Hsu, and Tamoni (2020), and Meyer-Gohde and Kliem (2022) among others.

therefore an unconditional test of term premia. The second requirement is that historical bond yields satisfy the expectations hypothesis once adjusted for term premia using the proposed model. This second requirement is therefore a conditional test of term premia, in the sense that it requires reliable state and term premia estimates for historical bond yields. We also show that the nonlinear structure of the model generates variation in key macro variables that is not explained in linear regressions of these macro variables on bond yields. That is, the model *endogenously* produces unspanned variation in macro variables as observed in U.S. data, and it therefore goes a long way in avoiding the critique of macro-finance term structure models raised in Joslin, Priebsch, and Singleton (2014). The model also provides a close fit to real bond yields, although they are not included when estimating the model. Taking the analysis beyond bond yields, we finally show that the model captures the overall evolution in the price-dividend ratio, and that excess equity returns are predictable based on the yield spread and the price-dividend ratio, as documented by Fama and French (1989). Here, and throughout we restrict relative risk aversion to 10, meaning that we do not rely on extreme levels of risk aversion to explain asset prices from macro fundamentals.

Two features of our model are essential to obtain these results. First, we use the flexible formulation of recursive preferences in Andreasen and Jørgensen (2020) to separate the intertemporal elasticity of substitution (IES), the relative risk aversion (RRA), *and* the timing attitude. In contrast, the standard formulation of recursive preferences only separates the IES and the RRA, while the timing attitude follows directly from the IES and RRA. With this flexible formulation of recursive preferences, demand shocks is found to be a key driver of bond yields, their embedded term premia, and key macro variables, but not of inflation. This is different from the standard formulation of recursive preferences, where bond yields and term premia are mainly determined by productivity shocks, as shown in Rudebusch and Swanson (2012). However, productivity shocks have a sizable effect on inflation and hence explain bond yield dynamics with too much variation in expected inflation, as emphasized by Duffee (2018).

Second, the New Keynesian model is typically solved using a third-order perturbation approximation, because it allows for time-varying term premia unlike the standard log-linear approximation widely used in macroeconomics. The perturbation approximation and its risk-adjustment for the variables of interest $\mathbf{y}_t$ are obtained at the deterministic steady state $\mathbf{y}_{ss}$, but the approximation is applied around the unconditional mean $\mathbb{E}[\mathbf{y}_t]$. This extrapolation typically works well when $\mathbf{y}_{ss}$ is reasonably close to $\mathbb{E}[\mathbf{y}_t]$. However, we generally find that this difference is large with recursive preferences estimated to match U.S. bond yields, and that this leads to poor performance of the perturbation approximation, even when considering expansions up to fifth order. We address this previously unnoticed shortcoming of the perturbation approximation by instead using the projection method of Judd (1992), which is computed on a grid around $\mathbb{E}[\mathbf{y}_t]$. To greatly improve its computational efficiency, we note that for a given stochastic discount factor the solution to all bond prices is given in closed form by ordinary least squares (OLS) and hence obtained instantaneously. The same holds for expected future short rates and hence for term premia. These closed form solutions represent a new methodological contribution to the projection literature and apply to many popular choices of basis functions and to any approximation order. We show that the estimates for several parameters in the New Keynesian model change substantially when using the projection approximation instead of the less accurate but widely used third-order perturbation approximation.

The remainder of this paper is organized as follows. Section II presents the New Keynesian model, while our estimation approach is described in Section III. The main empirical results are provided in Section IV, with various sensitivity results provided in Section V. Additional model implications are discussed in Section VI and Section VII concludes.

II A New Keynesian Model

A The Households

We consider an infinitely lived representative household with recursive preferences as in Epstein and Zin (1989) and Weil (1990). Using the formulation in Rudebusch and Swanson

(2012), the value function V_t is given by

$$V_t = u_t + \beta \left(\mathbb{E}_t[V_{t+1}^{1-\alpha}]\right)^{1/(1-\alpha)} \tag{1}$$

when the utility function $u_t > 0$ for all t.[2] The subjective discount factor $\beta \in (0, 1)$ and $\mathbb{E}_t[\cdot]$ denotes the conditional expectation in period t. The main purpose of $\alpha \in \mathbb{R} \setminus \{1\}$ is to endow the household with preferences for when uncertainty is resolved, unless $\alpha = 0$ and equation (1) reduces to expected utility. It follows from Kreps and Porteus (1978) that equation (1) implies preferences for early (late) resolution of uncertainty if $\alpha > 0$ ($\alpha < 0$) for $u_t > 0$, whereas the opposite sign restrictions apply when $u_t < 0$. Andreasen and Jørgensen (2020) further argue that the size of this timing attitude is proportional to α, meaning that numerically larger values of α generate stronger preferences for early (late) resolution of uncertainty.

The utility function $u_t \equiv u(c_t, l_t)$ is assumed to depend on the number of consumption units c_t bought in the goods market and the provided labor supply l_t to firms. Following Andreasen and Jørgensen (2020), we also include a constant u_0 in the utility function to account for utility from goods and services that are not acquired in the goods market. This could be utility from government spending or utility from goods produced and consumed within the household. As shown in Andreasen and Jørgensen (2020), the reason for introducing u_0 is to separately control the level of the utility function to disentangle the timing attitude α from relative risk aversion (RRA), which otherwise are tightly linked in the standard formulation of recursive preferences. Using a power specification to quantify the utility from market consumption c_t and similarly for leisure $1 - l_t$, we let

$$u(c_t, l_t) = \frac{d_t}{1-\chi}(c_t - bc_{t-1})^{1-\chi} + z_t^{1-\chi}\left(d_t n_t \varphi_0 \frac{(1-l_t)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}} + u_0\right). \tag{2}$$

The parameter $b \geq 0$ accommodates external consumption habits, which are included to capture autocorrelation in consumption growth. The variable d_t introduces shocks to the utility function to temporally increase or decrease the utility from a given level of c_t and l_t, implying that d_t

[2] When $u_t < 0$, $V_t = u_t - \beta\mathbb{E}_t[(-V_{t+1})^{1-\alpha}]^{\frac{1}{1-\alpha}}$ as in Rudebusch and Swanson (2012).

operates as a demand shock.[3] The variable n_t is also exogenous and temporally shifts the household's incentive between c_t and l_t, meaning that n_t may be interpreted as a labor supply shock. Market consumption grows with the rate of the productivity level z_t, and it is therefore necessary to scale the utility from leisure and nonmarket consumption by $z_t^{1-\chi}$ to ensure that these terms do not diminish relative to $\frac{1}{1-\chi}(c_t - bc_{t-1})^{1-\chi}$ along the balanced growth path. This scaling is discussed and motivated further in Rudebusch and Swanson (2012) and Andreasen and Jørgensen (2020).

The parameter $\chi > 0$ and the degree of habit formation b determine the steady state intertemporal elasticity of substitution (IES) as $\left(1 - b\mu_{z,ss}^{-1}\right)/\chi$, which measures the percentage change in market consumption growth for a one percent change in the real interest rate when ignoring uncertainty. The endogenous labor supply gives the household an additional margin to absorb shocks and this modifies existing expressions for RRA. Using the results in Swanson (2018), it follows that RRA in the steady state is

$$\begin{aligned} \text{RRA} &= \frac{\chi}{\left(1 - \frac{b}{\mu_{z,ss}}\right) + \chi\varphi\frac{\tilde{w}_{ss}(1-l_{ss})}{\tilde{c}_{ss}}} \\ &+ \frac{\alpha(1-\chi)}{(1-\chi)u_0\tilde{c}_{ss}^{\chi-1}\left(1 - \frac{b}{\mu_{z,ss}}\right)^{\chi} + \left(1 - \frac{b}{\mu_{z,ss}}\right) + \frac{1-\chi}{1-\frac{1}{\varphi}}\frac{\tilde{w}_{ss}^{*}(1-l_{ss})}{\tilde{c}_{ss}}} \end{aligned}$$

where $\tilde{c}_{ss} = c_t/z_t|_{ss}$ and $\tilde{w}_{ss}^* = w_t^*/z_t|_{ss}$ refer to the steady state (ss) of market consumption and the real frictionless wage w_t^* relative to the productivity level. Given that the IES is determined by χ, and α controls the strength of the timing attitude, the level of the utility function u_0 is the key parameter for determining RRA.

The real budget constraint is $c_t + \mathbb{E}_t\left[M_{t,t+1}x_{t+1}^{real}\right] = x_t^{real}/\pi_t + w_t^* l_t + Div_t + T_t$. That is, resources are spent on market consumption goods c_t and nominal state-contingent claims X_{t+1}, where $x_{t+1}^{real} \equiv X_{t+1}/P_t$ denotes the real value of these claims that are priced using the nominal stochastic discount factor $M_{t,t+1}$. The household's income is given by the real value of

[3] In the endowment models of Albuquerque, Eichenbaum, Luo, and Rebelo (2016) and Gomez-Cram and Yaron (2021), shocks to the utility function affect only asset prices and are therefore referred to as capturing valuation risk. In production-based equilibrium models, shocks to d_t generate also endogenous variation in consumption, inflation, etc., and this explains why these shocks are referred to as demand shocks within these models.

state-contingent claims bought in the previous period x_t^{real}/π_t, the real wage income $w_t^* l_t$, real dividend payments from firms Div_t, and real lump-sum transfers T_t. Here, π_t denotes the gross inflation rate.

The first-order conditions for the utility maximizing household are

$$\mathbb{E}_t \left[M_{t,t+1} R_t \right] = 1 \tag{3}$$

$$z_t^{1-\chi} n_t \varphi_0 \left(1 - l_t\right)^{-\frac{1}{\varphi}} = \left(c_t - bc_{t-1}\right)^{-\chi} w_t^*, \tag{4}$$

where R_t is gross one-period nominal interest rate. Equation (3) is the well-known consumption Euler-equation, while equation (4) is the optimality condition for the labor supply.

To incorporate wage stickiness, we follow Blanchard and Gali (2005), Rudebusch and Swanson (2008), among others and introduce a simple wage bargaining friction in the labor market. Specifically, we assume that the real market wage w_t is given by

$$w_t = \kappa_w \left(\widetilde{w}_{ss} z_t\right) + \left(1 - \kappa_w\right) w_t^*, \tag{5}$$

where $\kappa_w \in [0, 1)$ captures wage stickiness by smoothing the frictionless wage w_t^* in relation to the long-term wage along the balanced growth path $\widetilde{w}_{ss} z_t$. Although this simple wage rule does not explicitly introduce Nash bargaining between workers and firms, Blanchard and Gali (2005) argue that it is a parsimonious way to capture the essential features of real wage bargaining.

B The Firms

Output y_t is produced by a competitive representative firm, which combines differentiated intermediate goods $y_t(i)$ using $y_t = \left(\int_0^1 y_t(i)^{\frac{\eta-1}{\eta}} di\right)^{\frac{\eta}{\eta-1}}$ with $\eta > 1$. The demand for the ith good is given by $y_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\eta} y_t$, where $P_t \equiv \left(\int_0^1 P_t(i)^{1-\eta} di\right)^{\frac{1}{1-\eta}}$ denotes the aggregate price level and $P_t(i)$ is the price of the ith good.

Intermediate firms produce slightly differentiated goods using the production function

$y_t(i) = z_t a_t k_{ss}^{\theta} l_t(i)^{1-\theta}$, where k_{ss} and $l_t(i)$ denote capital and labor services at the ith firm, respectively. The variables a_t and z_t capture transitory and permanent productivity shocks, respectively. Each intermediate firm can freely adjust its labor demand at the given market wage w_t. Price stickiness is introduced as in Rotemberg (1982), where $\xi \geq 0$ controls the size of firms' real cost $\frac{\xi}{2}\left(P_t(i) / \left(P_{t-1}(i)\pi_{ss}\right) - 1\right)^2 y_t$ when changing $P_t(i)$. As in Rudebusch and Swanson (2012), each firm uses $\delta k_{ss} z_t$ units of output for investment to maintain a constant capital stock along the balanced growth path. The first-order conditions for profit maximization are

$$w_t = mc_t (1-\theta) z_t a_t k_{ss}^{\theta} l_t^{-\theta} \tag{6}$$

$$y_t(1-\eta) + \mathbb{E}_t\left[\xi M_{t,t+1}\pi_{t+1}\left(\frac{\pi_{t+1}}{\pi_{ss}} - 1\right)\frac{\pi_{t+1}}{\pi_{ss}} y_{t+1}\right] + mc_t y_t \eta = \xi\left(\frac{\pi_t}{\pi_{ss}} - 1\right) y_t \frac{\pi_t}{\pi_{ss}}, \tag{7}$$

where mc_t denotes marginal costs. Equation (6) determines labor demand, and equation (7) implies an aggregate supply relation between output and inflation.

The aggregate resource constraint reads $c_t + z_t \delta k_{ss} = \left(1 - \frac{\xi}{2}\left(\frac{\pi_t}{\pi_{ss}} - 1\right)^2\right) y_t$, where $\frac{\xi}{2}\left(\frac{\pi_t}{\pi_{ss}} - 1\right)^2 y_t$ is the output loss from price stickiness.[4]

C The Central Bank

The policy rate for the Federal Reserve is the short (one-period) nominal rate R_t. It is set according to the forward-looking Taylor-rule

$$R_t = R_{ss}\mathbb{E}_t\left[\exp\left\{\phi_{\pi}\log\left(\frac{\pi_{t+1}}{\pi_{ss}\pi_{t+1}^{*}}\right) + \phi_{\Delta c}\left(\Delta c_{t+1} - \Delta c_{ss}\right)\right\}\right], \tag{8}$$

based on a desire to stabilize future inflation π_{t+1} and future economic activity as measured by consumption growth Δc_{t+1}. To accommodate the large variation in inflation in post-war U.S. data, we follow the existing literature and introduce an exogenous inflation target π_t^* around the steady state inflation rate π_{ss}.

[4] Here we abstract from any costs linked to wage stickiness by setting the household transfers $T_t = l_t w_t - l_t w_t^* = \kappa_w l_t\left(\left(\widetilde{w}_{ss} z_t\right) - w_t^*\right)$ to ensure that the wage bill $w_t l_t$ paid by the firms is also the wage bill received by the household.

D Bond Pricing and Structural Shocks

The price in period t of a default-free zero-coupon bond $B_t^{(k)}$ maturing in k periods with a face value of one dollar is $B_t^{(k)} = \mathbb{E}_t\left[M_{t,t+1}B_{t+1}^{(k-1)}\right]$ for $k = 1, ..., n_b$ with $B_t^{(0)} = 1$. Using continuos compounding, the yield to maturity is $r_t^{(k)} = -\frac{1}{k}b_t^{(k)}$ with $b_t^{(k)} \equiv \log B_t^{(k)}$. Bond prices are therefore determined by the nominal stochastic discount factor, which reads

$$M_{t,t+1} = \beta \frac{d_{t+1}}{d_t}\left(\frac{c_{t+1} - bc_t}{c_t - bc_{t-1}}\right)^{-\chi}\left(\frac{\left(\mathbb{E}_t\left[V_{t+1}^{1-\alpha}\right]\right)^{\frac{1}{1-\alpha}}}{V_{t+1}}\right)^{\alpha}\frac{1}{\pi_{t+1}} \tag{9}$$

when $u_t > 0$ for all t. The term premium $TP_t^{(k)}$ is defined as the difference between $r_t^{(k)}$ and the expected future path of the short rate, i.e., $TP_t^{(k)} = r_t^{(k)} - \sum_{i=0}^{k-1}\mathbb{E}_t\left[r_{t+i}\right]$ with $r_t \equiv \log R_t$.

Finally, we use a standard specification for the five structural shocks by letting

$$\begin{aligned}
\log\log\left(\mu_{z,t+1}/\mu_{z,ss}\right) &= \rho_{\mu_z}\log\left(\mu_{z,t}/\mu_{z,ss}\right) + \sigma_{\mu_z}\epsilon_{\mu_z,t+1} \\
\log d_{t+1} &= \rho_d \log d_t + \sigma_d\epsilon_{d,t+1} \\
\log n_{t+1} &= \rho_n \log n_t + \sigma_n\epsilon_{n,t+1} \\
\log \pi^*_{t+1} &= \rho_{\pi^*}\log \pi^*_t + \sigma_{\pi^*}\epsilon_{\pi^*,t+1} \\
\log a_{t+1} &= \rho_a \log a_t + \sigma_a\epsilon_{a,t+1}
\end{aligned} \tag{10}$$

with $\mu_{z,t} \equiv z_t/z_{t-1}$, where $\epsilon_{\mu_z,t+1}$, $\epsilon_{d,t+1}$, $\epsilon_{n,t+1}$, $\epsilon_{\pi^*,t+1}$, and $\epsilon_{a,t+1}$ are standard normally distributed and independent across time, denoted $\mathcal{NID}(0,1)$, and mutually uncorrelated.

III Estimation Methodology

This section outlines how we estimate the New Keynesian model. We describe the considered data in Section III.A and present our solution method in Section III.B. The adopted estimation routine is discussed in Section III.C.

A Data

The model is estimated using quarterly U.S. data from 1961Q2 to 2019Q3, where we end the sample before the start of the COVID-19 pandemic. The dynamics of the macro economy is measured by i) detrended labor supply $\hat{l}_t$, ii) the detrended real wage $\hat{w}_t$ iii) consumption growth Δc_t, and iv) inflation π_t.[5] The nominal yield curve is represented by the 3-month, 1-, 3-, 5-, 7-, and 10-year bond yields. These yields are taken from Gürkaynak, Sack, and Wright (2007), except at the 3-month maturity where we use the implied rate on a 3-month Treasury Bill.

Taylor-rules as the one considered in equation (8) are obviously an approximation to the historical policy decisions taken by the Federal Reserve. In particular, the rule ignores the 'monetarist experiment' from 1979Q4 to 1982Q4 and the policy rate being constrained by the zero lower bound (ZLB) after 2008. We account for both of these deviations from the adopted interest rule in the following way.

First, during the monetarist experiment, the Federal Reserve targeted monetary aggregates and its policy decisions were therefore not well summarized by the interest rule in equation (8). To avoid that the highly volatile short rate in this period distorts the estimation, we greatly downweight the importance of the 3-month interest rate from 1979Q4 to 1982Q4 in the estimation. As explained below in Section III.C, the model is estimated using nonlinear filtering techniques, and we implement this downweighting by increasing the standard deviation of the measurement errors for the 3-month interest rate to 100 basis points from 1979Q4 to 1982Q4.

Second, the interest rule in equation (8) does not enforce the ZLB. A possible concern is therefore that the model may generate a short rate during the estimation that is clearly below zero although the empirical short rate is always positive. We guard against this possibility by lowering the standard deviation of the measurement in the 3-month rate to just 5 basis points at the ZLB. This also ensures that the signal-to-noise ratio for the 3-month rate remains reasonable

[5] The labor supply is measured by the total number of employees for total nonfarm payrolls, while the real wage is measured by the seasonally adjusted real hourly compensation in the nonfarm business sector. Both variables are detrended using the procedure in Hamilton (2018). The consumption growth rate is calculated from real per capita nondurables and service expenditures. Inflation is measured by the year-on-year growth rate in the consumer price index (CPI) excluding food and energy prices for all urban consumers.

at the ZLB, where the variability in the short rate is greatly reduced. We define the ZLB as spells where the 3-month interest rate is below 50 basis points.

In total, the estimation uses 10 variables, which we collect in the vector $\mathbf{y}_t^{obs}$ and express in annualized terms except for $\hat{l}_t$ and $\hat{w}_t$.

B Model Solution

The state variables for the model appear in the vector $\mathbf{x}_t = [\begin{array}{cccccc} \log\left(\frac{c_{t-1}}{z_{t-1}}\right) & \log\mu_{z,t} & \log d_t & \log n_t & \log\pi_t^* & \log a_t \end{array}]'$, where consumption is scaled by the productivity level z_t to ensure stationarity. All the remaining variables appear in the vector $\mathbf{y}_t$ with dimensions $n_y \times 1$. The exact solution is

$$\begin{aligned} \mathbf{y}_t &= \mathbf{g}(\mathbf{x}_t; \boldsymbol{\theta}) \\ \mathbf{x}_{t+1} &= \mathbf{h}(\mathbf{x}_t; \boldsymbol{\theta}) + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} \end{aligned} \tag{11}$$

with $\boldsymbol{\epsilon}_t = [\begin{array}{ccccc} \epsilon_{\mu_z,t} & \epsilon_{d,t} & \epsilon_{n,t} & \epsilon_{\pi^*,t} & \epsilon_{a,t} \end{array}]'$ and all structural coefficients collected in $\boldsymbol{\theta}$. The law of motions for the exogenous states are provided in equation (10), and the dynamics of $\log\left(\frac{c_{t-1}}{z_{t-1}}\right)$ is identical to the one for consumption scaled by z_t in $\mathbf{g}(\mathbf{x}_t; \boldsymbol{\theta})$. As a result, we only need to approximate the $\mathbf{g}$-function.

The solution to the New Keynesian model with asset prices is typically obtained using a third-order perturbation approximation, because it allows for non-zero and time-varying risk premia. Although this approximation is often found to be very accurate as shown in Aruoba, Fernandez-Villaverde, and Rubio-Ramirez (2006) and Caldara, Fernandez-Villaverde, Rubio-Ramirez, and Yao (2012), the inclusion of recursive preferences generates a large risk-adjustment that may reduce the accuracy of the approximation. The reason is that the risk-adjustment to capture term premia in the perturbation approximation is computed around the deterministic steady state of $\mathbf{y}_t$, denoted $\mathbf{y}_{ss}$, but applied around the unconditional mean of $\mathbf{y}_t$, denoted $\mathbb{E}[\mathbf{y}_t]$. This extrapolation typically works well when $\mathbf{y}_{ss}$ is reasonably close to $\mathbb{E}[\mathbf{y}_t]$. However, we generally find that this difference is large with recursive preferences estimated to match U.S. bond yields, and this leads to poor performance of the perturbation approximation.

For instance, in our baseline specification of the New Keynesian model, the annualized short rate has a determinist steady state of 18.50%, but an unconditional mean of just 4.80%.[6]

Based on this observation, we instead solve the model by the projection method of Judd (1992) using a second-order polynomial for the **g**-function. Feasibility obviously dictates this relative low approximation order with $n_x = 6$ states, $n_y = 45$ control variables, and $n_\varepsilon = 5$ structural shocks in the model. Nevertheless, as we show below in Section V.B, the global nature of this solution provides a more accurate approximation when compared to even a fifth order perturbation solution - in part because the risk correction in the projection approximation is computed on a grid around $\mathbb{E}[\mathbf{y}_t]$ and not around $\mathbf{y}_{ss}$. That is, we let

$$\mathbf{y}_t = \mathbf{g}_0 + \mathbf{g}_\mathbf{x}\mathbf{x}_t + \mathbf{g}_{\mathbf{xx}} vech\left(\mathbf{x}_t\mathbf{x}_t'\right), \tag{12}$$

where the unknown loadings for this projection approximation appear in $\mathbf{g}_0$, $\mathbf{g}_\mathbf{x}$, and $\mathbf{g}_{\mathbf{xx}}$ with dimensions $n_y \times 1$, $n_y \times n_x$, and $n_y \times n_x(n_x+1)/2$, respectively. We consider N grid points $\mathcal{X} = \{\mathbf{x}_i\}_{i=1}^N$ for the states and determine these unknown loadings by minimizing the model's Euler-equation errors on this grid. Details on the construction of the grid $\mathcal{X}$, evaluation of the conditional expectations in the Euler-equations, and the construction of the objective function are standard and provided in Appendix A.

For an efficient implementation, consider the partition $\mathbf{y}_t = \begin{bmatrix} (\mathbf{y}_t^a)' & (\mathbf{y}_t^b)' & (\mathbf{b}_t^{zc})' \end{bmatrix}'$, where $\mathbf{y}_t^a = [\ \log\left(\frac{c_t}{z_t}\right) \quad \log \pi_t \quad \log evf_t\ ]'$ contains the three macro variables required to solve the model without any bond yields, with $evf_t = \mathbb{E}_t\left[\left(V_t/z_t^{1-\chi}\right)^{1-\alpha}\right]$ being an auxiliary variable used to capture the recursive preferences. The vector $\mathbf{y}_t^b = [\ \log l_t \quad \log\left(\frac{w_t}{z_t}\right)\ ]'$ represents additional macro variables used in the estimation, while all log-transformed zero-coupon bond prices appear in $\mathbf{b}_t^{zc} = [\ b_t^{(1)} \quad b_t^{(2)} \quad \ldots \quad b_t^{(40)}\ ]'$. As shown in Appendix A, only the optimization step with respect to the unknown loadings related to $\mathbf{y}_t^a$ is nonlinear. Given the solution for $\mathbf{y}_t^a$, the projection loadings for $\mathbf{y}_t^b$ and $\mathbf{b}_t^{zc}$ are given in closed form by ordinary least squares (OLS) and hence obtained instantaneously. We also show that the same holds for expected future short

[6] The main problem appears to be that the perturbation method struggles to accurately approximate V_t and $\mathbb{E}_t\left[V_{t+1}^{1-\alpha}\right]$, which then spills over into large approximations errors for bond yields.

rates and hence for term premia. These closed form solutions for bond prices $\mathbf{b}_t^{zc}$ and expected short rates represent a new methodological contribution to the projection literature for models with a yield curve and extends to other basis functions than the one applied in equation (12) and to any approximation order. As a result, we are able to obtain a second-order projection approximation in just about 6 seconds on a standard desktop computer using a single CPU.[7] Exploiting this new computational trick is absolutely essential for making a formal estimation of our medium-sized New Keynesian model computational feasible with a global solution routine.

C Robustified Inference with Nonlinear Filtering

The approximated state space representation of the model includes nonlinear terms with the unobserved states $\mathbf{x}_t$, implying that the Kalman filter cannot be used to estimate the model. Instead, we rely on the central difference Kalman filter (CDKF) developed by Norgaard, Poulsen, and Ravn (2000), which is a nonlinear extension of the Kalman filter. The CDKF accommodates measurement errors in $\mathbf{y}_t^{obs}$, which in practice also include potential modelling errors. We specify these errors $\mathbf{v}_t$ to be uncorrelated Gaussian white noise, as typically assumed when estimating structural macroeconomic models and reduced-form DTSMs. That is, $\mathbf{v}_t \sim \mathcal{NID}(\mathbf{0}, \mathbf{R_v})$ where $\mathbf{R_v}$ is a diagonal covariance matrix, implying that $\mathbf{y}_t^{obs} = \mathbf{S}\mathbf{y}_t + \mathbf{v}_t$ for an appropriate selection matrix $\mathbf{S}$.[8]

A likelihood function can be derived from the CDKF under the assumption that the prediction errors for $\mathbf{y}_t^{obs}$ are Gaussian. However, this distributional specification does not hold exactly due to the nonlinear terms in the model solution. The CDKF therefore only provides a quasi log-likelihood function $\frac{1}{T}\sum_{t=1}^{T}\mathcal{L}_t(\boldsymbol{\theta})$, which can be used for a quasi maximum likelihood (QML) estimation, as suggested in Andreasen (2013).

A possible limitation of any QML approach (as with standard maximum likelihood) is its

[7] The processor is an AMD Ryzen 7 PRO 7840U. We reserve the use of multi-processing for the estimation of the structural parameters using the evolutionary optimization routine of Andreasen (2010) applied to the objective function described below in Section III.C.

[8] Unlike the more accurate particle filter proposed in Fernández-Villaverde and Rubio-Ramírez (2007), the updating rule for the states in the CDKF is linear, implying that the recursive filtering equations only depend on first and second moments. The CDKF approximates these moments by a deterministic sampling procedure, and this makes the CDKF computationally much faster than any particle filter and generally also more accurate than the well-known extended Kalman filter.

lack of robustness to model misspecifications other than the distributional assumption of the prediction errors for $\mathbf{y}_t^{obs}$. A possible source of misspecification in the New Keynesian model is that the same structural parameters in $\boldsymbol{\theta}$ determine the $\mathbf{g}$- and the $\mathbf{h}$-function. In contrast, the widely used reduced-form Gaussian DTSM basically allows these two functions to be determined by separate parameters, which in these models is essential to generate term premia dynamics that are not rejected by historical data (see Dai and Singleton (2002)).[9] In the context of the New Keynesian model, the $\mathbf{g}$-function is likely to prefer large and persistent shocks, because this helps the model to generate variation in term premia and hence fit medium- and long-term bond yields. On the other hand, small and less persistent shocks are typically needed in the $\mathbf{h}$-function to fit the state dynamics and ensure sensible unconditional properties of $\mathbf{y}_t^{obs}$. Given the inclusion of several bond yields with small measurement errors in the estimation, our experience is that the $\mathbf{g}$-function dominates the QML estimates of the New Keynesian model. This typically implies a tight in-sample fit to $\mathbf{y}_t^{obs}$, but it comes at the expense of unconditional variances for $\mathbf{y}_t^{obs}$ that are too large compared to U.S. data.

Thus, a robustified version of the standard QML estimator is required to obtain reliable estimates of our New Keynesian model. The solution we propose is to shrink the QML estimator to the unconditional first and second moments of $\mathbf{y}_t^{obs}$, because this is a simple and transparent way to increase the weight assigned to the $\mathbf{h}$-function in the estimation. We denote the shrinkage moments by $\frac{1}{T}\sum_{t=1}^{T}\mathbf{m}_t$ in the sample and by $\mathbb{E}\left[\mathbf{m}\left(\boldsymbol{\theta}\right)\right]$ in the model. The applied estimator is given by

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \ \frac{1}{T}\sum_{t=1}^{T}\mathcal{L}_t\left(\boldsymbol{\theta}\right) - \lambda\mathbf{f}_{1:T}\left(\boldsymbol{\theta}\right)'\mathbf{W}\mathbf{f}_{1:T}\left(\boldsymbol{\theta}\right), \tag{13}$$

where $\boldsymbol{\Theta}$ is the feasible domain of $\boldsymbol{\theta}$, $\mathbf{f}_{1:T}\left(\boldsymbol{\theta}\right) \equiv \frac{1}{T}\sum_{t=1}^{T}\mathbf{f}_t\left(\boldsymbol{\theta}\right)$ with $\mathbf{f}_t\left(\boldsymbol{\theta}\right) \equiv \mathbf{m}_t - \mathbb{E}\left[\mathbf{m}\left(\boldsymbol{\theta}\right)\right]$, and $\mathbf{W}$ is a diagonal weighting matrix containing the inverse of the standard errors for the shrinkage moments.[10]

[9] The only link between the $\mathbf{g}$- and $\mathbf{h}$-functions in the reduced-form Gaussian DTSM is the conditional covariance matrix of the states, but the effect of this covariance matrix on the $\mathbf{g}$-function is typically very small.

[10] Another possibility is to use the optimal weigthing matrix, but this version of equation (13) is not considered to avoid well-known small-sample distortions from estimating large covariance matrices in moment-based estimators. In our application, we estimate the diagonal elements in $\mathbf{W}$ using the Newey-West estimator with three lags.

The nature of the estimator in equation (13) is determined by $\lambda \geq 0$, which controls the weight assigned to the shrinkage moments relative to the sample average of $\mathcal{L}_t(\boldsymbol{\theta})$. We obviously recover the standard QML estimator when $\lambda = 0$, while $\hat{\boldsymbol{\theta}}$ converges to the generalized method of moments (GMM) estimator of Hansen (1982) when λ becomes sufficiently large. We consider a small amount of shrinkage by letting $\lambda = T$, which in our setting implies that shrinkage constitutes a small part of the objective function. Our results are not particularly sensitive to increasing the degree of shrinkage further, although we do find notable effects of shrinkage when compared to the standard QML estimator.[11]

To obtain closed-form expressions for the model-implied shrinkage moments, we apply the pruning scheme of Andreasen, Fernandez-Villaverde, and Rubio-Ramirez (2018) when setting up the state space system for the approximated version of the model.[12] This implies that the estimator in equation (13) can be implemented without resorting to simulation and belongs to the general class of extremum estimators. Its asymptotic properties are therefore easily derived in the Online Appendix.

We use two classes of shrinkage moments. The first class contains first and second unconditional moments of $\mathbf{y}_t^{obs}$, except for the mean of detrended labor supply and the real wage that are zero. These moments help to robustify the QML estimates, which we illustrate in a Monte Carlo study in the Online Appendix. Here we show that shrinkage towards these moments give smaller parameter biases than standard QML when the model is misspecified with the **g**- and **h**-functions determined by different structural parameters. On the other hand, without any model misspecification, we unexpectedly find no benefit of shrinkage, which mainly reduces the efficiency of the standard QML estimator that is nearly unbiased in this case. We benchmark these results to using an extreme degree of shrinkage with $\lambda = 10^6$, which corresponds to estimating the structural parameters by GMM and obtaining the states

[11] Note also that equation (13) belongs to the class of Laplace type or quasi-Bayesian estimators of Chernozhukov and Hong (2003), where a potentially misspecified log-likelihood function (as considered in our case) may be used within a Bayesian setting. When this estimator is combined with the endogenours prior specification in Christiano, Trabandt, and Walentin (2011) using a pre-sample of length T^*, we obtain the objective function in equation (13) with $\lambda = T^*/2$, as shown in the Online Appendix.

[12] The pruning scheme is not essential in our case, because the proposed model has only one endogenous state (i.e., $\log(c_{t-1}/z_{t-1})$), and partly for this reason appears to be stable when simulated without pruning. Details related to the pruned state space system and the derivations of its moments are provided in the Online Appendix.

afterwards by the CDKF. The Monte Carlo study shows that these GMM estimates of $\boldsymbol{\theta}$ display notable biases in finite samples and are clearly less efficient compared to the standard QML estimator ($\lambda = 0$) and the proposed estimator (with $\lambda = T$), both with and without model misspecification. These imprecise GMM estimates of the structural parameters also imply less accurate state estimates when compared to the proposed estimator with $\lambda = T$.

The second class of moments we include contain information about term premia as captured by the regression of Campbell and Shiller (1991). That is,

$$r_{t+m}^{(k-m)} - r_t^{(k)} = \alpha_k + \beta_k \frac{m}{k-m}\left(r_t^{(k)} - r_t^{(m)}\right) + u_{t+m,k}, \tag{14}$$

where $u_{t+m,k}$ is an error term and $m = 4$ to obtain annual changes in bond yields for our quarterly sample. The expectations hypothesis implies $\beta_k = 1$, but empirical estimates of the slope coefficients β_k are negative and decreasing with maturity, which is evidence of time-varying term premia. This implies that bond returns are predictable, because equation (14) is equivalent to the forecast regression $rx_{t+m}^{(k)} = \tilde{\alpha}_k + \tilde{\beta}_k \left(r_t^{(k)} - r_t^{(m)}\right) + \varepsilon_{t+m,k}$, where $rx_{t+m}^{(k)}$ is excess bond returns and $\tilde{\beta}_k \equiv \frac{m}{4}(1 - \beta_k)$. We consider β_k at the 2-, 3-, ..., 10-year maturity to assign more weight to the evidence against the expectations hypothesis than implied by the considered panel of bond yields used to compute the quasi log-likelihood function. These slope coefficients are represented in the shrinkage moments by including $Cov(r_{t+m}^{(k-m)} - r_t^{(k)}, r_t^{(k)} - r_t^{(m)})$ and $Var(r_t^{(k)} - r_t^{(m)})$ related to β_k at the included maturities.

IV Estimation Results

This section presents the main results for the proposed model. We proceed by discussing the estimated parameters in Section IV.A and the model fit in Sections IV.B to IV.D. The following three subsections explore how well the model matches key moments of bond yields that are *not* included in the estimation. The key mechanisms in the model are finally explained in Section IV.H.

A Estimated Parameters

Not all parameters in the New Keynesian model are well-identified from our set of observables, and these parameters are therefore determined by standard calibration arguments. Hence, we let $\delta = 0.025$, $l_{ss} = 0.33$, and $\theta = 0.4$ as typically assumed for the U.S. economy. We also set the ratio of capital to output in the steady state to 2.5 as in Rudebusch and Swanson (2012). The value of $\mu_{z,ss}$ is set to match the mean of consumption growth, implying that $\mu_{z,ss} = 1.0055$. Finally, we set the constant u_0 in the utility function to get a steady state RRA of 10 as considered in Bansal and Yaron (2004).

Given the large number of variables included in the estimation, we set the size of the measurement errors in $\mathbf{y}_t^{obs}$ using reasonable assumptions. The standard deviation for the measurement errors in all bond yields is set to 25 basis points, corresponding to 8.3% of the average standard deviations in the six considered yields. Similarly, we assume that 20% of the variation in labor supply, the real wage, and consumption growth is due to measurement errors, implying standard deviations of 58, 44, and 35 basis points, respectively. Inflation displays very similar dynamics to short-term bond yields, and we therefore set its measurement errors to 10% of its variation, as it gives a standard deviation of 24 basis points.

The first column in Table 1 reports the estimated parameters in the preferred version of the model. This baseline model is denoted $\mathcal{M}^{M,CS}$, where the superscripts indicate that shrinkage is applied based on the considered first and second unconditional moments of $\mathbf{y}_t^{obs}$ (by "M") and the selected Campbell-Shiller moments (by "CS"). The results for the other versions of the model will be discussed below in Section V. For our preferred model $\mathcal{M}^{M,CS}$ we find $\hat{\beta} = 0.991$ and a moderate degree of habit formation with $\hat{b} = 0.65$. For consumption growth, this implies autocorrelations of 0.65, 0.42, 0.27, and 0.17 at the first four lags, which are reasonably close to the responding sample moments of 0.47, 0.37, 0.39, and 0.18, respectively. For the curvature parameter in the utility function, we get $\hat{\chi} = 5.83$, which together with our estimate of b imply an IES of 0.06. This relatively low IES is consistent with Hall (1988) and Yogo (2004), who estimate the IES to between zero and 0.2.

The degree of substitutability among the intermediate goods is $\eta = 8.85$, giving a low and

Table 1: The Structural Parameters

This table shows the estimated parameters using quarterly data from 1961 to 2019, with asymptotic standard errors provided in parentheses. The first four observations are used to initialize the CDKF. The use of shrinkage is denoted on the model object by using the superscipts "M" for the unconditional moments and "CS" for the Campbell-Shiller moments. The subscript on the model object indicates that the standard specification of recursive preferences (EZ) is used in column 4 and that a third-order perturbation approximation (3th) to the model solution is used in column 5. No standard errors are provided for β, η, and $\phi_{\Delta c}$ when they are at their upper bounds of 0.999, 10, and 10, respectively. The timing premium is evaluated at the steady state and computed as in Andreasen and Jørgensen (2020).

	$\mathcal{M}^{M,CS}$	$\mathcal{M}^{M}$	$\mathcal{M}$	$\mathcal{M}^{M,CS}_{\text{EZ}}$	$\mathcal{M}^{M,CS}_{\text{3th}}$
	1	2	3	4	5
α	-11.460 (6.214)	-9.029 (1.664)	-9.032 (0.630)	-16.488 (3.788)	-14.250 (2.367)
β	0.991 (0.011)	0.999 –	0.988 (0.000)	0.993 (0.003)	0.948 (0.057)
b	0.653 (0.077)	0.574 (0.007)	0.841 (0.010)	0.866 (0.116)	0.625 (0.268)
φ	0.098 (0.065)	0.092 (0.003)	0.156 (0.003)	0.209 (0.045)	0.169 (0.347)
χ	5.828 (1.101)	7.104 (0.309)	5.042 (0.630)	1.948 (0.550)	15.633 (13.174)
ξ_{Calvo}	0.438 (0.086)	0.309 (0.003)	0.497 (0.026)	0.772 (0.008)	0.396 (0.230)
κ_w	0.917 (0.016)	0.909 (0.020)	0.834 (0.001)	0.867 (0.045)	0.945 (0.017)
η	8.848 (11.414)	10.000 –	10.000 –	8.208 (2.166)	9.774 (0.763)
ϕ_π	3.957 (2.321)	3.691 (0.880)	4.764 (1.558)	2.982 (0.830)	10.000 –
$\phi_{\Delta c}$	1.096 (0.442)	0.658 (0.563)	4.572 (1.532)	2.344 (2.600)	2.940 (0.799)
ρ_{μ_z}	0.491 (0.128)	0.373 (0.075)	0.540 (0.000)	0.770 (0.046)	0.218 (0.272)
ρ_d	0.984 (0.001)	0.985 (0.000)	0.968 (0.005)	0.978 (0.002)	0.998 (0.010)
ρ_n	0.986 (0.008)	0.985 (0.005)	0.984 (0.004)	0.967 (0.041)	0.980 (0.026)
ρ_{π^*}	0.995 (0.011)	0.989 (0.000)	0.985 (0.009)	0.997 (0.003)	0.987 (0.015)
ρ_a	0.983 (0.008)	0.981 (0.004)	0.984 (0.000)	0.990 (0.001)	0.996 (0.027)
π_{ss}	1.021 (0.012)	1.021 (0.004)	1.034 (0.002)	1.018 (0.003)	1.024 (0.007)
$\sigma_{\mu_z} \times 100$	0.330 (0.013)	0.378 (0.072)	0.403 (0.039)	0.297 (0.003)	0.609 (0.342)
σ_d	0.127 (0.058)	0.122 (0.008)	0.152 (0.006)	0.033 (0.010)	0.053 (0.114)
σ_n	0.044 (0.017)	0.064 (0.010)	0.087 (0.007)	0.035 (0.002)	0.108 (0.111)
$\sigma_{\pi^*} \times 100$	0.043 (0.036)	0.060 (0.016)	0.247 (0.042)	0.083 (0.001)	0.086 (0.075)
$\sigma_a \times 100$	0.236 (0.129)	0.218 (0.043)	0.677 (0.039)	0.395 (0.021)	0.148 (0.509)
Additional Info					
Timing Premium	2.81%	2.77%	4.65%	7.33%	1.13%
IES	0.06	0.06	0.03	0.07	0.02
RRA	10	10	10	34.70	10
u_0	-459	-273	-5,945	0	-913,824

plausible average price markup of 13%. To aid the interpretation of the price adjustment parameter ξ, Table 1 reports the corresponding Calvo parameter $\hat{\xi}_{Calvo}$ that gives the same slope of the aggregate supply relation as $\hat{\xi}$ when linearizing this relation.[13] We find $\hat{\xi}_{Calvo} = 0.44$, which corresponds to an average price duration of just 1.8 quarters, showing that our New Keynesian model does not impose a strong degree of price stickiness. Instead, the model relies more on real wage stickiness with $\hat{\kappa}_w = 0.92$, which is similar to the finding in Smets and Wouters (2007).

In relation to the labor supply, we find $\hat{\varphi} = 0.10$ and therefore a Frisch elasticity of $\varphi\left(1/l_{ss} - 1\right) = 0.20$ in the steady state with $l_{ss} = 0.33$. Most micro estimates of the Frisch labor supply elasticity for males are in the range from 0.10 to 0.40 according to Keane (2011), showing that our low estimate of ϕ is consistent with micro evidence. With $\hat{\phi}$ and $\hat{\chi}$ well below and above one, respectively, the utility function in equation (2) is negative. Hence, the estimate $\hat{\alpha} = -11.46$ implies preferences for early resolution of uncertainty and a realistic timing premium of 2.81%. As discussed in Epstein, Farhi, and Strzalecki (2014), this means that the household is willing to give up 2.81% of total lifetime consumption to have all uncertainty resolved in the following period. Given these estimates, we obtain the required relative risk aversion of 10 when the utility constant is $u_0 = -459$. To interpret the magnitude of this constant, we compute the periodic utility $u\left(c_t, l_t\right)$ with and without a constant in a simulated sample to find the fraction of utility from consumption and leisure, i.e., $u\left(c_t, l_t\right)|_{u_0=0} / u\left(c_t, l_t\right)$. This fraction has an average of 12%, but displays considerable variation over the business cycle with a 5th percentile of 3% and a 95th percentile of 26%. As a result, the constant u_0 ensures a high and therefore also more stable utility level relative to its mean when compared to the standard specification with $u_0 = 0$. As argued by Andreasen and Jørgensen (2020), this helps to keep the household's risk aversion at a low and plausible level.

The forward-looking central bank assigns more weight to stabilizing future inflation than economic activity with $\hat{\phi}_\pi = 3.96$ and $\hat{\phi}_{\Delta c} = 1.10$, which is a common finding in the literature. For the structural shocks, we find that permanent productivity shocks display low persistence with $\hat{\rho}_{\mu_z} = 0.49$, showing that the model does not rely on a strong long-run risk component. In contrast, the inflation target is highly persistent with $\hat{\rho}_{\pi^*} = 0.995$ and it has low volatility with $\hat{\sigma}_{\pi^*} = 0.00043$, implying that this shock captures long-run nominal risk. Finally, the preference

[13] The mapping between ξ and ξ_{Calvo} is $\xi = \frac{(1-\theta+\eta\theta)(\eta-1)\xi_{Calvo}}{(1-\xi_{Calvo})(1-\theta)\left(1-\xi_{Calvo}\beta\mu_{z,ss}^{1-\chi}\right)}$.

shock, the labor supply shock, and the stationary productivity shock are also highly persistent with first-order autocorrelation functions of about 0.98.

Overall, the estimated structural parameters appear reasonable and consistent with micro evidence as well as previous findings for the New Keynesian model.

B Model Fit

Figure 1 plots the data and the model-implied values (the thick line) for the 10 variables included in the estimation. The model captures the overall evolution in the labor supply $\hat{l}_t$ and the real wage $\hat{w}_t$ remarkably well despite our parsimonious specification of the labor market. This is also evident from the correlation between the data series and the model-implied series as denoted by τ in the figure, which is 0.87 for the labor supply and 0.85 for the real wage. The model provides an even closer fit to consumption growth Δc_t and inflation π_t with $\tau = 0.98$. For the short rate r_t, we see an outlier around 1980, which is expected due to our choice of downweighting the importance of the short rate in the estimation during the monetary policy experiment. We also find that the model fits the short rate closely at the zero lower bound as desired, and that all remaining bond yields are well matched by the model.

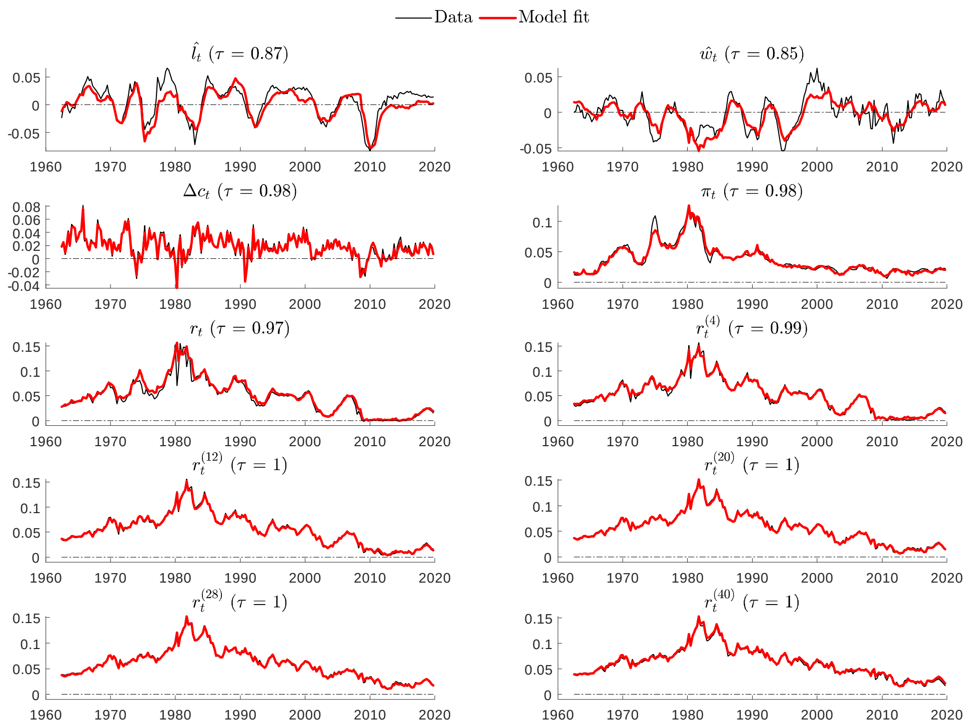
Accordingly, the proposed model with just five structural shocks provides a good in-sample fit to the 10 variables in $\mathbf{y}_t^{obs}$. To benchmark this result to the literature, recall that the New Keynesian model is typically estimated using as many shocks as observables. However, the dynamics of the yield curve is well captured by the first two principal components, implying that six structural shocks should be needed in our case. The fact that the proposed model matches the observed data using just five shocks is therefore very satisfying.

C Unconditional Stylized Moments

Table 2 explores the ability of $\mathcal{M}^{M,CS}$ to match the unconditional means and standard deviations in $\mathbf{y}_t^{obs}$ as included when estimating $\mathcal{M}^{M,CS}$. The first column in Table 2 shows U.S. sample moments and their bootstrapped 95% confidence bands, while the second column shows the corresponding population moments in $\mathcal{M}^{M,CS}$. We first note that $\mathcal{M}^{M,CS}$ reproduces the

Figure 1: Model Fit

This figure shows the observables in the data along with the model-implied values at the state estimates from the CDKF for the baseline model $\mathcal{M}^{M,CS}$, estimated on quarterly data from 1961 to 2019. The value of τ denotes the correlation between the empirical series and the model-implied series.



mean of inflation and the mean of all bond yields. For instance, the level of the 3-month yield is 4.80% vs. 4.58% in the data, and the level of the 10-year bond yield is 6.42% vs. 6.13% in the data. This implies that the average yield spread is 162 basis points in $\mathcal{M}^{M,CS}$ compared to 155 basis points in U.S. data.

The model is also successful in matching the standard deviation of the labor supply (2.78% vs. 2.88% in the data), the real wage (1.99% vs. 2.20% in the data), consumption growth (1.60% vs. 1.74% in the data), and inflation (2.58% vs. 2.43% in the data). The standard deviations of all bond yields are also well matched.[14]

Accordingly, $\mathcal{M}^{M,CS}$ explains the overall level and variability of the data. This shows that the satisfying in-sample fit in Figure 1 does not come at the expense of distorted unconditional

[14] Following the work of Piazzesi and Schneider (2007), the correlation between π_t and Δc_t has attracted some attention in the literature. We find $corr(\pi_t, \Delta c_t) = -0.16$ in $\mathcal{M}^{M,CS}$ and $corr(\pi_t, \Delta c_t) = -0.11$ in our quarterly sample from 1961 to 2019.

Table 2: Unconditional First and Second Moments

The data moments are for the U.S. from 1961 to 2019 with 95% confidence bands stated below. These bands are computed using a block bootstrap with 10,000 bootstrap samples using blocks of 60 observations. The model-implied moments are computed in closed form using the procedure in Andreasen et al. (2018). The version of the model is denoted by the superscipts "M" for applying shrinkage to the unconditional moments and "CS" to the Campbell-Shiller moments. The subscript on the model object indicates that the standard specification of recursive preferences (EZ) is used in column 5 and that a third-order perturbation approximation (3th) to the model solution is used in column 6. All means and standard deviations are stated in annualized percent, except for the standard deviation of $\hat{l}_t$ and $\hat{w}_t$ which are not annualized.

	Data	$\mathcal{M}^{M,CS}$	$\mathcal{M}^{M}$	$\mathcal{M}$	$\mathcal{M}^{M,CS}_{\text{EZ}}$	$\mathcal{M}^{M,CS}_{\text{3th}}$
	1	2	3	4	5	6
Means						
Δc_t	1.91 (1.39,2.44)	1.91	1.91	1.91	1.91	1.91
π_t	3.67 (1.93,5.41)	3.67	3.64	7.14	3.72	3.71
r_t	4.58 (2.21,6.96)	4.80	4.74	-1.92	4.78	4.54
$r_t^{(4)}$	5.04 (2.50,7.59)	4.93	4.97	-1.27	4.96	4.73
$r_t^{(12)}$	5.43 (2.91,7.95)	5.26	5.33	-0.26	5.44	5.25
$r_t^{(20)}$	5.70 (3.27,8.14)	5.58	5.61	0.26	5.79	5.64
$r_t^{(28)}$	5.91 (3.56,8.26)	5.91	5.88	0.66	6.04	5.97
$r_t^{(40)}$	6.13 (3.87,8.39)	6.42	6.29	1.24	6.29	6.38
Stds						
$\hat{l}_t$	2.88 (2.42,3.34)	2.78	2.91	6.49	2.91	2.36
$\hat{w}_t$	2.20 (1.65,2.75)	1.99	2.09	6.51	1.93	1.81
Δc_t	1.74 (1.49,1.98)	1.60	1.75	2.16	1.70	1.68
π_t	2.43 (1.18,3.67)	2.58	2.48	6.09	2.76	2.34
r_t	3.19 (2.07,4.32)	3.11	3.30	7.97	3.07	3.76
$r_t^{(4)}$	3.31 (2.16,4.46)	3.13	3.24	7.71	3.02	3.67
$r_t^{(12)}$	3.17 (2.02,4.32)	3.13	3.10	7.09	2.89	3.46
$r_t^{(20)}$	3.02 (1.89,4.15)	3.10	2.97	6.47	2.80	3.25
$r_t^{(28)}$	2.90 (1.78,4.01)	3.06	2.86	5.91	2.74	3.05
$r_t^{(40)}$	2.78 (1.68,3.88)	3.02	2.72	5.19	2.68	2.79

model dynamics. This indicates that the estimated state innovations $\hat{\boldsymbol{\epsilon}}_{t+1}$ evolve according to the assumed laws of motion in equation (10) and hence do not show evidence of overfitting. We draw the same conclusion from an inspection of $\hat{\boldsymbol{\epsilon}}_{t+1}$, which are basically uncorrelated as desired. For instance, the average of the absolute correlations between the five innovations is just 0.12, and the highest and lowest correlations are just 0.26 and −0.21, respectively.[15]

D Ordinary Campbell-Shiller Loadings

Another important question is whether $\mathcal{M}^{M,CS}$ can match the empirical pattern in the ordinary Campbell-Shiller loadings β_k from equation (14), which we include in the estimation. The top panel in Table 3 reports these loadings in the sample along with their bootstrapped 95% confidence bands. A very encouraging finding is that $\mathcal{M}^{M,CS}$ generates ordinary Campbell-Shiller loadings that are *negative* and track the empirical loadings remarkably well inside their 95% confidence bands, although the loadings in the model do not fall as much with time to maturity as in the data. This shows that the model satisfies the first requirement for a correct specification of term premia.

E Risk-Adjusted Campbell-Shiller Loadings

Next, we study moments that are not included in the estimation of $\mathcal{M}^{M,CS}$. A correct specification of term premia should imply that historical risk-adjusted bond yields satisfy the expectations hypothesis. Dai and Singleton (2002) show that this corresponds to testing whether the loading β_k^{Adj} is equal to one in the following version of the Campbell-Shiller regression

$$\begin{aligned} (15) \qquad & r_{t+m}^{(k-m)} - r_t^{(k)} - \left(TP_{t+m}^{(k-m)} - TP_t^{(k-m)}\right) + \frac{m}{k-m} TP_t^{fwd,(k-m)} \\ = \; & \alpha_k^{Adj} + \beta_k^{Adj} \frac{m}{k-m} \left(r_t^{(k)} - r_t^{(m)}\right) + u_{t+m,k}^{Adj}, \end{aligned}$$

where term premia are subtracted from long-term yields. Here, the variable $TP_t^{fwd,(k)} = f_{t,k} - \mathbb{E}_t\left[r_{t+k}\right]$ is the term premium in the forward rate $f_{t,k} \equiv -\log\left(B_t^{(k+1)}/B_t^{(k)}\right)$ and $u_{t+m,k}^{Adj}$ is an error term. That is, we substract the estimated term premia by the New Keynesian model from U.S. bond yields $r_{t+m}^{(k-m)} - r_t^{(k)}$ and regress them on a constant and the yield spread in the U.S.

The bottom panel in Table 3 uses term premia from $\mathcal{M}^{M,CS}$ to compute risk-adjusted

[15] The full correlation matrix for the structural innovations are reported in the Online Appendix.

Table 3: Ordinary and Risk-Adjusted Campbell-Shiller Loadings

The top panel shows the ordinary Campbell-Shiller regression loadings β_k for $m = 4$. The empirical values are computed using quarterly U.S. data from 1961 to 2019, with the 95 percent confidence interval provided in parentheses, computed using a block bootstrap where the regressand and the regressor in equation (14) are sampled jointly in blocks of 10 observations in 10,000 bootstrap samples. The model-implied Campbell-Shiller loadings are computed in closed form. The bottom panel shows the risk-adjusted Campbell-Shiller regression loadings β_k^{Adj} for $m = 4$ in quarterly U.S. data from 1961 to 2019 when using term premia from the particular version of the model with the 95 percent confidence interval provided in parentheses, computed using a block bootstrap where the regressand and the regressor in equation (15) are sampled jointly in blocks of 10 observations in 10,000 bootstrap samples.

	Data	$\mathcal{M}^{M,CS}$	$\mathcal{M}^{M}$	$\mathcal{M}$	$\mathcal{M}^{M,CS}_{\text{EZ}}$	$\mathcal{M}^{M,CS}_{\text{3th}}$
	1	2	3	4	5	6
Ordinary CS loadings						
β_{12}	−1.03 (−2.27,0.20)	−0.73	−0.28	0.49	−0.60	−0.72
β_{20}	−1.53 (−2.92,−0.14)	−0.74	−0.29	0.31	−0.36	−0.65
β_{28}	−1.93 (−3.43,−0.42)	−0.78	−0.30	0.20	−0.18	−0.64
β_{40}	−2.39 (−4.09,−0.70)	−0.82	−0.29	0.10	−0.01	−0.62
Adjusted CS loadings						
β_{12}^{Adj}	–	−0.50 (−1.57,0.99)	−0.33 (−1.41,1.21)	−0.57 (−1.78,1.35)	−0.90 (−1.93,0.57)	−0.04 (−1.73,1.65)
β_{20}^{Adj}	–	−0.49 (−1.74,1.23)	−0.34 (−1.58,1.36)	−0.59 (−1.72,1.11)	−1.42 (−2.60,0.20)	0.16 (−1.74,2.07)
β_{28}^{Adj}	–	−0.42 (−1.85,1.40)	−0.31 (−1.68,1.45)	−0.56 (−1.64,0.99)	−1.88 (−3.17,−0.15)	0.39 (−1.69,2.47)
β_{40}^{Adj}	–	−0.27 (−2.08,1.74)	−0.21 (−1.86,1.70)	−0.42 (−1.47,1.01)	−2.40 (−3.92,−0.52)	0.77 (−1.61,3.14)

Campbell-Shiller loadings β_k^{Adj} and their 95% confidence bands on the historical sample from 1961 to 2019. We find that these risk-adjusted loadings β_k^{Adj} are notably larger than the corresponding unadjusted loadings β_k, and hence closer to the desired value of one. For instance, at the 10-year maturity, the loading increases from $\beta_{40} = -2.39$ to $\beta_{40}^{Adj} = -0.27$. At the other maturities, we also find that the risk-adjusted Campbell-Shiller loadings are slightly below zero, but that the desired value of one generally is within the 95% confidence bands for β_k^{Adj}. The only exception is at the 1-year maturity, where the upper bound for the confidence interval is 0.99 and hence just below one. Subject to this minor qualification, term premia from the model are not rejected by the data, implying that $\mathcal{M}^{M,CS}$ also passes the second requirement for a correct specification of term premia.

F Inflation News in Bond Yields

Duffee (2018) shows that the quarterly news in bond yields $\tilde{r}_t^{(k)} = r_t^{(k)} - \mathbb{E}_{t-1}\left[r_t^{(k)}\right]$ is much more volatile than the quarterly news in expected inflation $\eta_{\pi,t}^{(k)} = \mathbb{E}_t\left[\frac{1}{k}\sum_{i=1}^{k}\pi_{t+i}\right] - \mathbb{E}_{t-1}\left[\frac{1}{k}\sum_{i=1}^{k}\pi_{t+i}\right]$. The inflation variance ratio $VR_\pi^{(k)} = \mathbb{V}[\eta_{\pi,t}^{(k)}]/\mathbb{V}[\tilde{r}_t^{(k)}]$ is therefore only around 0.15 in the U.S., whereas this ratio often exceeds one in various versions of the New Keynesian model (see Duffee (2018)). However, the time series for $\tilde{r}_t^{(k)}$ and $\eta_{\pi,t}^{(k)}$ in the U.S. display clear evidence of time-varying volatility, and Gomez-Cram and Yaron (2021) therefore refine the estimates in Duffee (2018) by correcting for heteroskedasticity. This increases $VR_\pi^{(k)}$ to 0.23, 0.22, and 0.19 at the 1-, 3-, and 5-year maturity, as shown in Table 4.[16]

The proposed model $\mathcal{M}^{M,CS}$ implies that the standard deviation in news to expected inflation is only 0.31 at the 1-year maturity, and falls steadily to 0.26 at the 5- and 10-year maturities. The standard deviation in news to bond yields is substantially higher at 0.49 at the 1-year maturity and 0.46 at the 3- and 5-year maturities, and hence within the reported 90% uncertainty bands for the corresponding data moments in Table 4. The model therefore implies fairly low inflation variance ratios of 0.41, 0.35, and 0.33 at the 1-, 3-, and 5-year maturities, respectively, which are just outside the reported 90% uncertainty bands. This shows that the proposed New Keynesian model goes a long way in addressing the critique of Duffee (2018). As with the risk-adjusted Campbell-Shiller loadings in Section IV.E, we emphasize that the model's ability to explain the inflation variance ratios is an out-of-sample test, in the sense that β_k^{Adj} and $VR_\pi^{(k)}$ are not included in the model estimation.

G Term Premia

The 10-year term premium implied by $\mathcal{M}^{M,CS}$ is shown at the top left graph in Figure 2, where the shaded bars denote NBER recessions. The graph reveals that the model generates the same overall pattern for the 10-year nominal term premium as found in the flexible 5-factor model of Adrian, Crump, and Moench (2013), which is a standard reduced-form Gaussian DTSM without any economic structure. The correlation between the two measures of term premium is

[16] We are grateful to Roberto Gomez-Cram for sharing these data moments.

Table 4: News to Inflation and Bond Yields

This table reports the annualized standard deviation of quarterly news in expected inflation $\eta_{\pi,t}^{(k)} = \mathbb{E}_t\left[\frac{1}{k}\sum_{i=1}^{k}\pi_{t+i}\right] - \mathbb{E}_{t-1}\left[\frac{1}{k}\sum_{i=1}^{k}\pi_{t+i}\right]$ in percent, the annualized standard deviation of quarterly news in bond yields $\tilde{r}_t^{(k)} = r_t^{(k)} - \mathbb{E}_{t-1}\left[r_t^{(k)}\right]$ in percent, and the inflation variance ratio $VR_{\pi}^{(k)} = \mathbb{V}[\eta_{\pi,t}^{(k)}]/\mathbb{V}[\tilde{r}_t^{(k)}]$. The data moments are from Gomez-Cram and Yaron (2021) (supported material provided by R. Gomez-Cram) with a correction for heteroskedasticity. These estimates are from 1962Q1 to 2018Q4, where the standard deviations in bond yields news use martingale forecasts, and the standard deviations in expected inflation news are obtained using the GDP deflator and surveys on current-quarter inflation. Figures in parentheses denote the 90 percent credibility interval. The corresponding model-implied moments are computed for the indicated versions of the New Keynesian model using a simulated sample of 100,000 observations to obtain the reported moments.

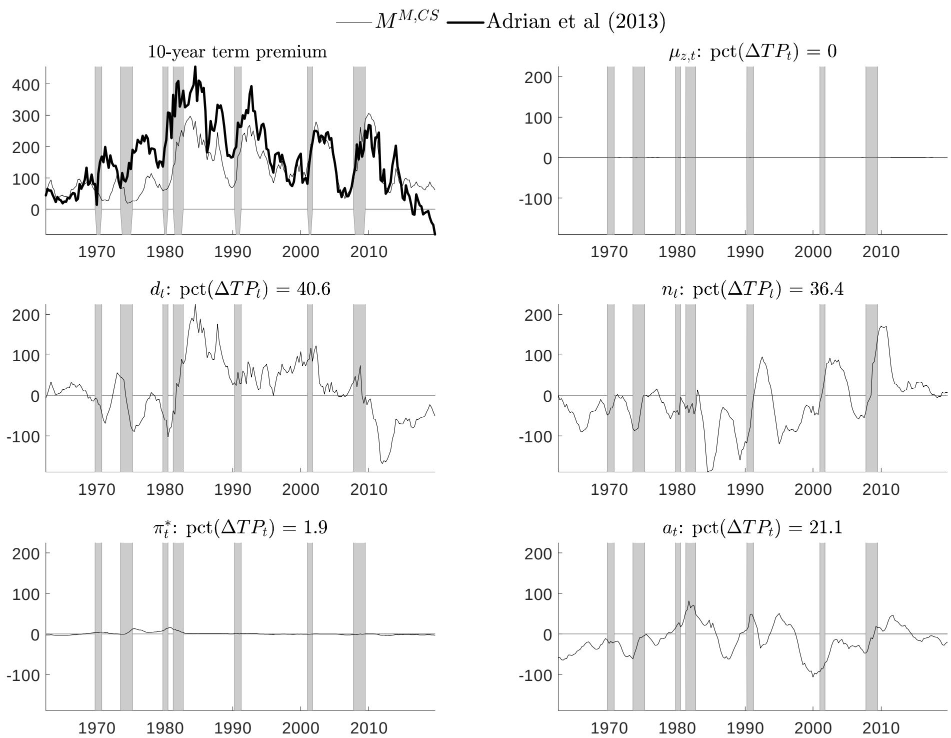
	Data	$\mathcal{M}^{M,CS}$	$\mathcal{M}^{M}$	$\mathcal{M}$	$\mathcal{M}_{\text{EZ}}^{M,CS}$	$\mathcal{M}_{\text{3th}}^{M,CS}$
	1	2	3	4	5	6
Std of inflation news						
1-year	0.22 [0.19,0.25]	0.31	0.39	0.94	0.27	0.37
3-year	0.21 [0.18,0.24]	0.27	0.35	0.88	0.27	0.34
5-year	0.20 [0.18,0.23]	0.26	0.34	0.85	0.27	0.32
10-year	-	0.26	0.31	0.77	0.27	0.28
Std of yield news						
1-year	0.46 [0.41,0.51]	0.49	0.62	1.86	0.48	0.45
3-year	0.44 [0.40,0.50]	0.46	0.56	1.39	0.43	0.44
5-year	0.46 [0.41,0.51]	0.46	0.53	1.24	0.39	0.43
10-year	-	0.45	0.48	1.01	0.33	0.41
Inflation variance ratio						
1-year	0.23 [0.16,0.32]	0.41	0.39	0.26	0.32	0.67
3-year	0.22 [0.16,0.32]	0.35	0.40	0.40	0.39	0.59
5-year	0.19 [0.14,0.27]	0.33	0.41	0.47	0.48	0.55
10-year	-	0.33	0.42	0.58	0.69	0.48

0.73, with the similarities being particularly strong since the early 1980s. Note also that both measures generally increase during NBER recessions and hence capture the counter-cyclical nature of term premia, with the two recessions during the 1970s as the main exceptions for $\mathcal{M}^{M,CS}$. The summary statistics for term premia are also very similar across the two models, as the 10-year nominal term premium in the model of Adrian et al. (2013) has a mean of 163 basis points and a standard deviation of 110 basis points, while the corresponding moments in $\mathcal{M}^{M,CS}$ are 151 and 100 basis points, respectively, in a simulated sample of $100,000$ observations.

To conduct a shock decomposition of the 10-year term premium, we condition on the

Figure 2: The 10-year Term Premium

The first graph shows the 10-year term premium in $\mathcal{M}^{M,CS}$ and in the model by Adrian et al. (2013). The remaining graphs show how each of the structural shocks in $\mathcal{M}^{M,CS}$ contribute to the variation in the 10-year term premium. The percentage of the total variation in the term premium explained by each shock, denoted pct(ΔTP_t), is computed using equation (16) based on the absolute variation in the series. The gray shaded bars denote NBER recessions, and term premium is expressed in annualized basis points.



nonlinear solution in equation (12) obtained using all five shocks and adopt the following procedure. First, let $TP^{(k),all}$ denote the term premium computed using the nonlinear solution and the estimated states $\hat{\mathbf{x}}_t$ from the CDKF using all shocks. Second, let $TP^{(k),-i}$ denote the term premium when omitting variation in $\hat{\mathbf{x}}_t$ due to the ith shock. The contribution to term premium by the ith shock is then $\Delta TP^{(k),i} \equiv TP^{(k),all} - TP^{(k),-i}$, which we plot in Figure 2 for each of the five shocks. To quantify the variation explained by a given shock, we compute the percentage of the total variation in the term premium explained by each shock as

$$\text{(16)} \qquad \text{pct}\left(TP^{(k),i}\right) = \frac{\sum_{t=1}^{T} \left| \Delta TP^{(k),i} - \overline{\Delta TP}^{(k),i} \right|}{\sum_{i=1}^{n_\varepsilon} \sum_{t=1}^{T} \left| \Delta TP^{(k),i} - \overline{\Delta TP}^{(k),i} \right|} \quad \text{for } i = \{1, 2, ..., n_\varepsilon\},$$

where $\overline{\Delta TP}^{(k),i}$ denotes the sample average of $\Delta TP^{(k),i}$. We find that demand and labor supply shocks are the two dominating forces, explaining 41% and 36%, respectively, of the variation in the 10-year term premium. Stationary productivity shocks account for 21%, whereas shocks to the inflation target π_t^* and the growth rate $\mu_{z,t}$ have hardly any effect on the 10-year term premium. The decomposition also shows that the elevated level of the term premium during the two recessions in the early 1980s are due to demand and stationary productivity shocks. For the 1991 and the 2001 recessions, labor supply shocks and stationary productivity shocks are the main drivers behind the term premium. More recently, we see that demand, labor supply, and stationary productivity shocks all contribute to generate a high term premium during the 2007-2009 Great Recession.

H Understanding the Key Mechanisms in the Model

At this point, we have shown that the proposed model i) explains historical bond yields, ii) matches unconditional properties of the data, iii) passes the two requirements for a correct specification of term premia, and iv) goes a long way in matching the level of news in expected inflation. Most of these features are also matched by reduced-form DTSMs, but these models offer little insights into the economic mechanisms that determine bond yields and especially term premia. The model we propose provides such a structural explanation, and this section analyzes the key mechanisms in $\mathcal{M}^{M,CS}$ that drive term premia. Our discussion is structured around Table 5, which shows how unconditional moments and Campbell-Shiller loadings (both ordinary and risk-adjusted) are affected when omitting one of the five shocks in the model.

We first note that permanent productivity shocks have a large impact on the level of bond yields, as omitting variation in $\mu_{z,t}$ (i.e., $\sigma_{\mu_z} = 0$) increases the average level of the yield curve by about five percentage points. For instance, the mean of the 10-year yield increases from 6.42% with all five shocks to 11.60% when omitting permanent productivity shocks. This effect arises from a precautionary saving motive, as omitting permanent productivity shocks make the economy less risky and this reduces the negative precautionary saving correction in bond yields. We also note that omitting permanent productivity shocks reduce the standard deviations in the

Table 5: Decomposing the Effects of the Structural Shocks

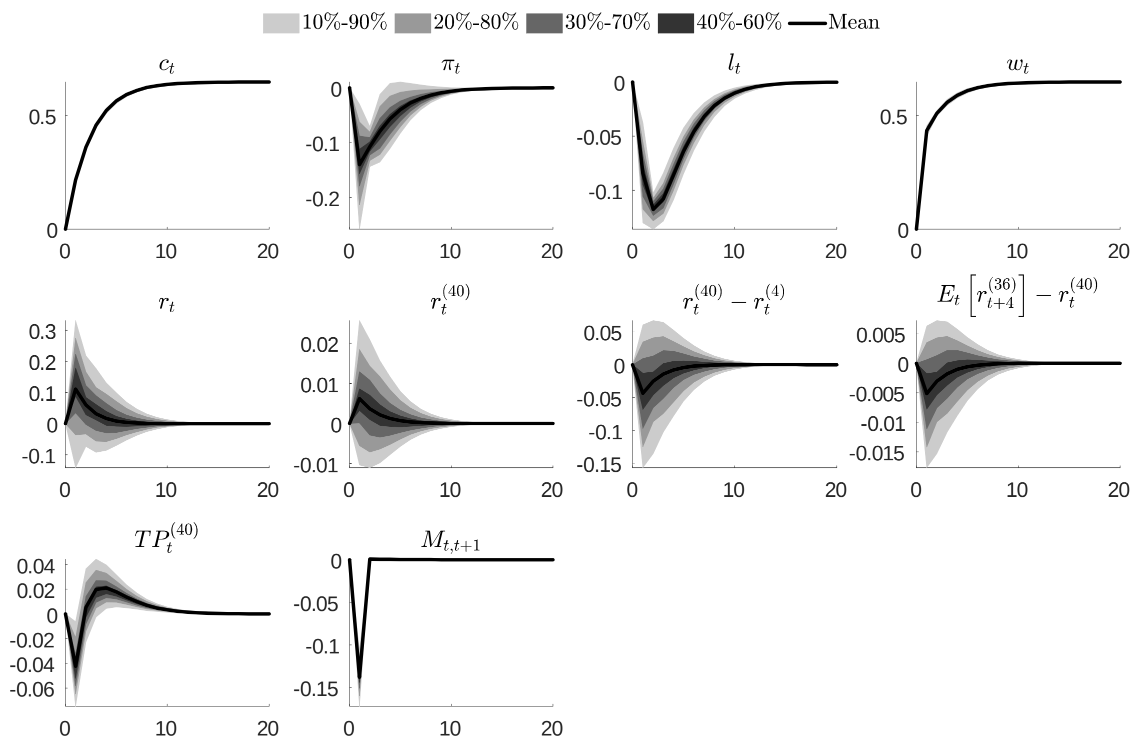
This table reports unconditional moments for $\mathcal{M}^{M,CS}$ when all structural shocks are present in column 1, and when each of the structural shocks are omitted in columns 2 to 6. The model-implied moments are computed in closed form using the procedure in Andreasen et al. (2018). All means and standard deviations are stated in annualized percent, except for the standard deviation of $\hat{l}_t$ and $\hat{w}_t$ which are not annualized.

	$\mathcal{M}^{M,CS}$	$\sigma_{\mu_z}=0$	$\sigma_d=0$	$\sigma_n=0$	$\sigma_{\pi^*}=0$	$\sigma_a=0$
	1	2	3	4	5	6
Means						
Δc_t	1.91	1.91	1.91	1.91	1.91	1.91
π_t	3.67	4.70	7.00	3.76	3.67	3.66
r_t	4.80	9.27	18.43	5.34	4.76	3.56
$r_t^{(4)}$	4.93	9.45	18.43	5.51	4.88	3.51
$r_t^{(12)}$	5.26	9.92	18.51	5.96	5.20	3.58
$r_t^{(20)}$	5.58	10.40	18.54	6.37	5.51	3.85
$r_t^{(28)}$	5.91	10.88	18.56	6.77	5.83	4.22
$r_t^{(40)}$	6.42	11.60	18.57	7.37	6.34	4.91
Stds						
$\hat{l}_t$	2.78	1.82	1.81	1.51	2.78	4.47
$\hat{w}_t$	1.99	1.30	1.32	1.52	1.98	2.90
Δc_t	1.60	0.80	1.42	1.44	1.60	1.88
π_t	2.58	2.50	2.44	2.45	0.82	3.61
r_t	3.11	3.49	2.66	2.84	1.77	5.16
$r_t^{(4)}$	3.13	3.45	2.60	2.97	1.85	3.97
$r_t^{(12)}$	3.13	3.41	2.51	3.07	1.93	3.33
$r_t^{(20)}$	3.10	3.39	2.45	3.08	1.93	3.24
$r_t^{(28)}$	3.06	3.37	2.39	3.07	1.94	3.25
$r_t^{(40)}$	3.02	3.36	2.30	3.05	1.96	3.34
Ordinary CS loadings						
β_{12}	-0.73	-0.51	1.02	-0.87	-0.74	-0.62
β_{20}	-0.74	-0.70	1.03	-1.07	-0.76	-0.55
β_{28}	-0.78	-0.78	1.03	-1.24	-0.81	-0.57
β_{40}	-0.82	-0.83	1.04	-1.40	-0.87	-0.62
Adjusted CS loadings						
β_{12}^{Adj}	-0.50	-0.63	-0.99	-0.93	-0.61	-0.29
β_{20}^{Adj}	-0.49	-0.76	-1.49	-1.59	-0.79	-0.13
β_{28}^{Adj}	-0.42	-0.76	-1.89	-2.15	-0.93	0.03
β_{40}^{Adj}	-0.27	-0.58	-2.35	-2.85	-1.07	0.26

three macro variables $\hat{l}_t$, $\hat{w}_t$, and Δc_t, whereas the variability in π_t is almost unaffected. On the other hand, the standard deviation in bond yields actually increase when imposing that $\sigma_{\mu_z}=0$. This may at first seem counter-intuitive, but this effect arises because the nonlinear policy function for bond yields changes with $\sigma_{\mu_z}=0$ such that they depend more strongly on the other structural shocks due to a sizable risk-adjustment in the model solution. The ordinary and risk-adjusted Campbell-Shiller loadings are not materially affected by letting $\sigma_{\mu_z} = 0$. This suggests that permanent productivity shocks are not the key source for generating bond return predictability and hence variability in term premia, consistent with the results in Figure 2.

Figure 3: Impulse Response Functions: A Permanent Productivity Shock

This figure shows the generalized impulse response functions (IRFs) for a positive one-standard deviation shock to $\mu_{z,t}$, where the various shadings cover the indicated fraction of the distribution for these IRFs obtained using 1,000 randomly generated initial states. The IRFs are computed in closed-form for the second order projection solution using the approach in Andreasen et al. (2018), except for $M_{t,t+1}$ which is computed using Monte Carlo integration with 1,000 draws. Except for the term premium $TP_t^{(40)}$ and the nominal stochastic discount factor $M_{t,t+1}$, all impulse response functions are expressed in percentage deviations (i.e., scaled by 100) from the steady state or from the deterministic growth path (in the case of consumption and the wage level). All bond yields and inflation are measured in annualized terms and the term premium $TP_t^{(40)}$ is expressed in annualized basis points.



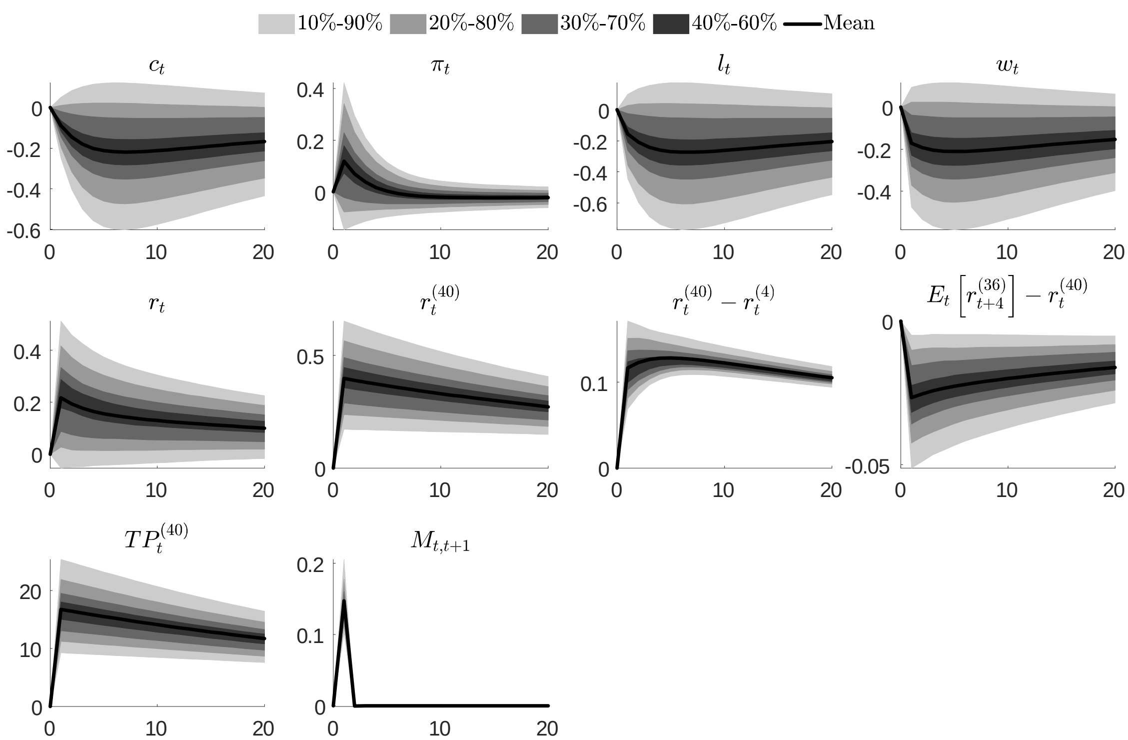
To further analyze the effects a change in productivity, Figure 3 shows the generalized impulse response functions (IRFs) following a positive one-standard deviation permanent productivity shock. These IRFs depend on the initial state $\mathbf{x}_t$ when the shock hits in period $t + 1$, and we therefore show fan charts for the distribution of these responses across 1,000 randomly generated initial states from a simulated sample of the model. Consistent with the results in Table 5, the 10-year term premium is basically unaffected by this shock. On the other hand, consumption and the wage level increase strongly, whereas inflation shows a small reduction. These findings are consistent with the empirical evidence provided in Altig, Christiano, Eichenbaum, and Linde (2011) using quarterly U.S. data from 1982 to 2008. However, the drop in the labor supply in Figure 3 is contrary to the finding in Altig et al. (2011), but Fisher (2006) shows that the reponse in the labor supply is sensitive to the considered sample, as it is negative when using data before 1982 but positive when using data after 1982. Accordingly, the IRFs to a permanent productivity shock appear consistent with existing empirical evidence. This suggests that the shock is well specified in the model, and that its responses do not appear to be distorted by letting z_t enter in the utility function in equation (2).

Our second observation is that demand shocks generate an upward sloping yield curve. This is seen from the third column in Table 5, which shows an almost flat yield curve close to the steady state of 18.50% when omitting demand shocks (i.e., $\sigma_d = 0$). In contrast, demand shocks often generate a downward sloping yield curve in other versions of the New Keynesian model (see, for instance, Nakata and Tanaka (2016)). To understand this implication of $\mathcal{M}^{M,CS}$, consider the IRFs in Figure 4 for a positive demand shock. We generally find that inflation π_t and the long-term bond yield $r_t^{(40)}$ increase, and hence that $B_t^{(40)}$ falls following a demand shock. On the other hand, the responses of consumption, labor supply, and the real wage may be either positive or negative. In contrast, when solving the model without accounting for risk (as typically done in the macro literature), we find almost exclusively positive responses in consumption, the labor supply, and the real wage to a demand shock, as shown in the Online Appendix. This means that the sizable risk-adjusted required in the model to match bond yields has notable macroeconomic effects. Despite these state-dependent responses, we see a large increase in the nominal stochastic discount factor $M_{t.t+1}$ due to $d_{t+1}/d_t > 1$ and an increase in $\left(\mathbb{E}_t\left[V_{t+1}^{1-\alpha}\right]\right)^{\frac{\alpha}{1-\alpha}}/V_{t+1}^{\alpha}$. As shown in Rudebusch and Swanson (2012), such a negative comovement between $M_{t,t+1}$ and the bond price $B_t^{(40)}$ over the lifetime of the bond generates a positive term premium. Thus, demand shocks make long-term bonds a risky investment for the household, and this explains why they generate a strong increase in the 10-year term premium $TP_t^{(40)}$ of about

15 basis points and an upward sloping yield curve.

Figure 4: Impulse Response Functions: A Demand Shock

This figure shows the generalized impulse response functions (IRFs) for a positive one-standard deviation shock to d_t, where the various shadings cover the indicated fraction of the distribution for these IRFs obtained using 1,000 randomly generated initial states. The IRFs are computed in closed-form for the second order projection solution using the approach in Andreasen et al. (2018), except for $M_{t,t+1}$ which is computed using Monte Carlo integration with 1,000 draws. Except for the term premium $TP_t^{(40)}$ and the nominal stochastic discount factor $M_{t,t+1}$, all impulse response functions are expressed in percentage deviations (i.e., scaled by 100) from the steady state or from the deterministic growth path (in the case of consumption and the wage level). All bond yields and inflation are measured in annualized terms and the term premium $TP_t^{(40)}$ is expressed in annualized basis points.



Another important observation from Figure 4 is that demand shocks increase the yield spread $r_t^{(40)} - r_t^{(4)}$ and reduce the expected change in the long yield $\mathbb{E}_t\left[r_{t+4}^{(36)}\right] - r_t^{(40)}$. Demand shocks therefore also help to generate the desired negative estimates of $\beta^{(k)}$ in the ordinary Campbell-Shiller regressions. Indeed, the ordinary Campbell-Shiller loadings are basically one without demand shocks, and the model-implied risk-adjustment in U.S. bond yields is negligible with risk-adjusted Campbell-Shiller loadings that almost coincide with the unadjusted slope coefficients in the first column in Table 3. This shows that demand shocks are a key driver for generating return predictability and variation in term premia in the model. Hence, when omitting these shocks we greatly reduce the variation in term premia and therefore also the standard deviation in bond yields, as shown in the third column in Table 5. In contrast, the standard deviation of inflation only drops slightly from 2.58% to 2.44% without demand shocks. These shocks therefore help to explain the variation in bond yields without materially affecting inflation, which is precisely what is required to match the low inflation variance ratios discussed in Section IV.F. Indeed, the inflation variance ratio for the 10-year bond yield increases from 0.33 to 0.65 without demand shocks in the model.

Given the importance of the demand shock for the New Keynesian model to match bond yields, we briefly explore how well this shock relates to the widely used consumer sentiment index from the University of Michigan. This index is obtained by surveying a cross-section of the population in the U.S. to capture their views about the current state of their own finances and the U.S. economy in general. The index therefore responds to aggregate shocks hitting the U.S. economy such as changes in productivity and government spendings, but also to variation within the household that are unrelated to such external shocks, as discussed in Lagerborg, Pappa, and Ravn (2023). This means that we should not expect a perfect correlation between this sentiment index and our measure of the demand shock. A regression of the consumer sentiment index on a constant and the estimated demand shock $\hat{d}_t$ from $\mathcal{M}^{M,CS}$ shows a positive correlation of 0.42, with a 95% confidence interval of $(0.24, 0.59)$ obtained using Newey-West standard errors with four lags. This shows that the demand shock is not simply some unobserved statistical factor extracted to fit bond yields, but instead is well connected to the level of consumer sentiment in the U.S. and hence has a clear economic interpretation.

The last columns in Table 5 show that i) labor supply shocks n_t unsurprisingly explain much of the variation in the labor supply, ii) shocks to π_t^* control much of the variation in inflation and bond yields, and that iii) stationary productivity shocks a_t mainly affect the level of bond yields and the variation in the labor supply, the real wage, inflation, and short-term bond yields.

V Sensitivity Analysis

This section explores various aspects of the baseline model $\mathcal{M}^{M,CS}$ that are new to the literature. We first examine the effects of the estimator with shrinkage in Section V.A, while Section V.B applies the standard formulation of recursive preferences. Finally, we discuss the accuracy of the projection solution in Section V.C and relate it to the standard perturbation approximation.

A The Effects of Shrinkage

The results presented so far are obtained when shrinking the traditional QML estimates of the structural parameters towards some stylized empirical moments for $\mathbf{y}_t^{obs}$. Given that this estimator is new to the literature on the New Keynesian model, we next explore the effects of gradually removing this form of shrinkage. We first re-estimate the model when only shrinking the QML estimates to the first and second unconditional moments of $\mathbf{y}_t^{obs}$ and not to the ordinary Campbell-Shiller loadings. This version of the model is denoted $\mathcal{M}^{M}$. Table 1 shows that this modification implies slightly higher curvature in the utility function ($\chi = 7.1$ vs. $\chi = 5.8$), lower price stickiness ($\xi_{Calvo} = 0.31$ vs. $\xi_{Calvo} = 0.44$), and a lower weight on consumption growth in the Taylor-rule ($\phi_{\Delta c} = 0.66$ vs. $\phi_{\Delta c} = 1.10$) when compared to $\mathcal{M}^{M,CS}$. The unconditional means and standard deviations in Table 2 are basically unchanged, whereas the standard deviations of news to inflation and bond yields in Table 4 increase and imply larger inflation variance ratios in $\mathcal{M}^{M}$ than in $\mathcal{M}^{M,CS}$. We further find that $\mathcal{M}^{M}$ generates bond return predictability with ordinary Campbell-Shiller loadings around -0.3, which are somewhat lower (in absolute terms) than in $\mathcal{M}^{M,CS}$ and the data. The risk-adjusted Campbell-Shiller loadings for $\mathcal{M}^{M}$ increase slightly compared to $\mathcal{M}^{M,CS}$, with their 95% confidence bands including the desired value of one, as seen from Table 3.

Another possibility is to abstract from any shrinkage and simply use QML for the estimation. We refer to this version of the model as $\mathcal{M}$. This modification gives somewhat larger changes in the estimated structural parameters in Table 1, with stronger habit formation ($b = 0.84$ vs. $b = 0.65$), less wage stickiness ($\kappa_w = 0.83$ vs. $\kappa_w = 0.92$), and a much larger weight on consumption growth in the Taylor-rule ($\phi_{\Delta c} = 4.57$ vs. $\phi_{\Delta c} = 1.10$) when compared to $\mathcal{M}^{M,CS}$. In addition, the standard deviations for all five structural shocks increase in $\mathcal{M}$ compared to $\mathcal{M}^{M,CS}$, which illustrates the tendency of QML to select large structural shocks in the New Keynesian model, as discussed in Section III.C. By construction, these estimates for $\mathcal{M}$ give a higher value of the log-likelihood function and therefore a closer in-sample fit than in both $\mathcal{M}^{M}$ and $\mathcal{M}^{M,CS}$ that apply shrinkage. But Table 2 shows that $\mathcal{M}$ completely misses the i) mean of inflation, ii) the mean of bond yields, and iii) generates too much volatility in the observables $\mathbf{y}_t^{obs}$, with all standard deviations exceeding their 95% confidence bands. This imply that the standard deviations of news in expected inflation and bond yields are much larger than in the data according to Table 4. Thus, the standard QML estimates display clear signs of overfitting, as the improved value of the quasi log-likelihood function compared to $\mathcal{M}^{M,CS}$ and $\mathcal{M}^{M}$ comes at the cost of distorting the fit to several unconditional properties of the model. This also distorts estimates of term premia in $\mathcal{M}$, as its ordinary Campbell-Shiller loadings are all positive and outside of the 95% confidence bands for the corresponding sample moments according to Table 3.

We therefore conclude that shrinking in the estimation of the structural parameters is essential for the New Keynesian model to provide a satisfying in-sample fit and match several unconditional aspects of bond yields and macroeconomic variables.

B Standard Formulation of Recursive Preferences

Our baseline model $\mathcal{M}^{M,CS}$ uses the flexible formulation of recursive preferences in Andreasen and Jørgensen (2020) with the utility constant u_0 to separate the IES, RRA, and the timing attitude. We next re-estimate the model with $u_0 = 0$ to recover the standard specification of recursive preferences applied in Rudebusch and Swanson (2012), Binsbergen, Fernandez-Villaverde, Koijen, and Rubio-Ramirez (2012), many among others. We refer to this version of the model as $\mathcal{M}_{\text{EZ}}^{M,CS}$, which we estimate using the same shrinkage specification as in our baseline model. Table 1 shows that this seemingly minor modification implies substantial changes in the estimated structural parameters. In particular, we find stronger habit formation

($b = 0.87$ vs. $b = 0.65$), less curvature in the utility function ($\chi = 1.95$ vs. $\chi = 5.83$), much more price stickiness ($\xi_{Calvo} = 0.77$ vs. $\xi_{Calvo} = 0.44$), a larger weight on consumption growth in the Taylor-rule ($\phi_{\Delta c} = 2.98$ vs. $\phi_{\Delta c} = 1.10$), much smaller demand shocks ($\sigma_d = 0.033$ vs. $\sigma_d = 0.127$), and larger stationary productivity shocks ($\sigma_a = 0.0040$ vs. $\sigma_a = 0.0024$) when compared to $\mathcal{M}^{M,CS}$. As a result, we obtain a high RRA of 35 in $\mathcal{M}_{\text{EZ}}^{M,CS}$ with $\alpha = -16.49$, as typically found when estimating the New Keynesian model with recursive preferences.

The unconditional means and standard deviations in Table 2 are almost perfectly matched by $\mathcal{M}_{\text{EZ}}^{M,CS}$. However, Table 3 shows that the model is unable to reproduce the ordinary Campbell-Shiller loadings at the 7- and 10-year maturities, and it misses completely the risk-adjusted Campbell-Shiller loadings, which are far from the desired value of one. These findings suggest that term premia are not appropriately specified in this version of the New Keynesian model. Finally, Table 4 shows that the inflation variance ratios increase monotonically with the time to maturity and clearly become too large. For instance, the inflation variance ratio is 0.48 and 0.69 at the 5- and 10-year maturity in $\mathcal{M}_{\text{EZ}}^{M,CS}$ but only 0.19 in US data at the 5-year maturity.

To conclude, the flexible formulation of recursive preferences in Andreasen and Jørgensen (2020) is required to enable the New Keynesian model to appropriately capture the dynamics of bond yields and their embedded term premia.

C Accuracy of the Model Solution

All versions of the New Keynesian model have so far been estimated using the second-order projection solution outlined in Section III.B. Table 6 studies the accuracy of this approximation on a grid of $S = 10,000$ simulated states and compare its performance to the standard perturbation approximation up to fifth order as made available by Levintal (2017). We report accuracy separately for the $n_m = 9$ equations defining the macro block of the model and the $n_b = 40$ equations defining bond prices. This is done using the standard summary statistics

$$\begin{aligned}\text{MAE} &= \log 10\left(\frac{1}{nS}\sum_{s=1}^{S}\sum_{i=1}^{n}|EE_{i,s}|\right) \\ \text{RMSE} &= \log 10\left(\frac{1}{n}\sum_{i=1}^{n}\sqrt{\frac{1}{S}\sum_{s=1}^{S}EE_{i,s}^{2}}\right) \\ \text{Max AE} &= \log 10\left(\frac{1}{n}\sum_{i=1}^{n}\max\left(\{|EE_{i,s}|\}_{s=1}^{S}\right)\right)\end{aligned} \tag{17}$$

for $n = n_m$ and $n = n_b$. Here, $EE_{i,s}$ is the ith unit-free Euler-equation error at state value $\boldsymbol{x}_s$, with log 10 denoting the base 10 logarithm (e.g., $\log 10(0.01) = -2$).

The top panel in Table 6 considers the macro block of the baseline model $\mathcal{M}^{M,CS}$, where the second-order perturbation approximation clearly does better than the certainty equivalent first-order solution. But accuracy does surprisingly not improve monotonically when increasing the approximaiton order for standard perturbation. Our second-order projection solution gives marginally higher accuracy than even a fifth order perturbation solution for the MAE and the RMSE (as indicated by the bold figures), whereas the second-order perturbation solution gives the lowest maximum absolute error. For the bond equations, we see larger differences between the two solution methods, where the second-order projection solution gives MAE$= -2.74$ and RMSE$= -2.34$, whereas we find -1.97 and -1.75, respectively, for a fifth order perturbation solution. Analogue to the macro block, we also note that the performance of standard perturbation for bond prices does not improve monotonically when increasing the approximation order.

The middle part of Table 6 considers $\mathcal{M}_{\text{EZ}}^{M,CS}$ with the standard specification of recursive preferences. The results show that the perturbation solution breaks down in this case, as the Euler-errors are very large and increase with the approximation order. In contrast, our second-order projection solution performs well also in this case.

To illustrate what drives this disappointing performance of the perturbation approximation, consider a restricted version of the baseline model without demand shocks by letting $\sigma_d = 0$. This parametrization is informative, because it implies that the deterministic steady state is close to the unconditional mean (as shown for bond yields in Section IV.H), and

Table 6: Accuracy of Model Solutions

This table shows the mean absolute errors (MAE), the root mean squared errors (RMSE), and the maximum absolute errors (Max AE) as defined in equation (17) for the nine equations defining the macro block of the New Keynesian model and for the 40 equations defining log-transformed bond prices for the 10-year yield curve. These accurary measures are computed on a simulated grid of 10,000 observations obtained from the considered version of the model using a second-order perturbation approximation. Conditional expectations in Euler-equations are evaluated using Gauss-Hermite quadrature with five points for each of the five structural shocks, giving a total of 5^5 integration points for a given state value. A bold figure indicates the best performing model along a given accuracy measure for a given model.

		Perturbation					Projection
		1st	2nd	3rd	4th	5th	2nd
		1	2	3	4	5	6
$\mathcal{M}^{M,CS}$							
Macro Eqs	MAE	−1.66	−2.59	−2.54	−2.78	−2.73	**−2.76**
	RMSE	−1.65	−2.48	−2.41	−2.46	−2.46	**−2.51**
	Max AE	−1.49	**−1.75**	−1.44	−1.26	−1.21	−1.47
Bond Eqs	MAE	−1.04	−2.12	−1.87	−2.09	−1.97	**−2.74**
	RMSE	−1.04	−2.00	−1.77	−1.79	−1.75	**−2.34**
	Max AE	−1.03	**−1.27**	−0.83	−0.61	−0.55	−1.21
$\mathcal{M}^{M,CS}_{\text{EZ}}$							
Macro Eqs	MAE	−1.42	11.56	6.43	$\pm\infty$	$\pm\infty$	**−2.40**
	RMSE	−1.20	13.46	8.02	$\pm\infty$	$\pm\infty$	**−1.89**
	Max AE	−0.46	15.45	9.84	$\pm\infty$	$\pm\infty$	**−0.79**
Bond Eqs	MAE	−1.29	−1.07	−1.06	217.18	204.95	**−2.63**
	RMSE	−1.29	−0.46	−0.09	$\pm\infty$	$\pm\infty$	**−2.00**
	Max AE	−1.28	0.78	1.50	221.17	208.94	**−0.89**
$\mathcal{M}^{M,CS}_{\sigma_d=0}$							
Macro Eqs	MAE	−1.75	−3.64	−3.60	−4.80	**−4.94**	−4.02
	RMSE	−1.75	−3.56	−3.57	−4.47	**−4.53**	−3.82
	Max AE	−1.63	−2.50	−2.58	−2.69	**−2.72**	−2.52
Bond Eqs	MAE	−1.12	−4.10	−3.17	−5.00	**−5.08**	−4.30
	RMSE	−1.12	−3.99	−3.17	−4.93	**−5.02**	−4.17
	Max AE	−1.12	−3.05	−3.10	−3.95	**−4.44**	−3.36

the perturbation approximation should therefore perform well in this case. This is confirmed in the bottom part of Table 6, as the fifth-order perturbation solution delivers a very high degree of accuracy in this case and clearly outperforms our second-order projection solution.

To show that the accuracy of the model solution has important consequences for the New Keynesian model, we finally re-estimate the model using a third-order perturbation solution, as widely applied in the literature (see Andreasen (2012), Rudebusch and Swanson (2012), among many others). This version of the model is denoted $\mathcal{M}^{M,CS}_{\text{3rd}}$. We find that this model performs well in matching the unconditional means and standard deviations in Table 2 and the Campbell-Shiller loadings in Table 3, but that the variance ratios in Table 4 are somewhat higher than with the projection solution. However, the structural parameters in Table 1 differ substantially from those obtained using the projection solution. For instance, with the third-order perturbation approximation the subjective discount factor becomes very low ($\beta = 0.948$ vs. $\beta = 0.991$) and the curvature in the utility function becomes very high ($\chi = 15.63$ vs. $\chi = 5.83$) when compared to $\mathcal{M}^{M,CS}$. Also, the weight on inflation in the Taylor-rule ϕ_π becomes implausible large and hits the imposed upper bound of 10.

Accordingly, the proposed second-order projection solution delivers high accuracy and ensures sensible structural estimates, whereas the performance of standard perturbation deteriorates when the deterministic steady state is too far from the unconditional mean.

VI Additional Model Implications

This section considers the baseline model $\mathcal{M}^{M,CS}$ and studies three additional aspects of the model that are not included in the estimation, and hence can be considered as representing over-identified restrictions. Joslin et al. (2014) criticize a wide class of macro-finance term structure models because they are unable to address the so-called "spanning puzzle". We summarize the arguments of Joslin et al. (2014) in Section VI.A and show that $\mathcal{M}^{M,CS}$ goes a long way in addressing this puzzle. The model's ability to capture the dynamics of real bond yields are examined in Section VI.B, while Section VI.C studies its implications for the price-dividend ratio and equity returns.

A The Yield Curve and Spanned Macro Variation

Many DTSMs with macro variables imply that bond yields are a linear combination of latent factors and observed macro variables. This linear mapping can (up to knife-edge restrictions) be inverted to express macro variables as a function of bond yields. All variation in macro variables is therefore spanned (i.e., explained) by bond yields. However, this implication is heavily criticized by Joslin et al. (2014) because regressions of macro variables on linear combination of bond yields such as the principal components typically generate R^2 values substantially below one. This constitutes the "spanning puzzle", which leads Joslin et al. (2014) to challenge the usefulness of most equilibrium models - including the standard New Keynesian model. The model we propose also implies that bond yields depend on macro variables, but this relation is not linear. Thus, the nonlinear structure of our model may generate what may appear to be unspanned variation in macro variables and hence resolve the spanning puzzle.

Table 7: Regression Evidence on the Spanning Hypothesis

This table shows the R^2 in percent from the regression $m_t = \alpha + \beta' \mathbf{pca}_t^{(K)} + u_t$, where m_t refers to a macro variable and $\mathbf{pca}_t^{(K)}$ of dimension $K \times 1$ contains the first K principal components of bond yields. The R^2s in the data are provided in column 1 and computed on quarterly U.S. data from 1961 to 2019, with the 95 percent confidence interval (shown below the estimate) computed using a block bootstrap, where the regressand and the regressors are sampled jointly in blocks of 50 observations in 10,000 bootstrap samples. The corresponding R^2s implied by the New Keynesian model $\mathcal{M}^{M,CS}$ are computed using a simulated sample of 100,000 observations. The simulated sample in column 2 is obtained using the nonlinear model solution (i.e., second-order projection) and with measurement errors for macro variables and bond yields of the same size as considered in the estimation. The same nonlinear model solution is used in column 3, but no measurement errors are added to the macro variables and bond yields. The simulated sample in column 4 uses a simplified second-order projection solution, where all terms that are quadratic in the states are omitted.

	Data	Nonlinear model & Measurement errors	Nonlinear model	Linear model
	1	2	3	4
$K = 5$				
$\hat{l}_t$	26.23 (10.62,60.97)	32.39	51.58	88.92
$\hat{w}_t$	16.53 (4.32,41.36)	34.18	73.74	94.61
Δc_t	30.40 (15.63,46.84)	3.44	18.76	74.06
π_t	64.08 (40.48,80.05)	84.57	86.36	98.95
$K = 6$				
$\hat{l}_t$	28.11 (11.42,63.42)	32.41	52.05	100.00
$\hat{w}_t$	17.31 (5.44,42.89)	34.21	79.64	100.00
Δc_t	30.40 (16.53,47.02)	3.46	68.56	100.00
π_t	64.34 (40.82,80.50)	84.57	90.85	100.00

We explore this possibility in Table 7 by regressing each of the four macro variables included in the model estimation on the first K principal components of bond yields. The model $\mathcal{M}^{M,CS}$ has five structural shocks, suggesting that at least five principal components are required to span variation in the macro variables. The first column in Table 7 shows that U.S. bond yields only explain a fairly small proportion of the variation in the labor supply $\hat{l}_t$ (26.2%), the real wage $\hat{w}_t$ (16.5%), and consumption growth Δc_t (30.4%), whereas inflation π_t displays some evidence of spanning with an R^2 of 64.1%. The results for the New Keynesian model are very similar to these regressions according to the second column in Table 7, where the corresponding values of the R^2 are 32.4% for $\hat{l}_t$, 34.2% for $\hat{w}_t$, 3.4% for Δc_t, and 84.6% for π_t. These figures are computed using a simulated sample of $100,000$ observations that include the same measurement errors $\mathbf{v}_t$ as applied when estimating $\mathcal{M}^{M,CS}$. Measurement errors help to generate unspanned variation, as emphasized in Bauer and Rudebusch (2017), and we therefore omit these errors in the third column in Table 7. The R^2 values then increase to 51.6% for $\hat{l}_t$, 73.7% for $\hat{w}_t$, 18.8% for Δc_t, and 86.4% for π_t

To evaluate the degree of unspanned variation that is generated by the nonlinear structure of the New Keynesian model, the fourth column in Table 7 reports the R^2 values when omitting the quadratic terms in the second-order projection solution and the measurement errors. The R^2 values now increase to 88.9% for $\hat{l}_t$, 94.6% for $\hat{w}_t$, 74.1% for Δc_t, and 99.0% for π_t. We do not achieve perfect spanning (i.e., $R^2 = 1$) in this case because consumption habits introduce $\log\left(c_{t-1}/z_{t-1}\right)$ as an endogenous state variable. The last part of Table 7 shows that an additional principal component with $K = 6$ allows us to capture the information in this endogenous state and achieve perfect spanning.

Accordingly, the nonlinear structure of the New Keynesian model implies that the first five principal components of bond yields explain a fairly small proportion of the variation in $\hat{l}_t$ and Δc_t but a much larger proportion of the variation in π_t, which is consistent with U.S. data. For the real wage, the nonlinearities in the model have a somewhat smaller effect, and measurement errors are therefore needed to reduce the R^2 to the level seen in U.S. data. This shows that the proposed model goes a long way in resolving the spanning puzzle. This result may also serve as a more theoretical motivation for imposing the restrictions in Joslin et al. (2014), because these restrictions make their reduced-form DTSM more consistent with an equilibrium model of the type proposed in this paper.

B Real Bond Yields and Real Term Premia

Figure 5 shows selected real bond yields $r^{(k)}_{real,t}$ from $\mathcal{M}^{M,CS}$, with the corresponding U.S. yields provided by Gürkaynak, Sack, and Wright (2010) from 1999. We see that the model gives a remarkable close fit to historical real yields across all maturities, with correlations between the data and the model-implied series of 0.9 or higher. We also note that model-implied real yields are generally slightly below historical yields. This finding is consistent with liquidity risk premia in U.S. real yields, which make these yields higher than implied by a frictionless market (see Abrahams, Adrian, Crump, Moench, and Yu (2016), D'Amico, Kim, and Wei (2018), Andreasen, Christensen, and Riddell (2021), among others). These liquidity premia are generally estimated to be large in the beginning of the 2000s and around the 2007-2009 Great Recession, which may explain why we at these episodes see somewhat larger differences between our model and U.S. real yields.

Table 8 summarizes key unconditional moments linked to real bond yields, with the model-implied moments computed using a simulated sample of 100,000 observations. We find that the mean and standard deviation of real bond yields in $\mathcal{M}^{M,CS}$ match closely the corresponding empirical moments. Note in particularly that $\mathcal{M}^{M,CS}$ has an upward sloping average real yield curve of $1.89\% - 1.13\% = 0.76\%$, which is very similar to the corresponding empirical slope of $1.59\% - 0.78\% = 0.81\%$. As for nominal bond yields, we compute real term premia as $TP^{(k)}_{real,t} = r^{(k)}_{real,t} - \sum_{i=0}^{k-1} \mathbb{E}_t \left[r_{real,t+i} \right]$. The fifth and sixth column in Table 8 show that $\mathcal{M}^{M,CS}$ generates real term premia, which on average are positive and very volatile. For instance, at the 10-year maturity, the real term premium has a mean of 82 basis points and a standard deviation of 57 basis points. Finally, the inflation risk premium is given by the difference between nominal and real term premia, i.e., $TP^{(k)}_t - TP^{(k)}_{real,t}$, and seen to be positive on average and almost as volatile as real term premia.

To summarize, the proposed model provides a close fit to historical real bond yields and their first and second unconditional moments, although none of these moments are included when estimating the New Keynesian model.

Figure 5: Real Bond Yields

This figure shows real bond yields in the U.S. as provided by Gürkaynak et al. (2010) and model-implied real yields from $\mathcal{M}^{M,CS}$ evaluated at the estimated states from the CDKF. The value of τ denotes the correlation between the empirical series and the model-implied series. Shaded areas denote NBER recessions.

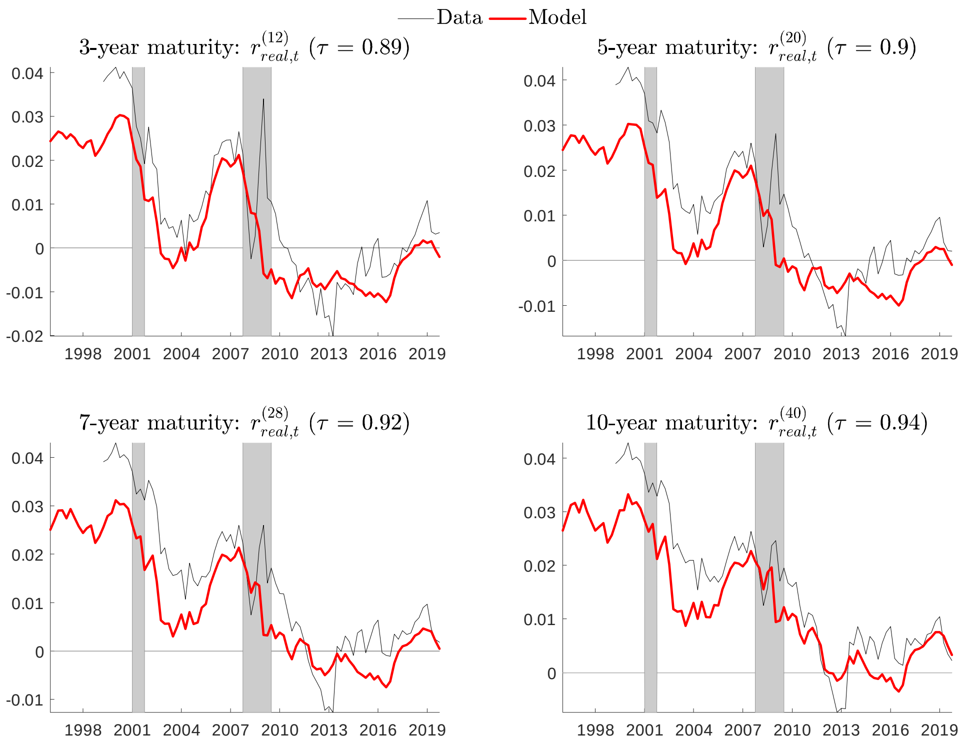


Table 8: Moments Related to Real Bond Yields

This table shows first and second unconditional moments related to real yields in the U.S. from 1999 to 2019 and in the model $\mathcal{M}^{M,CS}$. The model-implied unconditional moments are computed using a simulated series of 100,000 observations. All bond yields are reported in percentages, while all risk premia are in annualized basis points.

	U.S. Data (1999-2019)		Model: $\mathcal{M}^{M,CS}$					
	Real yields		Real yields		Real term premia		Inflation risk premia	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
	1	2	3	4	5	6	7	8
3-year	0.78	1.55	1.13	1.38	7.1	17.9	29.1	17.5
5-year	1.10	1.48	1.33	1.37	26.3	29.4	41.4	27.2
7-year	1.34	1.39	1.54	1.36	47.7	41.0	52.8	35.4
10-year	1.59	1.27	1.89	1.36	82.0	56.8	69.4	44.7

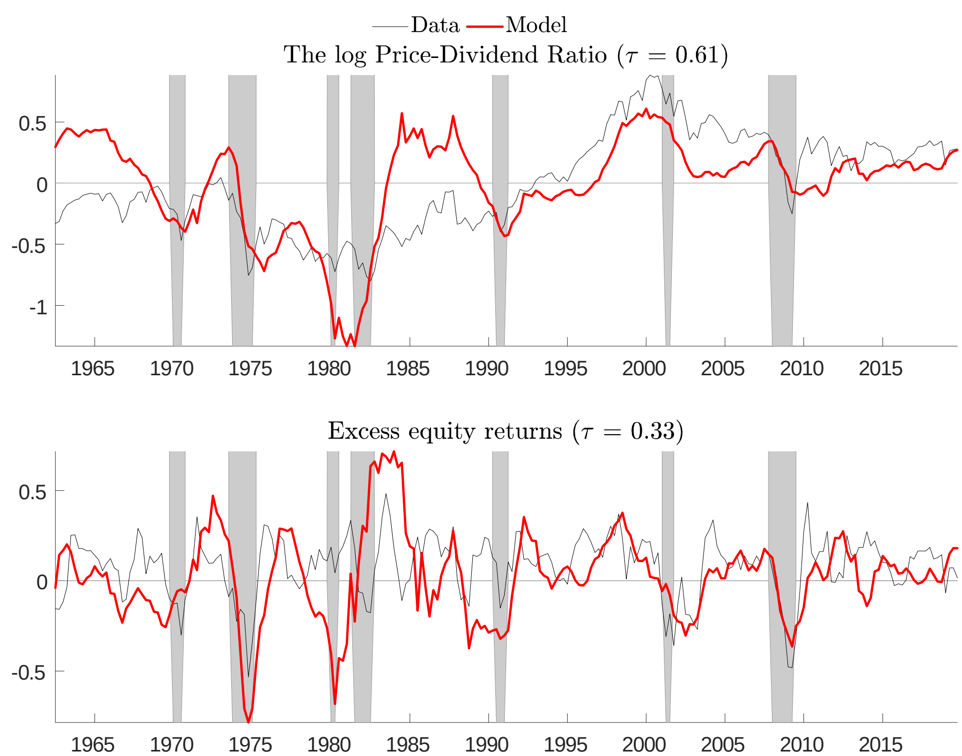
C Equity Returns

We define the equity return r_t^{eq} as a claim on consumption which constitutes dividends. The real equity price P_t^{eq} is then given by $\mathbb{E}_t\left[M^{real}_{t,t+1}e^{r^{eq}_{t+1}}\right] = 1$, where $e^{r^{eq}_{t+1}} = \left(C_{t+1} + P^{eq}_{t+1}\right)/P^{eq}_t$

and $M^{real}_{t,t+1} = M_{t,t+1}\pi_{t+1}$. Figure 6 shows the log-transformed price-dividend ratio pd_t in the model when evaluated at the estimated states and the corresponding series in the U.S. obtained from the Center for Research in Security and Prices (CRSP) with dividends reinvested in the risk-free rate. The unconditional mean of pd_t in $\mathcal{M}^{M,CS}$ is 4.35 and hence somewhat higher than the sample average of 3.61 for pd_t in our U.S. sample, and we therefore plot the demeaned series in the top graph in Figure 6 to facilitate the comparison. We find that $\mathcal{M}^{M,CS}$ fits the overall evolution in the price-dividend ratio fairly well, particularly during the 1970s and after 1990. The main exception relates to the 1980s, where the model either gives a too low or too high value of pd_t. The correlation between the data and the model-implied series is $\tau = 0.61$. For the considered sample, the model implies a standard deviation of 0.39 for pd_t, which is nearly identical to corresponding sample moment of 0.38 in the U.S.

Figure 6: The Price-Dividend Ratio and Excess Equity Returns

The top graph shows the log-transformed price-dividend ratio for the U.S. and in $\mathcal{M}^{M,CS}$. The bottom graph shows the annualized realized excess equity returns in the U.S. and in $\mathcal{M}^{M,CS}$. The model-implied values are evaluated at the estimated states from the CDKF and shaded areas denote NBER recessions.



The bottom graph in Figure 6 considers the realized excess equity returns $r_t^{ex} \equiv r_t^{eq} - r_{real,t-1}^{(1)}$, where we plot annualized returns. We find that $\mathcal{M}^{M,CS}$ provides a reasonable fit to U.S. excess equity returns from CRSP, although the model gives too volatile returns during the 1980s. As a result, the model implies a standard deviation of 23.9% for excess equity returns, whereas the corresponding sample moment is somewhat lower at 16.4%. The overall correlation between the data and the model-implied series is $\tau = 0.33$.

An important aspect of equity returns relate to their predictability. We explore this issue for excess equity returns, where we regress the accumulated excess returns j periods ahead on the current 10-year yield spread $r_t^{(40)} - r_t^{(4)}$ and the price-dividend ratio pd_t following the work of Fama and French (1989). That is, we consider the forecast regression

$$\sum_{i=1}^{j} r_{t+i}^{ex} = \alpha_j + \beta_j \left(r_t^{(40)} - r_t^{(4)} \right) + \delta_j pd_t + \upsilon_{t+j}.$$

Table 9 reports the slope coefficients in this regression at selected forecast horizons, with bootstrapped 95% confidence bands provided in parentheses. To ease the comparison with the model, we scale all variables in this regression to have zero mean and a unit variance. The results show that excess returns in the U.S. are predictable by the 10-year yield spread at the 2- and 3-year horizon and by the price-dividend ratio at the 3-year horizon. We obtain the corresponding model-implied regression loadings in $\mathcal{M}^{M,CS}$ from a simulated sample of 100,000 observations, which we split into samples of the same length $T = 234$ as in the empirical data to account for well-known small-sample biases. The reported results are then the average across these $\lfloor 100,000/T \rfloor$ regressions, where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x. We find that the model generates the desired positive slope coefficients for the yield spread, which increase with the forecast horizon although less strongly when compared to the data. For the price-dividend ratio, we find the desired negative slope coefficients, which closely match the value in the data at all forecast horizons. As a result, the model generates R^2 values of the same magnitude as in empirical data.

To summarize, the proposed New Keynesian model matches several stylized properties of equity prices and equity returns, although these moments are not included when estimating the

Table 9: Equity Return Predictability

This table shows the regression coefficients and the R^2 when regressing future excess equity returns on the slope of the yield curve and the log-transformed price-dividend ratio. For the empirical data moments, these regressions are computed on quarterly U.S. data from 1961 to 2019. Figures in parentheses for the data moments refer the 95 percent confidence bands obtained by resampling the regressant and the regressors jointly in blocks of $2 \times j$ using 10,000 samples of the same length $T = 234$ as in the empirical sample. The corresponding model-implied moments for $\mathcal{M}^{M,CS}$ are computed from a simulated sample of 100,000 observations, which are split into samples of length T. The reported results are the average across these $\lfloor 100,000/T \rfloor$ regressions. All variables in the regression are demeaned and scaled to have a unit standard deviation.

	U.S. Data			Model: $\mathcal{M}^{M,CS}$		
	β_j	δ_j	R^2_j	β_j	δ_j	R^2_j
1-year	0.59 (−0.24,1.51)	−0.56 (−1.50,0.33)	0.04 (0.00,0.21)	0.12	−0.37	0.04
2-year	1.40 (0.16,2.67)	−1.69 (−3.28,0.07)	0.14 (0.01,0.40)	0.35	−1.55	0.12
3-year	2.27 (0.74,3.46)	−2.53 (−3.88,−0.22)	0.25 (0.04,0.55)	0.52	−2.57	0.20

model.

VII Conclusion

This paper addresses a long-standing ambition in the literature by formulating a New Keynesian model to provide a structural explanation for variation in bond yields and their term premia. The model explains bond yields with a low level of news in expected inflation and plausible term premia. This implies that the slope of the yield curve predicts future bond yields, and that risk-adjusted historical bond yields satisfy the expectations hypothesis. The model goes a long way in explaining the spanning puzzle and matches key moments for real bond yields. Finally, the model also captures the overall evolution of the price-dividend ratio and excess equity returns, and it implies that these returns can be forecasted based on slope of the yield curve and the price-dividend ratio.

A The Projection Approximation

Let $\widetilde{c}_t = \frac{c_t}{z_t}$, $\tilde{V}_t = \frac{V_t}{z_t^{1-\chi}}$, $\widetilde{w}_t = \frac{w_t}{z_t}$, and $\widetilde{y}_t = \frac{y_t}{z_t}$. The equilibrium conditions for a scaled version of the New Keynesian model without trending variables are then easily shown to be given by

EQ 0	$evf_t = \mathbb{E}_t\left[\left(\tilde{V}_{t+1}\mu_{z,t+1}^{1-\chi}\right)^{1-\alpha}\right]$
EQ 1	$\tilde{V}_t = u_0 + d_t\left[\frac{1}{1-\chi}\left(\widetilde{c}_t - b\widetilde{c}_{t-1}\mu_{z,t}^{-1}\right)^{1-\chi} + n_t\varphi_0\frac{(1-l_t)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}}\right] + \beta evf_t^{\frac{1}{1-\alpha}}$
EQ 2	$\widetilde{w}_t = \kappa_w\widetilde{w}_{ss} + (1-\kappa_w)\left[n_t\varphi_0(1-l_t)^{-\frac{1}{\varphi}}\left(\widetilde{c}_t - b\widetilde{c}_{t-1}\mu_{z,t}^{-1}\right)^{\chi}\right]$
EQ 3	$\mathbb{E}_t\left[M_{t,t+1}\right]R_t = 1$
EQ 4	$\widetilde{w}_t = mc_t(1-\theta)a_t k_{ss}^{\theta} l_t^{-\theta}$
EQ 5	$\frac{-1}{\eta mc_t}\left(\xi\left(\frac{\pi_t}{\pi_{ss}}-1\right)\frac{\pi_t}{\pi_{ss}} - (1-\eta) - \mathbb{E}_t\left[\xi M_{t,t+1}\pi_{t+1}\left(\frac{\pi_{t+1}}{\pi_{ss}}-1\right)\frac{\pi_{t+1}}{\pi_{ss}}\frac{\widetilde{y}_{t+1}}{\widetilde{y}_t}\mu_{z,t+1}\right]\right) = 1$
EQ 6	$R_t = R_{ss}\mathbb{E}_t\left[\exp\left\{\phi_\pi\log\left(\frac{\pi_{t+1}}{\pi_{ss}\pi_{t+1}^*}\right) + \phi_{\Delta c}\log\left(\frac{\widetilde{c}_{t+1}}{\widetilde{c}_t}\right) + \log\mu_{z,t+1} - \log\mu_{z,ss}\right\}\right]$
EQ 7	$\widetilde{c}_t + \delta k_{ss} = \left(1 - \frac{\xi}{2}\left(\frac{\pi_t}{\pi_{ss}}-1\right)^2\right)\widetilde{y}_t$
EQ 8	$\widetilde{y}_t = a_t k_{ss}^{\theta} l_t^{1-\theta}$

where evf_t is an auxiliary variable, $\Delta c_{ss} = \log\mu_{z,ss}$, and

$$M_{t,t+1} = \beta\left(\frac{\left[\mathbb{E}_t\left[\left(\tilde{V}_{t+1}\mu_{z,t+1}^{(1-\chi)}\right)^{1-\alpha}\right]\right]^{\frac{1}{1-\alpha}}}{\tilde{V}_{t+1}\mu_{z,t+1}^{(1-\chi)}}\right)^{\alpha}\frac{d_{t+1}}{d_t}\frac{\left(\widetilde{c}_{t+1} - b\mu_{z,t}^{-1}\widetilde{c}_t\right)^{-\chi}}{\left(\widetilde{c}_t - b\mu_{z,t-1}^{-1}\widetilde{c}_{t-1}\right)^{-\chi}}\mu_{z,t+1}^{-\chi}\frac{1}{\pi_{t+1}}.$$

We first realize that it is sufficient to solve for $\widetilde{c}_t$, π_t, and evf_t. The reason is that given these variables, EQ 7 implies $\widetilde{y}_t = (\widetilde{c}_t + \delta k_{ss}) / \left(1 - \frac{\xi}{2}\left(\frac{\pi_t}{\pi_{ss}^{\nu}} - 1\right)^2\right)$ and EQ 8 gives $l_t = \left(\frac{\widetilde{y}_t}{a_t k_{ss}^{\theta}}\right)^{\frac{1}{1-\theta}}$. From EQ 2 we directly have $\widetilde{w}_t$, and $mc_t = \frac{\widetilde{w}_t}{(1-\theta)a_t k_{ss}^{\theta} l_t^{-\theta}}$ follows form EQ 4. Finally, EQ 1 gives $\tilde{V}_t$ and EQ 6 gives R_t after evaluating the conditional expectation. Accordingly, the model can be condensly expressed by the three Euler-equations in EQ 0, EQ 3, and EQ 5. We evaluate conditional expectations using the monomial (non-product) integration rule of Judd, Maliar, and Maliar (2011) to obtain the Euler-equation errors $\mathbf{r}(\mathbf{x}_i) = \begin{bmatrix} r^{EQ0}(\mathbf{x}_i) & r^{EQ3}(\mathbf{x}_i) & r^{EQ5}(\mathbf{x}_i)\end{bmatrix}'$ for each point $\mathbf{x}_i$ in the grid $\mathcal{X}$. The Euler-errors for EQ3 and EQ5 are expressed in unit-free terms. A similar scaling is not adopted for EQ 0, because preliminary results showed that the level of evf_t then became badly identified and most often implausible large in the approximation. As for the grid, we express $\mathbf{x}_i$ in deviation from the deterministic steady state. The work of Judd, Maliar, Maliar, and Valero (2014) is used to construct a multidimensional Smolyak grid of order 2 for the hypercube $\prod_{i=1}^{n_x}[\mu_i - 1.5\sigma_i, \mu_i + 1.5\sigma_i]$, where μ_i and σ_i are the unconditional mean and standard deviation, respectively, of the i state variable. For the five shocks, the model is expressed such that $\mu_i = 0$ and we easily obtain σ_i from equation (10). For the last element $\widetilde{c}_{t-1}$ in the state vector, we use a second-order perturbation approximation and the results in Andreasen et al. (2018) to compute its unconditional mean μ_i in deviation from the deterministic steady state, while we use a first-order perturbation approximation to get σ_i. Applying a standard log-transformation to all variables, all loadings related to the three variables in $\mathbf{y}_t^a$ are then obtained by nonlinearly minimizing $\sum_{i=1}^{N} \frac{\mathbf{r}(\mathbf{x}_i)'\mathbf{r}(\mathbf{x}_i)}{\left(1+\mathbf{x}_i'\mathbf{x}_i\right)^2}$, where $N = 85$ is the number of grid points for our configuration. The scaling $1/\left(1+\mathbf{x}_i'\mathbf{x}_i\right)^2$ is introduced to assign less weight to points far from the steady state to ensure that the approximation is very accurate close to the steady state.

Given this solution for $\mathbf{y}_t^a$, we can compute the additional macro variables in $\mathbf{y}^b\left(\mathbf{x}_i\right)$ for all grid points and run the OLS regression

$$\mathbf{y}^b\left(\mathbf{x}_i\right) = \mathbf{g}_0^b + \mathbf{g}_{\mathbf{x}}^b\mathbf{x}_i + \mathbf{g}_{\mathbf{xx}}^b vech\left(\mathbf{x}_i\mathbf{x}_i'\right) + \mathbf{u}_i^b$$

to instantaneously get the loadings $\mathbf{g}_0^b$, $\mathbf{g}_{\mathbf{x}}^b$, and $\mathbf{g}_{\mathbf{xx}}^b$ with dimensions 2×1, $2\times n_x$, and $2\times n_x\left(n_x+1\right)/2$, respectively.

To obtain all bond prices, for $k=1$ we can compute $B_t^{(1)}\left(\mathbf{x}_i\right) = \mathbb{E}_t\left[M_{t,t+1}\left(\mathbf{x}_i\right)\right]$ for all grid points using the solution for $\mathbf{y}_t^a$ and numerical integration. We then compute $b_t^{(1)}\left(\mathbf{x}_i\right) = \log B_t^{(1)}\left(\mathbf{x}_i\right)$ and regress $b_t^{(1)}\left(\mathbf{x}_i\right)$ on a constant, $\mathbf{x}_i$, and $vech\left(\mathbf{x}_i\mathbf{x}_i'\right)$ to get the loadings for $b_t^{(1)}$. For $k=2$, we use the these loadings to compute $B_t^{(2)} = \mathbb{E}_t\left[M_{t,t+1}\left(\mathbf{x}_i\right)B_{t+1}^{(1)}\left(\mathbf{x}_i\right)\right]$ in a similar manner. Then we compute $b_t^{(2)}\left(\mathbf{x}_i\right) = \log B_t^{(2)}\left(\mathbf{x}_i\right)$ and regress $b_t^{(2)}\left(\mathbf{x}_i\right)$ on a constant, $\mathbf{x}_i$, and $vech\left(\mathbf{x}_i\mathbf{x}_i'\right)$ to get the loadings for $b_t^{(2)}$. We continue in this way until we have the entire 10-year yield curve.

Finally, for expected short rates $r^e_{t|k} \equiv \mathbb{E}_t\left[r_{t+k}\right]$, we first compute $r^e_{t|1}\left(\mathbf{x}_i\right) = \mathbb{E}_t\left[-b^{(1)}_{t+1}\left(\mathbf{x}_i\right)\right]$ for all grid points using the solution for $b^{(1)}_t$ and numerical integration. Then we regress $r^e_{t|1}\left(\mathbf{x}_i\right)$ on a constant, $\mathbf{x}_i$, and $vech\left(\mathbf{x}_i\mathbf{x}_i'\right)$ to get the loadings for $r^e_{t|1}$. We then use these loadings and numerical integration to compute $r^e_{t|2}\left(\mathbf{x}_i\right) = \mathbb{E}_t\left[r^e_{t+1|1}\left(\mathbf{x}_i\right)\right]$ for all grid points, and the required loadings are obtained by regressing $r^e_{t|2}\left(\mathbf{x}_i\right)$ on a constant, $\mathbf{x}_i$, and $vech\left(\mathbf{x}_i\mathbf{x}_i'\right)$. We continue in this way until we have all the desired short rate expectations needed to compute term premia at the desired horizon.

References

Abrahams, M.; T. Adrian; R. K. Crump; E. Moench; and R. Yu. Decomposing Real and Nominal Yield Curves. *Journal of Monetary Economics*, 84 (2016), 182–200.

Adrian, T.; R. K. Crump; and E. Moench. Pricing the Term Structure with Linear Regressions. *Journal of Financial Economics*, 110 (2013), 110–138.

Albuquerque, R.; M. Eichenbaum; V. X. Luo; and S. Rebelo. Valuation Risk and Asset Pricing. *Journal of Finance*, 71 (2016), 2861–2904.

Altig, D.; L. J. Christiano; M. Eichenbaum; and J. Linde. Firm-Specific Capital, Nominal Rigidities and the Business Cycle. *Review of Economic Dynamics*, 14 (2011), 225–247.

Andreasen, M. M. How to Maximize The Likelihood Function for a DSGE Model. *Computational Economics*, 35 (2010), 127–154.

Andreasen, M. M. On the Effects of Rare Disasters and Uncertainty Shocks for Risk Premia in Non-Linear DSGE Models. *Review of Economic Dynamics*, 15 (2012), 295–316.

Andreasen, M. M. Non-Linear DSGE Models and the Central Difference Kalman Filter. *Journal of Applied Econometrics*, 28 (2013), 929–955.

Andreasen, M. M.; J. H. E. Christensen; and S. Riddell. The TIPS Liquidity Premium. *Review of Finance*, 25 (2021), 1639–1675.

Andreasen, M. M.; J. Fernandez-Villaverde; and J. F. Rubio-Ramirez. The Pruned State Space System for Non-Linear DSGE Models: Theory and Empirical Applications to Estimation. *Review of Economic Studies*, 85 (2018), 1–49.

Andreasen, M. M., and K. Jørgensen. The Importance of Timing Attitudes in Consumption-Based Asset Pricing Models. *Journal of Monetary Economics*, 111 (2020), 95–117.

Aruoba, S. B.; J. Fernandez-Villaverde; and J. F. Rubio-Ramirez. Comparing Solution Methods for Dynamic Equilibrium Economies. *Journal of Economic Dynamics and Control*, 30 (2006), 2477–2508.

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

Bauer, M. D., and G. D. Rudebusch. Resolving the Spanning Puzzle in Macro-Finance Term Structure Models. *Review of Finance*, 21 (2017), 511–553.

Bekaert, G.; S. Cho; and A. Moreno. New Keynesian Macroeconomics and the Term Structure. *Journal of Money, Credit and Banking*, 42 (2010), 33–62.

Binsbergen, J. H. V.; J. Fernandez-Villaverde; R. S. Koijen; and J. Rubio-Ramirez. The Term Structure of Interest Rates in a DSGE Model with Recursive Preferences. *Journal of Monetary Economics*, 59 (2012), 634–648.

Blanchard, O., and J. Gali. Real Wage Rigidities and the New Keynesian Model. *Journal of Money, Credit and Banking*, 39 (2005), 35–65.

Bretscher, L.; A. Hsu; and A. Tamoni. Fiscal Policy Driven Bond Risk Premia. *Journal of Financial Economics*, 138 (2020), 53–73.

Caldara, D.; J. Fernandez-Villaverde; J. F. Rubio-Ramirez; and W. Yao. Computing DSGE Models with Recursive Preferences and Stochastic Volatility. *Review of Economic Dynamics*, 15 (2012), 188–206.

Campbell, J. Y., and R. J. Shiller. Yield Spread and Interest Rate Movements: A Bird's Eye View. *The Review of Economic Studies*, 58 (1991), 495–514.

Chernozhukov, V., and H. Hong. An MCMC Approach to Classical Estimation. *Journal of Econometrics*, 115 (2003), 293–346.

Christiano, L. J.; M. Trabandt; and K. Walentin. Introducing Financial Frictions and Unemployment into a Small Open Economy Model. *Journal of Economic Dynamic and Control*, 35 (2011), 1999–2041.

Cochrane, J. H. Financial Markets and the Real Economy. *Handbook of the equity risk premium, Chapter 7*, 237–330.

Dai, Q., and K. J. Singleton. Expectation Puzzles, Time-Varying Risk Premia and Affine Models of the Term Structure. *Journal of Financial Economics*, 63 (2002), 415–441.

D'Amico, S.; D. H. Kim; and M. Wei. Tips from TIPS: The Informational Content of Treasury Inflation-Proctected Security Prices. *Journal of Financial and Quantitative Analysis*, 53 (2018), 395–436.

Dew-Becker, I. Bond Pricing with a Time-Varying Price of Risk in an Estimated Medium-Scale Bayesian DSGE Model. *Journal of Money Credit and Banking*, 46 (2014), 837–888.

Duffee, G. R. Expected Inflation and Other Determinants of Treasury Yields. *The Journal of Finance*, 73 (2018), 2139–2180.

Epstein, L. G.; E. Farhi; and T. Strzalecki. How Much Would You Pay to Resolve Long-Run Risk? *American Economic Review*, 104 (2014), 2680–2697.

Epstein, L. G., and S. E. Zin. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica*, 57 (1989), 937–969.

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

Fernández-Villaverde, J., and J. F. Rubio-Ramírez. Estimating Macroeconomic Models: A Likelihood Approach. *Review of Economic Studies*, 74 (2007), 1–46.

Fisher, J. D. M. The Dynamic Effects of Neutral and Investments-Specific Technology Shocks. *Journal of Political Economy*, 114, no. 3 (2006), 413–451.

Gomez-Cram, R., and A. Yaron. How Important are Inflation Expectations for the Nominal Yield Curve? *The Review of Financial Studies*, 34 (2021), 985–1045.

Graeve, F. D.; M. Emiris; and R. Wouters. A Structural Decomposition of the US Yield Curve. *Journal of Monetary Economics*, 56 (2009), 545–559.

Gürkaynak, R.; B. Sack; and J. Wright. The U.S. Treasury Yield Curve: 1961 to the Present. *Journal of Monetary Economics*, 54 (2007), 2291–2304.

Gürkaynak, R. S.; B. Sack; and J. H. Wright. The TIPS Yield Curve and Inflation Compensation. *American Economic Journal: Macroeconomics*, 2 (2010), 70–92.

Hall, R. E. Intertemporal Substitution in Consumption. *Journal of Political Economy*, 96 (1988), 339–357.

Hamilton, J. D. Why You Should Never Use the Hodrick-Prescott Filter. *The Review of Economics and Statistics*, 100 (2018), 831–843.

Hansen, L. P. Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 50 (1982), 1029–1054.

Hordahl, P.; O. Tristani; and D. Vestin. The Yield Curve and Macroeconomic Dynamics. *The Economic Journal*, 118 (2008), 1937–1970.

Joslin, S.; M. A. Priebsch; and K. J. Singleton. Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks. *The Journal of Finance*, 69 (2014), 1197–1233.

Judd, K. L. Projection Methods for Solving Aggregate Growth Models. *Journal of Economic Theory*, 58 (1992), 410–452.

Judd, K. L.; L. Maliar; and S. Maliar. Numerically Stable and Accurate Stochastic Simulation Approaches for Solving Dynamic Economic Models. *Quantitative Economics*, 2 (2011), 173–210.

Judd, K. L.; L. Maliar; S. Maliar; and R. Valero. Smolyak Method for Solving Dynamic Economic Models: Lagrange Interpolation, Anisotropic Grid and Adaptive Domain. *Journal of Economic Dynamics and Control*, 44 (2014), 92–123.

Keane, M. P. Labor Supply and Taxes: A Survey. *Journal of Economic Literature*, 49 (2011), 961–1075.

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

Kung, H. Macroeconomic Linkages between Monetary Policy and the Term Structure of Interest Rates. *Journal of Financial Economics*, 115 (2015), 42–57.

Lagerborg, A.; E. Pappa; and M. O. Ravn. Sentimantal Business Cycles. *Review of Economic Studies*, 90 (2023), 1358–1393.

Levintal, O. Fifth-Order Perturbation Solution to DSGE Models. *Journal of Economics Dynamic and Control*, 80 (2017), 1–16.

Meyer-Gohde, A., and M. Kliem. (Un)expected Monetary Policy Shocks and Term Premia. *Journal of Applied Econometrics*, 37 (2022), 477–499.

Nakata, T., and H. Tanaka. Equilibrium Yield Curves and the Interest Rate Lower Bound. *Finance and Economic Discussion Series, Federal Reserve Board, Washington D. C.*

Norgaard, M.; N. K. Poulsen; and O. Ravn. Advances in Derivative-Free State Estimation for Nonlinear Systems. *Automatica*, 36 (2000), 1627–1638.

Piazzesi, M., and M. Schneider. *Equilibrium Yield Curves, in NBER Macroeconomics Annual 2006*, Vol. 21. Cambridge MA: MIT press (2007).

Rotemberg, J. J. Monopolistic Price Adjustment and Aggregate Output. *Review of Economic Studies*, 49 (1982), 517–531.

Rudebusch, G. D., and E. T. Swanson. Examining the Bond Premium Puzzle with a DSGE Model. *Journal of Monetary Economics*, 55 (2008), 111–126.

Rudebusch, G. D., and E. T. Swanson. The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks. *American Economic Journal: Macroeconomics*, 4 (2012), 1–43.

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

Swanson, E. T. Risk Aversion, Risk Premia, and the Labor Margin with Generalized Recursive Preferences. *Review of Economic Dynamics*, 28 (2018), 290–321.

Weil, P. Nonexpected Utility in Macroeconomics. *Quarterly Journal of Economics*, 105 (1990), 29–42.

Yogo, M. Estimating the Elasticity of Intertemporal Substitution When Instruments Are Weak. *The Review of Economics and Statistics*, 86 (2004), 797–810.

Online Appendix (not for publication): The New Keynesian Model and Bond Yields

Martin M. Andreasen

November 28, 2024

Contents

1 The New Keynesian DSGE Model

1.1 Household

The objective of the representative household is to

$$\max \mathbb{E}_t \sum_{i=0}^{\infty} \beta^i \mathcal{U}(c_{t+i} - h_{t+i}, l_{t+i})$$

subject to the real budget constraint

$$\mathbb{E}_t \left[\frac{M_{t,t+1} X_{t+1}}{P_t} \right] + c_t = \frac{X_t}{P_t} + w_t^* l_t + Div_t + T_t,$$

where c_t is real consumption, h_t is habit stock (taken to be exogenous to the household), l_t is labor supply, w_t^* is the real frictionless wage, P_t is the price level, Div_t is a real dividends from firms, and T_t denotes lump-sum transfers. Letting $x_{t+1}^{real} = X_{t+1}/P_t$ denote the real value of state contingent claims, this can be written as

$$\mathbb{E}_t \left[M_{t,t+1} x_{t+1}^{real} \right] + c_t = \frac{X_t}{P_t} \frac{P_{t-1}}{P_{t-1}} + w_t^* l_t + Div_t + T_t,$$

$\Updownarrow$

$$\mathbb{E}_t \left[M_{t,t+1} x_{t+1}^{real} \right] + c_t = \frac{x_t^{real}}{P_t} \frac{P_{t-1}}{1} + w_t^* l_t + Div_t + T_t,$$

$\Updownarrow$

$$\mathbb{E}_t \left[M_{t,t+1} x_{t+1}^{real} \right] + c_t = \frac{x_t^{real}}{\pi_t} + w_t^* l_t + Div_t + T_t,$$

where $\pi_t = P_{t+1}/P_t$ is the gross inflation rate. Following Rudebusch & Swanson (2012), recursive Epstein-Zin-Weil preferences implies that the value function of the representative household is given by

$$V_t = \mathcal{U}(c_t - h_t, l_t) + \beta \left(\mathbb{E}_t \left[V_{t+1}^{1-\alpha} \right] \right)^{\frac{1}{1-\alpha}},$$

whenever $\mathcal{U}(c_t, l_t) > 0$ (The arguments are similar for the reverse case). The maximization problem can be formulated as a Lagrangean, where the household chooses state-contingent plans for consumption, labor, and asset holdings, i.e. $\left(c_t, l_t, x_{t+1}^{real}\right)$, to maximize V_0 subject to the infinite sequence of state-contingent constraints. I.e.,

$$\mathcal{L} = V_0 + \mathbb{E}_t \sum_{i=0}^{\infty} \beta^i \gamma_{t+i} \left[\mathcal{U}\left(c_{t+i} - h_{t+i}, l_{t+i}\right) + \beta \left(\mathbb{E}_{t+i} \left[V_{t+1+i}^{1-\alpha} \right] \right)^{\frac{1}{1-\alpha}} - V_{t+i} \right]$$

$$+\mathbb{E}_t \sum_{i=0}^{\infty} \beta^i \lambda_{t+i} \left[\frac{x_{t+i}^{real}}{\pi_{t+i}} + l_{t+i} w_{t+i}^* + Div_{t+i} + T_{t+i} - \mathbb{E}_t \left[M_{t+i,t+i+1} x_{t+1+i}^{real} \right] - c_{t+i} \right],$$

where γ_t and λ_t are lagrange multipliers. The first order conditions are then given by

$$\frac{\partial \mathcal{L}}{\partial c_t} = \gamma_t \mathcal{U}_c \left(c_t - h_t, l_t\right) - \lambda_t = 0 \tag{1}$$

$$\frac{\partial \mathcal{L}}{\partial l_t} = \gamma_t \mathcal{U}_l \left(c_t - h_t, l_t\right) + \lambda_t w_t^* = 0 \tag{2}$$

$$\frac{\partial \mathcal{L}}{\partial x_{t+1}^{real}} = \mathbb{P}(s) \left[\beta \lambda_{t+1}(s) \frac{1}{\pi_{t+1}(s)} - \lambda_t M_{t,t+1}(s) \right] = 0 \tag{3}$$

$$\frac{\partial \mathcal{L}}{\partial V_{t+1}} = \gamma_t \beta \mathbb{E}_t \left[V_{t+1}^{1-\alpha} \right]^{\frac{\alpha}{1-\alpha}} \mathbb{P}(s) V_{t+1}^{-\alpha} - \gamma_{t+1} \mathbb{P}(s) \beta = 0 \tag{4}$$

$$\frac{\partial\mathcal{L}}{\partial\lambda_t} = \frac{x_t^{real}}{\pi_t} + l_t w_t + Div_t - \mathbb{E}_t\left[M_{t,t+1}x_{t+1}^{real}\right] - c_t = 0 \quad (5)$$

$$\frac{\partial\mathcal{L}}{\partial\gamma_t} = \mathcal{U}\left(c_t - h_t, l_t\right) + \beta\left(\mathbb{E}_t\left[V_{t+1}^{1-\alpha}\right]\right)^{\frac{1}{1-\alpha}} - V_t = 0, \quad (6)$$

where s indexes the (discrete) states in period $t+1$ and $\mathbb{P}(s)$ denotes the attached probability. Rewriting (1) as

$$\gamma_t\mathcal{U}_c\left(c_t - h_t, l_t\right) = \lambda_t$$

and substituting into (2) yields

$$\gamma_t\mathcal{U}_l\left(c_t - h_t, l_t\right) + \gamma_t\mathcal{U}_c\left(c_t - h_t, l_t\right) w_t^* = 0$$

$\Updownarrow$

$$\frac{1}{w_t^*} = -\frac{\mathcal{U}_c\left(c_t - h_t, l_t\right)}{\mathcal{U}_l\left(c_t - h_t, l_t\right)}, \quad (7)$$

which can be interpreted as an intratemporal optimality condition, as it relates consumption and labor hours within a given period t. Further, rewriting (3) as

$$\beta\frac{\lambda_{t+1}\left(s\right)}{\lambda_t}\frac{1}{\pi_{t+1}\left(s\right)} = M_{t,t+1}\left(s\right).$$

Inserting the expressions for λ_t (and λ_{t+1}) from (1) yields

$$M_{t,t+1} \equiv \beta\frac{\gamma_{t+1}}{\gamma_t}\frac{\mathcal{U}_c\left(c_{t+1} - h_{t+1}, l_{t+1}\right)}{\mathcal{U}_c\left(c_t - h_t, l_t\right)}\frac{1}{\pi_{t+1}}$$

for all states. Now, (4) implies that

$$\left(\mathbb{E}_t\left[V_{t+1}^{1-\alpha}\right]\right)^{\frac{\alpha}{1-\alpha}} V_{t+1}^{-\alpha} = \frac{\gamma_{t+1}}{\gamma_t}.$$

Hence, we get

$$M_{t,t+1} = \beta\left(\frac{\left(\mathbb{E}_t\left[V_{t+1}^{1-\alpha}\right]\right)^{\frac{1}{1-\alpha}}}{V_{t+1}}\right)^{\alpha}\frac{\mathcal{U}_c\left(c_{t+1} - h_{t+1}, l_{t+1}\right)}{\mathcal{U}_c\left(c_t - h_t, l_t\right)}\frac{1}{\pi_{t+1}}. \quad (8)$$

Given the stochastic discount factor, we have

$$\mathbb{E}_t\left[M_{t,t+1}R_t\right] = 1 \quad (9)$$

where R_t is the gross one-period risk-free short rate. From (6) we have that

$$V_t = \mathcal{U}\left(c_t - h_t, l_t\right) + \beta\left(\mathbb{E}_t\left[V_{t+1}^{1-\alpha}\right]\right)^{\frac{1}{1-\alpha}}. \quad (10)$$

1.1.1 Separable utility

The considered utility function is given by

$$\mathcal{U}\left(c_t - h_t, l_t\right) = u_0 z_t^{(1-\chi)(1-\chi_0)} + d_t\left[\frac{1}{1-\chi}\left(\frac{c_t - bc_{t-1}}{\left(\widetilde{c}_{ss}z_t\right)^{\chi_0}}\right)^{1-\chi} + u_0^d \times z_t^{(1-\chi)(1-\chi_0)} + z_t^{(1-\chi)(1-\chi_0)} n_t\varphi_0\frac{\left(1-l_t\right)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}}\right]$$

using the external habits $h_t = bc_{t-1}$. Here, d_t denotes preference shocks, which we specify below, and $\widetilde{c}_{ss}z_t$ is the level of consumption along the balanced growth path. The variable n_t is an exogenous labor supply shock. Hence, for $\chi_0 = 0$, we get

$$\mathcal{U}\left(c_t - h_t, l_t\right) = u_0 z_t^{1-\chi} + d_t\left[\frac{1}{1-\chi}\left(c_t - bc_{t-1}\right)^{1-\chi} + u_0^d z_t^{1-\chi} + z_t^{1-\chi} n_t\varphi_0\frac{\left(1-l_t\right)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}}\right]$$

as in Rudebusch & Swanson (2012), provided $u_0^d, u_0 = 0$, and $b = 0$. For $\chi_0 = 1$, we get

$$\mathcal{U}\left(c_t - h_t, l_t\right) = u_0 + d_t \left[\frac{1}{1-\chi} \left(\frac{c_t - bc_{t-1}}{\widetilde{c}_{ss} z_t} \right)^{1-\chi} + u_0^d + n_t \varphi_0 \frac{(1-l_t)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}} \right]$$

which is very similar to the specification in Andreasen, Fernandez-Villaverde & Rubio-Ramirez (2018). Note that we throughout the present paper impose $\chi_0 = 0$ and $u_0^d = 0$, although both parameters are incorporated in this Online Appendix.

Hence, equation (7) is given by

$$\frac{1}{w_t^*} = -\frac{\mathcal{U}_c\left(c_t - h_t, l_t\right)}{\mathcal{U}_l\left(c_t - h_t, l_t\right)}$$

$\Updownarrow$

$$-\mathcal{U}_l\left(c_t - h_t, l_t\right) = \mathcal{U}_c\left(c_t - h_t, l_t\right) w_t^*$$

$\Updownarrow$

$$z_t^{(1-\chi)(1-\chi_0)} d_t n_t \varphi_0 \left(1 - l_t\right)^{-\frac{1}{\varphi}} = d_t \left(\frac{c_t - bc_{t-1}}{\left(\widetilde{c}_{ss} z_t\right)^{\chi_0}} \right)^{-\chi} \frac{w_t^*}{\left(\widetilde{c}_{ss} z_t\right)^{\chi_0}}$$

$\Updownarrow$

$$w_t^* = z_t^{(1-\chi)(1-\chi_0)} n_t \varphi_0 \left(1 - l_t\right)^{-\frac{1}{\varphi}} \left(\frac{c_t - bc_{t-1}}{\left(\widetilde{c}_{ss} z_t\right)^{\chi_0}} \right)^{\chi} \left(\widetilde{c}_{ss} z_t\right)^{\chi_0}$$

Equation (8) becomes

$$M_{t,t+1} = \beta \left(\frac{\left(\mathbb{E}_t \left[V_{t+1}^{1-\alpha}\right]\right)^{\frac{1}{1-\alpha}}}{V_{t+1}} \right)^{\alpha} \frac{d_{t+1}}{d_t} \frac{\left(\frac{c_{t+1} - bc_t}{\left(\widetilde{c}_{ss} z_{t+1}\right)^{\chi_0}}\right)^{-\chi} z_{t+1}^{-\chi_0}}{\left(\frac{c_t - bc_{t-1}}{\left(\widetilde{c}_{ss} z_t\right)^{\chi_0}}\right)^{-\chi} z_t^{-\chi_0}} \frac{1}{\pi_{t+1}}.$$

Equation (9) is not changed and finally equation (10) is

$$\begin{aligned} V_t &= u_0 z_t^{(1-\chi)(1-\chi_0)} + d_t \left[\frac{1}{1-\chi} \left(\frac{c_t - bc_{t-1}}{\left(\widetilde{c}_{ss} z_t\right)^{\chi_0}} \right)^{1-\chi} + u_0^d \times z_t^{(1-\chi)(1-\chi_0)} + z_t^{(1-\chi)(1-\chi_0)} n_t \varphi_0 \frac{(1-l_t)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}} \right] \\ &\quad + \beta \left(\mathbb{E}_t \left[V_{t+1}^{1-\alpha}\right]\right)^{\frac{1}{1-\alpha}}. \end{aligned}$$

1.2 Firms

Firms are modeled by a two-layer structure. The perfect competitive final good producer bundles together a continuum of intermediates good $y_t(i)$ indexed by $i \in [0, 1]$. Taking prices as given, the final good producer maximizes profits, i.e., solves the problem

$$\max_{y_t(i)} \Pi = P_t y_t - \int_0^1 P_t(i) y_t(i) di,$$

subject to the production function $y_t = \left(\int_0^1 y_t(i)^{\frac{\eta-1}{\eta}} di\right)^{\frac{\eta}{\eta-1}}$, where y_t denotes final output and $\eta > 1$. The Lagrangian for solving this problem reads

$$\mathcal{L} = P_t y_t - \int_0^1 P_t(i) y_t(i) di + \Psi \left(\left(\int_0^1 y_t(i)^{\frac{\eta-1}{\eta}} di \right)^{\frac{\eta}{\eta-1}} - y_t \right).$$

First order conditions are

$$\frac{\partial \mathcal{L}}{\partial y_t(i)} = -P_t(i) + \Psi \frac{\eta}{\eta-1} \left(\int_0^1 y_t(i)^{\frac{\eta-1}{\eta}} di \right)^{\frac{\eta}{\eta-1}-1} \frac{\eta-1}{\eta} y_t(i)^{\frac{\eta-1}{\eta}-1} = 0$$

$$\frac{\partial \mathcal{L}}{\partial \Psi} = \left(\int_0^1 y_t(i)^{\frac{\eta-1}{\eta}} di \right)^{\frac{\eta}{\eta-1}} - y_t = 0.$$

From the first equation,

$$P_t(i) = \Psi \left(\int_0^1 y_t(i)^{\frac{\eta-1}{\eta}} di \right)^{\frac{1}{\eta-1}} y_t(i)^{\frac{-1}{\eta}}$$

$\Updownarrow$

$$\left(\frac{P_t(i)}{\Psi} \right)^{\eta} = \left(\int_0^1 y_t(i)^{\frac{\eta-1}{\eta}} di \right)^{\frac{\eta}{\eta-1}} y_t(i)^{-1} = y_t y_t(i)^{-1}$$

$\Updownarrow$

$$y_t(i) = y_t \left(\frac{P_t(i)}{\Psi} \right)^{-\eta}.$$

Substituting this into the second first order condition gives

$$y_t = \left(\int_0^1 y_t(i)^{\frac{\eta-1}{\eta}} di \right)^{\frac{\eta}{\eta-1}} = \left(\int_0^1 \left(y_t \left(\frac{P_t(i)}{\Psi} \right)^{-\eta} \right)^{\frac{\eta-1}{\eta}} di \right)^{\frac{\eta}{\eta-1}} = y_t \Psi^{\eta} \left(\int_0^1 P_t(i)^{1-\eta} di \right)^{\frac{\eta}{\eta-1}}$$

$\Updownarrow$

$$\Psi = \left(\int_0^1 P_t(i)^{1-\eta} di \right)^{\frac{1}{1-\eta}} \equiv P_t,$$

where P_t is the overall price level. Hence, demand for the i'th intermediate good is

$$y_t(i) = y_t \left(\frac{P_t(i)}{P_t} \right)^{-\eta}.$$

Intermediate goods producers maximize profits, i.e., the dividend transfers to households. We have for the ith intermediate firm that

$$\begin{aligned} Div_t(i) &= \left(\frac{P_t(i)}{P_t} \right) y_t(i) - w_t l_t(i) - \frac{\xi}{2} \left(\frac{P_t(i)}{P_{t-1}(i)} \frac{1}{\pi_{ss}^{\nu}} - 1 \right)^2 y_t - z_t \delta k_{ss} \\ &= \left(\frac{P_t(i)}{P_t} \right)^{1-\eta} y_t - w_t l_t(i) - \frac{\xi}{2} \left(\frac{P_t(i)}{P_{t-1}(i)} \frac{1}{\pi_{ss}^{\nu}} - 1 \right)^2 y_t - z_t \delta k_{ss} \end{aligned}$$

where w_t is the real wage payed by the firms to the workers. Intermediate producers have technology given by $y_t(i) = z_t a_t k_{ss}^{\theta} l_t(i)^{1-\theta}$ available, where z_t and a_t are exogeneous technology processes presented below. Price stickiness is modeled by the Rotemberg scheme. Hence, the ith intermediate firm solves

$$\max_{L_t(i), P_t(i)} \mathbb{E}_t \sum_{k=0}^{\infty} M_{t,t+k}^{\text{real}} \left(\left(\frac{P_{t+k}(i)}{P_{t+k}} \right)^{1-\eta} y_{t+k} - w_{t+k} l_{t+k}(i) - \frac{\xi}{2} \left(\frac{P_{t+k}(i)}{P_{t+k-1}(i)} \frac{1}{\pi_{ss}^{\nu}} - 1 \right)^2 y_{t+k} \right) - z_{t+k} \delta k_{ss},$$

subject to $y_t(i) = y_t \left(\frac{P_t(i)}{P_t} \right)^{-\eta}$ and $y_t(i) = z_t a_t k_{ss}^{\theta} l_t(i)^{1-\theta}$. Here, $M_{t,t+k}^{\text{real}}$ denotes the real stochastic discount factor. The Lagrangian reads

$$\mathcal{L} = \mathbb{E}_t \sum_{k=0}^{\infty} M_{t,t+k}^{\text{real}} \left(\left(\frac{P_t(i)}{P_t} \right)^{1-\eta} y_{t+k} - w_{t+k} l_{t+k}(i) - \frac{\xi}{2} \left(\frac{P_{t+k}(i)}{P_{t+k-1}(i)} \frac{1}{\pi_{ss}^{\nu}} - 1 \right)^2 y_{t+k} \right) - z_{t+k} \delta k_{ss} +$$

$$+\mathbb{E}_t \sum_{k=0}^{\infty} M^{\text{real}}_{t,t+k} mc_{t+k}(i) \left(z_{t+k} a_{t+k} k_{ss}^{\theta} l_{t+k}(i)^{1-\theta} - \left(\frac{P_{t+k}(i)}{P_{t+k}} \right)^{-\eta} y_{t+k} \right)$$

where $mc_t(i)$ is the lagrange multiplier and can be interpreted as the marginal cost of production. The first order condition is

$$\frac{\partial \mathcal{L}}{\partial l_t(i)} = -w_t + mc_t(i)\,(1-\theta)\, z_t a_t k_{ss}^{\theta} l_t(i)^{-\theta} = 0,$$

or simply

$$w_t = mc_t(i)\,(1-\theta)\, z_t a_t k_{ss}^{\theta} l_t(i)^{-\theta}.$$

The first-order condition with respect to $P_t(i)$ is

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial P_t(i)} &= (1-\eta) \left(\frac{P_t(i)}{P_t} \right)^{-\eta} \frac{1}{P_t} y_t - \xi \left(\frac{P_t\,(i)}{P_{t-1}\,(i)} \frac{1}{\pi_{ss}^{\nu}} - 1 \right) \frac{y_t}{P_{t-1}\,(i)} \frac{1}{\pi_{ss}^{\nu}} \\
&\quad - \mathbb{E}_t \left[\xi M^{\text{real}}_{t,t+1} \left(\frac{P_{t+1}\,(i)}{P_t\,(i)} \frac{1}{\pi_{ss}^{\nu}} - 1 \right) \left(\frac{-P_{t+1}\,(i)}{P_t\,(i)^2} \right) \frac{1}{\pi_{ss}^{\nu}} y_{t+1} \right] \\
&\quad + \eta mc_t\,(i) \left(\frac{P_t(i)}{P_t} \right)^{-\eta-1} \frac{y_t}{P_t} \\
&= 0
\end{aligned}$$

$\Updownarrow$

$$\begin{aligned}
&(1-\eta) \left(\frac{P_t(i)}{P_t} \right)^{-\eta} \frac{1}{P_t} y_t \\
&+ \mathbb{E}_t \left[\xi M^{\text{real}}_{t,t+1} \left(\frac{P_{t+1}\,(i)}{P_t\,(i)} \frac{1}{\pi_{ss}^{\nu}} - 1 \right) \frac{P_{t+1}\,(i)}{P_t\,(i)^2} \frac{1}{\pi_{ss}^{\nu}} y_{t+1} \right] \\
&+ \eta mc_t\,(i) \left(\frac{P_t(i)}{P_t} \right)^{-\eta-1} \frac{y_t}{P_t} \\
= \; & \xi \left(\frac{P_t\,(i)}{P_{t-1}\,(i)} \frac{1}{\pi_{ss}^{\nu}} - 1 \right) \frac{y_t}{P_{t-1}\,(i)} \frac{1}{\pi_{ss}^{\nu}}
\end{aligned}$$

All firms are identical, so we can drop the index i. Hence,

$$\begin{aligned}
&(1-\eta) \left(\frac{P_t}{P_t} \right)^{-\eta} \frac{1}{P_t} y_t \\
&+ \mathbb{E}_t \left[\xi M^{\text{real}}_{t,t+1} \left(\frac{P_{t+1}}{P_t} \frac{1}{\pi_{ss}^{\nu}} - 1 \right) \frac{P_{t+1}}{P_t^2} \frac{1}{\pi_{ss}^{\nu}} y_{t+1} \right] \\
&+ \eta mc_t \left(\frac{P_t}{P_t} \right)^{-\eta-1} \frac{y_t}{P_t} \\
= \; & \xi \left(\frac{P_t}{P_{t-1}} \frac{1}{\pi_{ss}^{\nu}} - 1 \right) \frac{y_t}{P_{t-1}} \frac{1}{\pi_{ss}^{\nu}}
\end{aligned}$$

$\Updownarrow$

$$\begin{aligned}
& (1-\eta)\frac{1}{P_t}y_t \\
& +\mathbb{E}_t\left[\xi M^{\text{real}}_{t,t+1}\left(\frac{\pi_{t+1}}{\pi^{\nu}_{ss}}-1\right)\frac{\pi_{t+1}}{P_t}\frac{1}{\pi^{\nu}_{ss}}y_{t+1}\right] \\
& +\eta mc_t\frac{y_t}{P_t} \\
= \quad & \xi\left(\frac{\pi_t}{\pi^{\nu}_{ss}}-1\right)\frac{y_t}{P_{t-1}}\frac{1}{\pi^{\nu}_{ss}}
\end{aligned}$$

$\Updownarrow$

$$\begin{aligned}
& (1-\eta)\,y_t \\
& +\mathbb{E}_t\left[\xi M^{\text{real}}_{t,t+1}\left(\frac{\pi_{t+1}}{\pi^{\nu}_{ss}}-1\right)\frac{\pi_{t+1}}{P_t}\frac{P_t}{\pi^{\nu}_{ss}}y_{t+1}\right] \\
& +\eta mc_t\frac{y_t}{P_t}P_t \\
= \quad & \xi\left(\frac{\pi_t}{\pi^{\nu}_{ss}}-1\right)\frac{y_t}{P_{t-1}}\frac{1}{\pi^{\nu}_{ss}}P_t
\end{aligned}$$

$\Updownarrow$

$$(1-\eta)\,y_t+\mathbb{E}_t\left[\xi M^{\text{real}}_{t,t+1}\left(\frac{\pi_{t+1}}{\pi^{\nu}_{ss}}-1\right)\frac{\pi_{t+1}}{\pi^{\nu}_{ss}}y_{t+1}\right]+\eta mc_t y_t=\xi\left(\frac{\pi_t}{\pi^{\nu}_{ss}}-1\right)y_t\frac{\pi_t}{\pi^{\nu}_{ss}}$$

Or equivalently

$$(1-\eta)\,y_t+\mathbb{E}_t\left[\xi M^{\text{real}}_{t,t+1}\underbrace{\frac{\pi_{t+1}}{\pi_{t+1}}}_{1}\left(\frac{\pi_{t+1}}{\pi^{\nu}_{ss}}-1\right)\frac{\pi_{t+1}}{\pi^{\nu}_{ss}}y_{t+1}\right]+\eta mc_t y_t=\xi\left(\frac{\pi_t}{\pi^{\nu}_{ss}}-1\right)y_t\frac{\pi_t}{\pi^{\nu}_{ss}}$$

$\Updownarrow$

$$(1-\eta)\,y_t+\mathbb{E}_t\left[\xi M_{t,t+1}\pi_{t+1}\left(\frac{\pi_{t+1}}{\pi^{\nu}_{ss}}-1\right)\frac{\pi_{t+1}}{\pi^{\nu}_{ss}}y_{t+1}\right]+\eta mc_t y_t=\xi\left(\frac{\pi_t}{\pi^{\nu}_{ss}}-1\right)y_t\frac{\pi_t}{\pi^{\nu}_{ss}}$$

Note, that without sticky prices, i.e. $\xi=0$, we have $(1-\eta)\,y_t+\eta mc_t y_t=0$, or $mc_t=(\eta-1)/\eta$. Thus, we have that the markup is $1/mc_t$.

1.3 Central Bank

We assume a Taylor rule of the form

$$R_t=R_{ss}\mathbb{E}_t\left[\exp\left\{\phi_\pi\log\left(\frac{\pi_{t+1}}{\pi_{ss}\pi^*_{t+1}}\right)+\phi_{\Delta c}\left(\Delta c_{t+1}-\Delta c_{ss}\right)\right\}\right].$$

1.4 Aggregation and Market Clearing

For the labor market, we follow Blanchard & Gali (2005), Rudebusch & Swanson (2008), among others and introduce a simple wage bargaining friction in the labor market. Specifically, we assume that the real market wage faced by the firms w_t is given by

$$w_t=\kappa_w\left(\widetilde{w}_{ss}z_t\right)+(1-\kappa_w)\,w^*_t,$$

where the parameter $\kappa_w \in [0,1)$ captures the notion of wage stickiness by smoothing the real frictionless wage w_t^* in relation to the long-term equilibrium real wage $\widetilde{w}_{ss} z_t$. Inserting for w_t^* we get

$$w_t = \kappa_w \left(w_{ss} z_t\right) + (1-\kappa_w)\left[z_t^{(1-\chi)(1-\chi_0)} n_t \varphi_0 \left(1-l_t\right)^{-\frac{1}{\varphi}} \left(\frac{c_t - b c_{t-1}}{\left(\widetilde{c}_{ss} z_t\right)^{\chi_0}}\right)^{\chi} \left(\widetilde{c}_{ss} z_t\right)^{\chi_0}\right]$$

Although this simple rule does not explicitly introduce Nash bargaining between workers and firms, Blanchard & Gali (2005) argue that it is a simple way to capture the essential features of real wage bargaining.

From the household budget constraint, workers receive the frictionless wage w_t^* and not the actual wage w_t paid by the firms. To eliminate any resource costs linked to wage stickiness, we let the real transfers to the household be

$$\begin{aligned} T_t &= l_t w_t - l w_t^* \\ &= l_t \left(w_t - w_t^*\right) \\ &= l_t \left(\kappa_w \left(\widetilde{w}_{ss} z_t\right) + (1-\kappa_w) w_t^* - w_t^*\right) \\ &= l_t \left(\kappa_w \left(\widetilde{w}_{ss} z_t\right) - \kappa_w w_t^*\right) \\ &= \kappa_w l_t \left(\left(\widetilde{w}_{ss} z_t\right) - w_t^*\right), \end{aligned}$$

to ensure that the wage bill $w_t l_t$ paid by the firm is also the wage bill received by the household. From the budget restriction we have

$$\mathbb{E}_t\left[M_{t,t+1} x_{t+1}^{real}\right] + c_t = \frac{x_t^{real}}{\pi_t} + w_t^* l_t + Div_t + T_t$$

$\Downarrow$

$$0 + c_t = 0 + w_t^* l_t + Div_t + l_t w_t - l w_t^*$$

$\Updownarrow$

$$Div_t = c_t - w_t l_t$$

given that the amount of state contigent claims x_t are in zero net supply. The expression for dividends is

$$Div_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{1-\eta} y_t - w_t l_t(i) - \frac{\xi}{2}\left(\frac{P_t(i)}{P_{t-1}(i)}\frac{1}{\pi_{ss}^{\nu}} - 1\right)^2 y_t - z_t \delta k_{ss}$$

$\Downarrow$

$$\begin{aligned} Div_t &= y_t - w_t l_t - \frac{\xi}{2}\left(\frac{\pi_t}{\pi_{ss}^{\nu}} - 1\right)^2 y_t - z_t \delta k_{ss} \\ &= \left(1 - \frac{\xi}{2}\left(\frac{\pi_t}{\pi_{ss}^{\nu}} - 1\right)^2\right) y_t - w_t l_t - z_t \delta k_{ss} \end{aligned}$$

Combining the two expressions for dividends, we get

$$c_t - w_t l_t = \left(1 - \frac{\xi}{2}\left(\frac{\pi_t}{\pi_{ss}^{\nu}} - 1\right)^2\right) y_t - w_t l_t - z_t \delta k_{ss}$$

$\Updownarrow$

$$c_t + z_t \delta k_{ss} = \left(1 - \frac{\xi}{2}\left(\frac{\pi_t}{\pi_{ss}^{\nu}} - 1\right)^2\right) y_t.$$

1.5 Exogeneous Shocks

The exogenous shocks are given by

$$\log\left(\frac{\mu_{z,t+1}}{\mu_{z,ss}}\right)=\rho_{\mu_z}\log\left(\frac{\mu_{z,t}}{\mu_{z,ss}}\right)+\sigma_{\mu_z}\epsilon_{\mu_z,t+1}$$

$$\log d_{t+1}=\rho_d\log d_t+\sigma_d\epsilon_{d,t+1}$$

$$\log n_{t+1}=\rho_n\log n_t+\sigma_n\epsilon_{n,t+1}$$

$$\log\pi^*_{t+1}=\rho_{\pi^*}\log\pi^*_t+\sigma_{\pi^*}\epsilon_{\pi^*,t+1}$$

$$\log a_{t+1}=\rho_a\log a_t+\sigma_a\epsilon_{a,t+1}$$

1.6 Model Equations

The equations for the baseline implementation are summarized in the table below.

	Household
1	$V_t=u_0z_t^{(1-\chi)(1-\chi_0)}+d_t\left[\frac{1}{1-\chi}\left(\frac{c_t-bc_{t-1}}{(\widetilde{c}_{ss}z_t)^{\chi_0}}\right)^{1-\chi}+u_0^d\times z_t^{(1-\chi)(1-\chi_0)}+z_t^{(1-\chi)(1-\chi_0)}n_t\varphi_0\frac{(1-l_t)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}}\right]+\beta\left(\mathbb{E}_t\left[V_{t+1}^{1-\alpha}\right]\right)^{\frac{1}{1-\alpha}}$
2	$w_t=\kappa_w\left(\widetilde{w}_{ss}z_t\right)+(1-\kappa_w)\left[z_t^{(1-\chi)(1-\chi_0)}n_t\varphi_0\left(1-l_t\right)^{-\frac{1}{\varphi}}\left(\frac{c_t-bc_{t-1}}{(\widetilde{c}_{ss}z_t)^{\chi_0}}\right)^{\chi}\left(\widetilde{c}_{ss}z_t\right)^{\chi_0}\right]$
3	$\mathbb{E}_t\left[M_{t,t+1}R_t\right]=1$
	$M_{t,t+1}=\beta\left(\frac{\left[\mathbb{E}_t\left[V_{t+1}^{1-\alpha}\right]\right]^{\frac{1}{1-\alpha}}}{V_{t+1}}\right)^{\alpha}\frac{d_{t+1}}{d_t}\frac{\left(\frac{c_{t+1}-bc_t}{(\widetilde{c}_{ss}z_{t+1})^{\chi_0}}\right)^{-\chi}z_{t+1}^{-\chi_0}}{\left(\frac{c_t-bc_{t-1}}{(\widetilde{c}_{ss}z_t)^{\chi_0}}\right)^{-\chi}z_t^{-\chi_0}}\frac{1}{\pi_{t+1}}$
	Firm
4	$w_t=mc_t\left(1-\theta\right)z_ta_tk_{ss}^{\theta}l_t^{-\theta}$
5	$(1-\eta)\,y_t+\mathbb{E}_t\left[\xi M_{t,t+1}^{\text{real}}\left(\frac{\pi_{t+1}}{\pi_{ss}^{\nu}}-1\right)\frac{\pi_{t+1}}{\pi_{ss}^{\nu}}y_{t+1}\right]+\eta mc_ty_t=\xi\left(\frac{\pi_t}{\pi_{ss}^{\nu}}-1\right)y_t\frac{\pi_t}{\pi_{ss}^{\nu}}$
	Central bank
6	$R_t=R_{ss}\mathbb{E}_t\left[\exp\left\{\phi_\pi\log\left(\frac{\pi_{t+1}}{\pi_{ss}\pi^*_{t+1}}\right)+\phi_{\Delta c}\left(\Delta c_{t+1}-\Delta c_{ss}\right)\right\}\right]$
	Aggregation
7	$c_t+z_t\delta k_{ss}=\left(1-\frac{\xi}{2}\left(\frac{\pi_t}{\pi_{ss}^{\nu}}-1\right)^2\right)y_t$
8	$y_t=z_ta_tk_{ss}^{\theta}l_t^{1-\theta}$
	Links
9	$\left(c_{t-1}\right)_{t+1}=c_t$
10-	**Shocks**

To these equations, we add the recursive equations for bond prices, i.e. $B_t^{(k)}=\mathbb{E}_t\left[M_{t,t+1}B_{t+1}^{(k-1)}\right]$.

1.7 Detrending

1.7.1 Separable case

Define the stationary variables $\widetilde{c}_t=\frac{c_t}{z_t}$, $\tilde{V}_t=\frac{V_t}{z_t^{(1-\chi)(1-\chi_0)}}$, $\widetilde{w}_t=\frac{w_t}{z_t}$, $\widetilde{y}_t=\frac{y_t}{z_t}$

EQ 1

$$V_t=u_0z_t^{(1-\chi)(1-\chi_0)}+d_t\left[\frac{1}{1-\chi}\left(\frac{c_t-bc_{t-1}}{(\widetilde{c}_{ss}z_t)^{\chi_0}}\right)^{1-\chi}+u_0^d\times z_t^{(1-\chi)(1-\chi_0)}+z_t^{(1-\chi)(1-\chi_0)}n_t\varphi_0\frac{(1-l_t)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}}\right]+\beta\left(\mathbb{E}_t\left[V_{t+1}^{1-\alpha}\right]\right)^{\frac{1}{1-\alpha}}$$

$\Updownarrow$

$$\frac{V_t}{z_t^{(1-\chi)(1-\chi_0)}} = u_0 + d_t \left[\frac{1}{1-\chi} \left(\frac{c_t - bc_{t-1}}{\widetilde{c}_{ss}^{\chi_0} z_t^{\chi_0}} \right)^{1-\chi} \frac{1}{z_t^{(1-\chi)(1-\chi_0)}} + u_0^d + n_t \varphi_0 \frac{(1-l_t)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}} \right] + \frac{1}{z_t^{(1-\chi)(1-\chi_0)}} \beta \left(\mathbb{E}_t \left[V_{t+1}^{1-\alpha} \right] \right)^{\frac{1}{1-\alpha}}$$

$\Updownarrow$

$$\frac{V_t}{z_t^{(1-\chi)(1-\chi_0)}} = u_0 + d_t \left[\frac{1}{1-\chi} \left(\frac{c_t - bc_{t-1}}{\widetilde{c}_{ss}^{\chi_0} z_t^{\chi_0} z_t^{1-\chi_0}} \right)^{1-\chi} + u_0^d + n_t \varphi_0 \frac{(1-l_t)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}} \right] + \frac{1}{z_t^{(1-\chi)(1-\chi_0)}} \beta \left(\mathbb{E}_t \left[V_{t+1}^{1-\alpha} \right] \right)^{\frac{1}{1-\alpha}}$$

$\Updownarrow$

$$\frac{V_t}{z_t^{(1-\chi)(1-\chi_0)}} = u_0 + d_t \left[\frac{1}{1-\chi} \left(\frac{c_t - bc_{t-1}}{\widetilde{c}_{ss}^{\chi_0} z_t} \right)^{1-\chi} + u_0^d + n_t \varphi_0 \frac{(1-l_t)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}} \right] + \frac{1}{z_t^{(1-\chi)(1-\chi_0)}} \beta \left(\mathbb{E}_t \left[V_{t+1}^{1-\alpha} \right] \right)^{\frac{1}{1-\alpha}}$$

$\Updownarrow$

$$\tilde{V}_t = u_0 + d_t \left[\frac{1}{1-\chi} \left(\frac{\widetilde{c}_t - b\widetilde{c}_{t-1}\mu_{z,t}^{-1}}{\widetilde{c}_{ss}^{\chi_0}} \right)^{1-\chi} + u_0^d + n_t \varphi_0 \frac{(1-l_t)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}} \right] + \beta \left(\mathbb{E}_t \left[\left(\frac{V_{t+1}}{z_{t+1}^{(1-\chi)(1-\chi_0)}} \frac{z_{t+1}^{(1-\chi)(1-\chi_0)}}{z_t^{(1-\chi)(1-\chi_0)}} \right)^{1-\alpha} \right] \right)^{\frac{1}{1-\alpha}}$$

$\Updownarrow$

$$\tilde{V}_t = u_0 + d_t \left[\frac{1}{1-\chi} \left(\frac{\widetilde{c}_t - b\widetilde{c}_{t-1}\mu_{z,t}^{-1}}{\widetilde{c}_{ss}^{\chi_0}} \right)^{1-\chi} + u_0^d + n_t \varphi_0 \frac{(1-l_t)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}} \right] + \beta \left(\mathbb{E}_t \left[\left(\tilde{V}_{t+1} \mu_{z,t+1}^{(1-\chi)(1-\chi_0)} \right)^{1-\alpha} \right] \right)^{\frac{1}{1-\alpha}}$$

EQ 2

$$w_t = \kappa_w \left(\widetilde{w}_{ss} z_t \right) + (1-\kappa_w) \left[z_t^{(1-\chi)(1-\chi_0)} n_t \varphi_0 (1-l_t)^{-\frac{1}{\varphi}} \left(\frac{c_t - bc_{t-1}}{(\widetilde{c}_{ss} z_t)^{\chi_0}} \right)^{\chi} (\widetilde{c}_{ss} z_t)^{\chi_0} \right]$$

$\Updownarrow$

$$\frac{w_t}{z_t} = \kappa_w \widetilde{w}_{ss} + (1-\kappa_w) \left[z_t^{(1-\chi)(1-\chi_0)} n_t \varphi_0 (1-l_t)^{-\frac{1}{\varphi}} \left(\frac{c_t - bc_{t-1}}{(\widetilde{c}_{ss} z_t)^{\chi_0}} \right)^{\chi} (\widetilde{c}_{ss} z_t)^{\chi_0} \frac{1}{z_t} \right]$$

$\Updownarrow$

$$\frac{w_t}{z_t} = \kappa_w \widetilde{w}_{ss} + (1-\kappa_w) \left[z_t^{1-\chi_0-\chi+\chi\chi_0-1+\chi_0-\chi\chi_0} n_t \varphi_0 (1-l_t)^{-\frac{1}{\varphi}} \left(\frac{c_t - bc_{t-1}}{(\widetilde{c}_{ss})^{\chi_0}} \right)^{\chi} (\widetilde{c}_{ss})^{\chi_0} \right]$$

$\Updownarrow$

$$\widetilde{w}_t = \kappa_w \widetilde{w}_{ss} + (1-\kappa_w) \left[z_t^{-\chi} n_t \varphi_0 (1-l_t)^{-\frac{1}{\varphi}} \left(\frac{c_t - bc_{t-1}}{(\widetilde{c}_{ss})^{\chi_0}} \right)^{\chi} (\widetilde{c}_{ss})^{\chi_0} \right]$$

$\Updownarrow$

$$\widetilde{w}_t = \kappa_w \widetilde{w}_{ss} + (1-\kappa_w) \left[n_t \varphi_0 (1-l_t)^{-\frac{1}{\varphi}} \left(\frac{\frac{c_t}{z_t} - b\frac{z_{t-1}}{z_t}\frac{c_{t-1}}{z_{t-1}}}{(\widetilde{c}_{ss})^{\chi_0}} \right)^{\chi} (\widetilde{c}_{ss})^{\chi_0} \right]$$

$\Updownarrow$

$$\widetilde{w}_t = \kappa_w \widetilde{w}_{ss} + (1-\kappa_w) \left[n_t \varphi_0 (1-l_t)^{-\frac{1}{\varphi}} \left(\frac{\frac{c_t}{z_t} - b\frac{z_{t-1}}{z_t}\frac{c_{t-1}}{z_{t-1}}}{(\widetilde{c}_{ss})^{\chi_0}} \right)^{\chi} (\widetilde{c}_{ss})^{\chi_0} \right]$$

$\Updownarrow$

$$\widetilde{w}_t = \kappa_w \widetilde{w}_{ss} + (1-\kappa_w) \left[n_t \varphi_0 (1-l_t)^{-\frac{1}{\varphi}} \left(\frac{\widetilde{c}_t - b\widetilde{c}_{t-1}\mu_{z,t}^{-1}}{(\widetilde{c}_{ss})^{\chi_0}} \right)^{\chi} (\widetilde{c}_{ss})^{\chi_0} \right]$$

EQ 3

$\mathbb{E}_t \left[M_{t,t+1} R_t \right] = 1$

where

$$M_{t,t+1} = \beta \left(\frac{\left[\mathbb{E}_t \left[V_{t+1}^{1-\alpha} \right] \right]^{\frac{1}{1-\alpha}}}{V_{t+1}} \right)^{\alpha} \frac{d_{t+1}}{d_t} \frac{\left(\frac{c_{t+1} - bc_t}{\widetilde{c}_{ss}^{\chi_0} z_{t+1}^{\chi_0}} \right)^{-\chi} z_{t+1}^{-\chi_0}}{\left(\frac{c_t - bc_{t-1}}{\widetilde{c}_{ss}^{\chi_0} z_t^{\chi_0}} \right)^{-\chi} z_t^{-\chi_0}} \frac{1}{\pi_{t+1}}$$

$\Updownarrow$

$$M_{t,t+1} = \beta \left(\frac{\left[\mathbb{E}_t \left[\left(V_{t+1} \frac{z_{t+1}^{(1-\chi)(1-\chi_0)}}{z_{t+1}^{(1-\chi)(1-\chi_0)}} \right)^{1-\alpha} \right] \right]^{\frac{1}{1-\alpha}}}{V_{t+1} \frac{z_{t+1}^{(1-\chi)(1-\chi_0)}}{z_{t+1}^{(1-\chi)(1-\chi_0)}}} \right)^{\alpha} \frac{d_{t+1}}{d_t} \frac{\left(\frac{c_{t+1} - bc_t}{z_{t+1}^{\chi_0}} \frac{z_{t+1}^{1-\chi_0}}{z_{t+1}^{1-\chi_0}} \right)^{-\chi}}{\left(\frac{c_t - bc_{t-1}}{z_t^{\chi_0}} \frac{z_t^{1-\chi_0}}{z_t^{1-\chi_0}} \right)^{-\chi}} \left(\frac{z_{t+1}}{z_t} \right)^{-\chi_0} \frac{1}{\pi_{t+1}}$$

$\Updownarrow$

$$M_{t,t+1}=\beta\left(\frac{\left[\mathbb{E}_t\left[\left(\tilde{V}_{t+1}z_{t+1}^{(1-\chi)(1-\chi_0)}\right)^{1-\alpha}\right]\right]^{\frac{1}{1-\alpha}}}{\tilde{V}_{t+1}z_{t+1}^{(1-\chi)(1-\chi_0)}}\right)^{\alpha}\frac{d_{t+1}}{d_t}\frac{\left(\frac{c_{t+1}-bc_t}{z_{t+1}}\frac{z_{t+1}^{1-\chi_0}}{1}\right)^{-\chi}}{\left(\frac{c_t-bc_{t-1}}{z_t}\frac{z_t^{1-\chi_0}}{1}\right)^{-\chi}}\left(\frac{z_{t+1}}{z_t}\right)^{-\chi_0}\frac{1}{\pi_{t+1}}$$

$\Updownarrow$

$$M_{t,t+1}=\beta\left(\frac{\left[\mathbb{E}_t\left[\left(\tilde{V}_{t+1}\frac{z_{t+1}^{(1-\chi)(1-\chi_0)}}{z_t^{(1-\chi)(1-\chi_0)}}\right)^{1-\alpha}\right]\right]^{\frac{1}{1-\alpha}}}{\tilde{V}_{t+1}\frac{z_{t+1}^{(1-\chi)(1-\chi_0)}}{z_t^{(1-\chi)(1-\chi_0)}}}\right)^{\alpha}\frac{d_{t+1}}{d_t}\frac{\left(\widetilde{c}_{t+1}-b\mu_{z,t}^{-1}\widetilde{c}_t\right)^{-\chi}}{\left(\widetilde{c}_t-b\mu_{z,t-1}^{-1}\widetilde{c}_{t-1}\right)^{-\chi}}\frac{z_{t+1}^{-\chi(1-\chi_0)}}{z_t^{-\chi(1-\chi_0)}}\left(\frac{z_{t+1}}{z_t}\right)^{-\chi_0}\frac{1}{\pi_{t+1}}$$

$\Updownarrow$

$$M_{t,t+1}=\beta\left(\frac{\left[\mathbb{E}_t\left[\left(\tilde{V}_{t+1}\mu_{z,t+1}^{(1-\chi)(1-\chi_0)}\right)^{1-\alpha}\right]\right]^{\frac{1}{1-\alpha}}}{\tilde{V}_{t+1}\mu_{z,t+1}^{(1-\chi)(1-\chi_0)}}\right)^{\alpha}\frac{d_{t+1}}{d_t}\frac{\left(\widetilde{c}_{t+1}-b\mu_{z,t}^{-1}\widetilde{c}_t\right)^{-\chi}}{\left(\widetilde{c}_t-b\mu_{z,t-1}^{-1}\widetilde{c}_{t-1}\right)^{-\chi}}\mu_{z,t+1}^{-\chi(1-\chi_0)}\mu_{z,t+1}^{-\chi_0}\frac{1}{\pi_{t+1}}$$

$\Updownarrow$

$$M_{t,t+1}=\beta\left(\frac{\left[\mathbb{E}_t\left[\left(\tilde{V}_{t+1}\mu_{z,t+1}^{(1-\chi)(1-\chi_0)}\right)^{1-\alpha}\right]\right]^{\frac{1}{1-\alpha}}}{\tilde{V}_{t+1}\mu_{z,t+1}^{(1-\chi)(1-\chi_0)}}\right)^{\alpha}\frac{d_{t+1}}{d_t}\frac{\left(\widetilde{c}_{t+1}-b\mu_{z,t}^{-1}\widetilde{c}_t\right)^{-\chi}}{\left(\widetilde{c}_t-b\mu_{z,t-1}^{-1}\widetilde{c}_{t-1}\right)^{-\chi}}\mu_{z,t+1}^{-\chi(1-\chi_0)-\chi_0}\frac{1}{\pi_{t+1}}$$

EQ 4

$w_t = mc_t\left(1-\theta\right)z_t a_t k_{ss}^{\theta} l_t{}^{-\theta}$

$\Updownarrow$

$\frac{w_t}{z_t} = mc_t\left(1-\theta\right)a_t k_{ss}^{\theta} l_t^{-\theta}$

$\Updownarrow$

$\widetilde{w}_t = mc_t\left(1-\theta\right)a_t k_{ss}^{\theta} l_t^{-\theta}$

EQ 5

$\left(1-\eta\right)y_t+\mathbb{E}_t\left[\xi M_{t,t+1}^{\text{real}}\left(\frac{\pi_{t+1}}{\pi_{ss}^{\nu}}-1\right)\frac{\pi_{t+1}}{\pi_{ss}^{\nu}}y_{t+1}\right]+\eta mc_t y_t=\xi\left(\frac{\pi_t}{\pi_{ss}^{\nu}}-1\right)y_t\frac{\pi_t}{\pi_{ss}^{\nu}}$

$\Updownarrow$

$\left(1-\eta\right)\frac{y_t}{z_t}+\mathbb{E}_t\left[\xi M_{t,t+1}^{\text{real}}\left(\frac{\pi_{t+1}}{\pi_{ss}^{\nu}}-1\right)\frac{\pi_{t+1}}{\pi_{ss}^{\nu}}\frac{y_{t+1}}{z_{t+1}}\frac{z_{t+1}}{z_t}\right]+\eta mc_t\frac{y_t}{z_t}=\xi\left(\frac{\pi_t}{\pi_{ss}^{\nu}}-1\right)\frac{y_t}{z_t}\frac{\pi_t}{\pi_{ss}^{\nu}}$

$\Updownarrow$

$\left(1-\eta\right)\widetilde{y}_t+\mathbb{E}_t\left[\xi M_{t,t+1}^{\text{real}}\left(\frac{\pi_{t+1}}{\pi_{ss}^{\nu}}-1\right)\frac{\pi_{t+1}}{\pi_{ss}^{\nu}}\widetilde{y}_{t+1}\mu_{z,t+1}\right]+\eta mc_t\widetilde{y}_t=\xi\left(\frac{\pi_t}{\pi_{ss}^{\nu}}-1\right)\widetilde{y}_t\frac{\pi_t}{\pi_{ss}^{\nu}}$

or as in the Appendix

$\left(1-\eta\right)+\mathbb{E}_t\left[\xi M_{t,t+1}\pi_{t+1}\left(\frac{\pi_{t+1}}{\pi_{ss}^{\nu}}-1\right)\frac{\pi_{t+1}}{\pi_{ss}^{\nu}}\frac{\widetilde{y}_{t+1}}{\widetilde{y}_t}\mu_{z,t+1}\right]+\eta mc_t=\xi\left(\frac{\pi_t}{\pi_{ss}^{\nu}}-1\right)\frac{\pi_t}{\pi_{ss}^{\nu}}$

$\Updownarrow$

$\eta mc_t=\xi\left(\frac{\pi_t}{\pi_{ss}^{\nu}}-1\right)\frac{\pi_t}{\pi_{ss}^{\nu}}-\left(1-\eta\right)-\mathbb{E}_t\left[\xi M_{t,t+1}\pi_{t+1}\left(\frac{\pi_{t+1}}{\pi_{ss}^{\nu}}-1\right)\frac{\pi_{t+1}}{\pi_{ss}^{\nu}}\frac{\widetilde{y}_{t+1}}{\widetilde{y}_t}\mu_{z,t+1}\right]$

$\Updownarrow$

$1 = \frac{1}{\eta mc_t} \left\{ \xi \left(\frac{\pi_t}{\pi_{ss}^\nu} - 1 \right) \frac{\pi_t}{\pi_{ss}^\nu} - (1-\eta) - \mathbb{E}_t \left[\xi M_{t,t+1} \pi_{t+1} \left(\frac{\pi_{t+1}}{\pi_{ss}^\nu} - 1 \right) \frac{\pi_{t+1}}{\pi_{ss}^\nu} \frac{\widetilde{y}_{t+1}}{\widetilde{y}_t} \mu_{z,t+1} \right] \right\}$

EQ 6
$R_t = R_{ss} \mathbb{E}_t \left[\exp \left\{ \phi_\pi \log \left(\frac{\pi_{t+1}}{\pi_{ss} \pi_{t+1}^*} \right) + \phi_{\Delta c} \left(\Delta c_{t+1} - \Delta c_{ss} \right) \right\} \right]$
Note that

$$\begin{aligned} \Delta c_t &= \log \left(c_t / c_{t-1} \right) \\ &= \log \left(\frac{\widetilde{c}_t z_t}{\widetilde{c}_{t-1} z_{t-1}} \right) \\ &= \log \left(\frac{\widetilde{c}_t}{\widetilde{c}_{t-1}} \right) + \log \mu_{z,t} \end{aligned}$$

so no transformation of Δc_t is needed. Also

$$\Delta c_{t+1} = \log \left(\frac{\widetilde{c}_{t+1}}{\widetilde{c}_t} \right) + \log \mu_{z,t+1}$$

So we get

$$R_t = R_{ss} \mathbb{E}_t \left[\exp \left\{ \phi_\pi \log \left(\frac{\pi_{t+1}}{\pi_{ss} \pi_{t+1}^*} \right) + \phi_{\Delta c} \left(\log \left(\frac{\widetilde{c}_{t+1}}{\widetilde{c}_t} \right) + \log \mu_{z,t+1} - \log \mu_{z,ss} \right) \right\} \right]$$

as $\Delta c_{ss} = \log \mu_{z,ss}$.

EQ 7
$c_t + z_t \delta k_{ss} = \left(1 - \frac{\xi}{2} \left(\frac{\pi_t}{\pi_{ss}^\nu} - 1 \right)^2 \right) y_t$
$\Updownarrow$
$\widetilde{c}_t + \delta k_{ss} = \left(1 - \frac{\xi}{2} \left(\frac{\pi_t}{\pi_{ss}^\nu} - 1 \right)^2 \right) \widetilde{y}_t$

EQ 8
$y_t = z_t a_t k_{ss}^\theta l_t^{1-\theta}$
$\Updownarrow$
$\widetilde{y}_t = a_t k_{ss}^\theta l_t^{1-\theta}$

The remaining equations in the model do not need to be detrended.

	Household
1	$\tilde{V}_t = u_0 + d_t \left[\frac{1}{1-\chi} \left(\frac{\widetilde{c}_t - b\widetilde{c}_{t-1}\mu_{z,t}^{-1}}{\widetilde{c}_{ss}^{\chi_0}} \right)^{1-\chi} + u_0^d + n_t\varphi_0 \frac{(1-l_t)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}} \right] + \beta \left(\mathbb{E}_t \left[\left(\tilde{V}_{t+1}\mu_{z,t+1}^{(1-\chi)(1-\chi_0)} \right)^{1-\alpha} \right] \right)^{\frac{1}{1-\alpha}}$
2	$\widetilde{w}_t = \kappa_w \widetilde{w}_{ss} + (1-\kappa_w) \left[n_t\varphi_0 (1-l_t)^{-\frac{1}{\varphi}} \left(\frac{\widetilde{c}_t - b\widetilde{c}_{t-1}\mu_{z,t}^{-1}}{(\widetilde{c}_{ss})^{\chi_0}} \right)^{\chi} (\widetilde{c}_{ss})^{\chi_0} \right]$
3	$\mathbb{E}_t [M_{t,t+1}R_t] = 1$
	$M_{t,t+1} = \beta \left(\frac{\left[\mathbb{E}_t \left[\left(\tilde{V}_{t+1}\mu_{z,t+1}^{(1-\chi)(1-\chi_0)} \right)^{1-\alpha} \right] \right]^{\frac{1}{1-\alpha}}}{\tilde{V}_{t+1}\mu_{z,t+1}^{(1-\chi)(1-\chi_0)}} \right)^{\alpha} \frac{d_{t+1}}{d_t} \frac{\left(\widetilde{c}_{t+1} - b\mu_{z,t}^{-1}\widetilde{c}_t\right)^{-\chi}}{\left(\widetilde{c}_t - b\mu_{z,t-1}^{-1}\widetilde{c}_{t-1}\right)^{-\chi}} \mu_{z,t+1}^{-\chi(1-\chi_0)-\chi_0} \frac{1}{\pi_{t+1}}$
	Firm
4	$\widetilde{w}_t = mc_t (1-\theta) a_t k_{ss}^{\theta} l_t^{-\theta}$
5	$(1-\eta)\widetilde{y}_t + \mathbb{E}_t \left[\xi M_{t,t+1}^{\text{real}} \left(\frac{\pi_{t+1}}{\pi_{ss}^{\nu}} - 1 \right) \frac{\pi_{t+1}}{\pi_{ss}^{\nu}} \widetilde{y}_{t+1}\mu_{z,t+1} \right] + \eta mc_t \widetilde{y}_t = \xi \left(\frac{\pi_t}{\pi_{ss}^{\nu}} - 1 \right) \widetilde{y}_t \frac{\pi_t}{\pi_{ss}^{\nu}}$
	Central bank
6	$R_t = R_{ss}\mathbb{E}_t \left[\exp \left\{ \phi_\pi \log \left(\frac{\pi_{t+1}}{\pi_{ss}\pi_{t+1}^*} \right) + \phi_{\Delta c} \left(\log \left(\frac{\widetilde{c}_{t+1}}{\widetilde{c}_t} \right) + \log \mu_{z,t+1} - \log \mu_{z,ss} \right) \right\} \right]$
	Aggregation
7	$\widetilde{c}_t + \delta k_{ss} = \left(1 - \frac{\xi}{2} \left(\frac{\pi_t}{\pi_{ss}^{\nu}} - 1 \right)^2 \right) \widetilde{y}_t$
8	$\widetilde{y}_t = a_t k_{ss}^{\theta} l_t^{1-\theta}$
	Links
9	$(\widetilde{c}_{t-1})_{t+1} = \widetilde{c}_t$
10-	**Shocks**

To these equations, we add the recursive equations for bond prices, i.e. $B_t^{(k)} = \mathbb{E}_t \left[M_{t,t+1} B_{t+1}^{(k-1)} \right]$.

1.8 Steady state

Some steady state values are calibrated to certain values, i.e. $KoY = \frac{k_{ss}}{4y_{ss}}$, l_{ss}, π_{ss}, $\mu_{z,ss}$.

1.8.1 Separable utility case

Notice, that in steady state we have

$$M_{t,t+1} = \beta \left(\frac{\left[\mathbb{E}_t \left[\left(\tilde{V}_{t+1}\mu_{z,t+1}^{(1-\chi)(1-\chi_0)} \right)^{1-\alpha} \right] \right]^{\frac{1}{1-\alpha}}}{\tilde{V}_{t+1}\mu_{z,t+1}^{(1-\chi)(1-\chi_0)}} \right)^{\alpha} \frac{d_{t+1}}{d_t} \frac{\left(\widetilde{c}_{t+1} - b\mu_{z,t}^{-1}\widetilde{c}_t\right)^{-\chi}}{\left(\widetilde{c}_t - b\mu_{z,t-1}^{-1}\widetilde{c}_{t-1}\right)^{-\chi}} \mu_{z,t+1}^{-\chi(1-\chi_0)-\chi_0} \frac{1}{\pi_{t+1}}$$

⇓

$$M_{ss,ss} = \beta \frac{\mu_{z,ss}^{-\chi(1-\chi_0)-\chi_0}}{\pi_{ss}}$$

Then it follows from equation (3) that

$$\mathbb{E}_t [M_{t,t+1}R_t] = 1$$

⇓

$$R_{ss} = \frac{1}{M_{ss}}$$

From equation (8), we have that $\widetilde{y}_{ss} = a_{ss}k_{ss}^{\theta}l_{ss}^{1-\theta}$. Dividing by $k_{ss} \neq 0$ and multiplying by 4

$$\frac{4\widetilde{y}_{ss}}{k_{ss}} = \frac{1}{KoY} = 4a_{ss}k_{ss}^{\theta-1}l_{ss}^{1-\theta}.$$

Then, as $a_{ss} = 1$, we get

$$\frac{1}{k_{ss}^{\theta-1}} = 4KoYl_{ss}^{1-\theta}$$

$\Updownarrow$

$$k_{ss}^{1-\theta} = 4KoYl_{ss}^{1-\theta}$$

$\Updownarrow$

$$k_{ss} = l_{ss}\left(4 \cdot KoY\right)^{\frac{1}{1-\theta}}.$$

From equation (8) then

$$\widetilde{y}_{ss} = k_{ss}^{\theta}l_{ss}^{1-\theta},$$

and from equation (7)

$$\widetilde{c}_{ss} + \delta k_{ss} = \left(1 - \frac{\xi}{2}\left(\frac{\pi_{ss}}{\pi_{ss}^{\nu}} - 1\right)^2\right)\widetilde{y}_{ss}$$

$\Updownarrow$

$$\widetilde{c}_{ss} = \left(1 - \frac{\xi}{2}\left(\frac{\pi_{ss}}{\pi_{ss}^{\nu}} - 1\right)^2\right)\widetilde{y}_{ss} - \delta k_{ss}$$

From equation (5) we have

$$\left(1-\eta\right)\widetilde{y}_{ss} + \xi M_{ss}^{\text{real}}\left(\frac{\pi_{ss}}{\pi_{ss}^{\nu}} - 1\right)\frac{\pi_{ss}}{\pi_{ss}^{\nu}}\widetilde{y}_{ss}\mu_{z,ss} + \eta mc_{ss}\widetilde{y}_{ss} = \xi\left(\frac{\pi_{ss}}{\pi_{ss}^{\nu}} - 1\right)\widetilde{y}_{ss}\frac{\pi_{ss}}{\pi_{ss}^{\nu}}$$

$\Updownarrow$

$$\eta mc_{ss}\widetilde{y}_{ss} = \xi\left(\frac{\pi_{ss}}{\pi_{ss}^{\nu}} - 1\right)\widetilde{y}_{ss}\frac{\pi_{ss}}{\pi_{ss}^{\nu}} - \left(1-\eta\right)\widetilde{y}_{ss} - \xi M_{ss}^{\text{real}}\left(\frac{\pi_{ss}}{\pi_{ss}^{\nu}} - 1\right)\frac{\pi_{ss}}{\pi_{ss}^{\nu}}\widetilde{y}_{ss}\mu_{z,ss}$$

$\Updownarrow$

$$mc_{ss} = \frac{1}{\eta\widetilde{y}_{ss}}\left[\xi\left(\frac{\pi_{ss}}{\pi_{ss}^{\nu}} - 1\right)\widetilde{y}_{ss}\frac{\pi_{ss}}{\pi_{ss}^{\nu}} - \left(1-\eta\right)\widetilde{y}_{ss} - \xi M_{ss}^{\text{real}}\left(\frac{\pi_{ss}}{\pi_{ss}^{\nu}} - 1\right)\frac{\pi_{ss}}{\pi_{ss}^{\nu}}\widetilde{y}_{ss}\mu_{z,ss}\right]$$

Inserted in Equation (4) we then get

$$\widetilde{w}_{ss} = mc_{ss}\left(1-\theta\right)a_t k_{ss}^{\theta}l_t^{-\theta}$$

The parameter φ_0 is set to imply the choosen steady state for labor, i.e. l_{ss}. Hence, from equation (2) we get

$$\widetilde{w}_{ss} = \kappa_w\widetilde{w}_{ss} + \left(1-\kappa_w\right)\left[\varphi_0\left(1-l_{ss}\right)^{-\frac{1}{\varphi}}\left(\frac{\widetilde{c}_{ss} - b\widetilde{c}_{ss}\mu_{z,ss}^{-1}}{\left(\widetilde{c}_{ss}\right)^{\chi_0}}\right)^{\chi}\left(\widetilde{c}_{ss}\right)^{\chi_0}\right]$$

$\Updownarrow$

$$\widetilde{w}_{ss} = \varphi_0\left(1-l_{ss}\right)^{-\frac{1}{\varphi}}\left(\frac{\widetilde{c}_{ss} - b\widetilde{c}_{ss}\mu_{z,ss}^{-1}}{\left(\widetilde{c}_{ss}\right)^{\chi_0}}\right)^{\chi}\left(\widetilde{c}_{ss}\right)^{\chi_0}$$

$\Updownarrow$

$$\frac{\widetilde{w}_{ss}}{\left(\widetilde{c}_{ss}\right)^{\chi_0}}\left(\frac{\widetilde{c}_{ss} - b\widetilde{c}_{ss}\mu_{z,ss}^{-1}}{\left(\widetilde{c}_{ss}\right)^{\chi_0}}\right)^{-\chi} = \varphi_0\left(1-l_{ss}\right)^{-\frac{1}{\varphi}}$$

$\Updownarrow\Downarrow$

$$\varphi_0 \left(1 - l_{ss}\right)^{-\frac{1}{\varphi}} = \left(\frac{\widetilde{c}_{ss} - b\widetilde{c}_{ss}\mu_{z,ss}^{-1}}{\widetilde{c}_{ss}^{\chi_0}}\right)^{-\chi} \frac{\widetilde{w}_{ss}}{\widetilde{c}_{ss}^{\chi_0}}$$

$\Updownarrow$

$$\varphi_0 = \frac{\left(\frac{\widetilde{c}_{ss} - b\widetilde{c}_{ss}\mu_{z,ss}^{-1}}{\widetilde{c}_{ss}^{\chi_0}}\right)^{-\chi} \widetilde{w}_{ss}}{\left(1 - l_{ss}\right)^{-\frac{1}{\varphi}} \widetilde{c}_{ss}^{\chi_0}}$$

Finally, from equation (1)

$$\tilde{V}_{ss} = u_0 + \frac{1}{1-\chi}\left(\frac{\widetilde{c}_{ss} - b\widetilde{c}_{ss}\mu_{z,ss}^{-1}}{\widetilde{c}_{ss}^{\chi_0}}\right)^{1-\chi} + u_0^d + \varphi_0 \frac{\left(1 - l_{ss}\right)^{1-\frac{1}{\varphi}}}{1 - \frac{1}{\varphi}} + \beta\left(\tilde{V}_{ss}\mu_{z,ss}^{(1-\chi)(1-\chi_0)}\right)$$

$\Updownarrow$

$$V_{ss} = \frac{1}{1 - \beta\mu_{z,ss}^{(1-\chi)(1-\chi_0)}}\left[u_0 + u_0^d + \frac{1}{1-\chi}\left(\frac{\widetilde{c}_{ss} - b\widetilde{c}_{ss}\mu_{z,ss}^{-1}}{\widetilde{c}_{ss}^{\chi_0}}\right)^{1-\chi} + \varphi_0 \frac{\left(1 - l_{ss}\right)^{1-\frac{1}{\varphi}}}{1 - \frac{1}{\varphi}}\right].$$

2 Digression on The New Keynesian Model

2.1 Risk Aversion at the Steady State

We follow Swanson (2012) and compute two measures of relative risk aversion in our model. With recursive Epstein-Zin preferences controlled by α, there are two measures of relative risk aversion. The first measure RRA^c applies when there is no upper bound for labor and therefore total household wealth A_t equals the present discounted value of consumption. The other measure RRA^{cl} applies when the upper bound for the household's time endowment is well-specified, meaning that total household wealth $\tilde{A}_t$ equals the present discounted value of leisure plus consumption.

Throughout this section we use the notational convention in Swanson (2012) where a variable in the steady state is denoted without a subscript. For instance c is the steady state value of c_t. Moreover, $u = u\left(c, l\right)$.

The general formulas with external habit formation are (see Swanson (2012), page 24, eq 53 and eq 54)

$$RRA^{cl} = \frac{-u_{11} + \lambda u_{12}}{u_1} \frac{c + w\left(1 - l\right)}{1 + w\lambda} + \alpha \frac{\left(c + w\left(1 - l\right)\right) u_1}{u}$$

$$RRA^{c} = c\left(\frac{-u_{11} + \lambda u_{12}}{u_1} \frac{1}{1 + w\lambda} + \alpha \frac{u_1}{u}\right)$$

where

$$w = -\frac{u_2}{u_1}$$

$$\lambda = \frac{wu_{11} + u_{12}}{u_{22} + wu_{12}}$$

Note that

$$RRA^{cl} = \tfrac{-u_{11} + \lambda u_{12}}{u_1} \tfrac{c + w(1-l)}{1 + w\lambda} + \alpha \tfrac{(c + w(1-l)) u_1}{u}$$

$$= \tfrac{c + w(1-l)}{c} c\left[\tfrac{-u_{11} + \lambda u_{12}}{u_1} \tfrac{1}{1 + w\lambda} + \alpha \tfrac{u_1}{u}\right]$$

$\Updownarrow$

$$RRA^{cl} = \left(1 + \frac{w}{c}\left(1 - l\right)\right) RRA^c.$$

Here, we use the notation that

- $u =$ the utility index
- $u_1 =$ the partial derivative of u with respect to consumption
- $u_2 =$ the partial derivative of u with respect to hours worked
- $w =$ the steady state wage level
- $c =$ the steady state consumption level
- $l =$ hours worked

Note that these formulas also apply in our case even though we have wage strickiness. The reason is that the household derives its FOC in a frictionless setting, and the steady state of wage is not affected by the presence of wage stickiness, i.e., when using the notation above we have $w_t^*/z_t = w_t/z_t$ in the steady state. Also, Swanson (2012) shows that the above formulas also hold with balanced growth (see Swanson (2012) page 48). Recall that our utility function reads (ignoring d_t and n_t as $d_{ss} = n_{ss} = 1$)

$$u\left(c_t - bc_{t-1}, 1 - l_t\right) = \left(u_0^d + u_0\right) z_t^{(1-\chi)(1-\chi_0)} + \frac{1}{1-\chi}\left(\frac{c_t - bc_{t-1}}{\widetilde{c}_{ss}^{\chi_0} z_t^{\chi_0}}\right)^{1-\chi} + z_t^{(1-\chi)(1-\chi_0)}\varphi_0\frac{\left(1-l_t\right)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}}$$

Hence, from now on all steady state variables refer to those in the normalized economy without trends. Hence, in the steady state we have

$$u = \left(u_0^d + u_0\right) + \frac{1}{1-\chi}\left(\frac{c_t - bc\mu_{z,t}^{-1}}{c^{\chi_0}}\right)^{1-\chi} + \varphi_0\frac{(1-l)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}} = \left(u_0^d + u_0\right) + \frac{c^{\chi_0(\chi-1)}}{1-\chi}\left(c - bc\mu_z^{-1}\right)^{1-\chi} + \varphi_0\frac{(1-l)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}}$$

$$u_1 = c^{\chi_0(\chi-1)}\left(c - bc\mu_z^{-1}\right)^{-\chi}$$

$$u_{11} = -\chi c^{\chi_0(\chi-1)}\left(c - bc\mu_z^{-1}\right)^{-\chi-1}$$

$$u_2 = -\phi_0\left(1-l\right)^{-\frac{1}{\varphi}}$$

$$u_{22} = -\frac{1}{\varphi}\left(-\phi_0\right)\left(1-l\right)^{-\frac{1}{\varphi}-1}(-1) = -\frac{1}{\varphi}\phi_0\left(1-l\right)^{-\frac{1}{\varphi}-1}$$

$$u_{12} = 0$$

Thus

$$w = -\frac{u_2}{u_1} = -\frac{-\phi_0(1-l)^{-\frac{1}{\varphi}}}{c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi}} = \frac{\phi_0(1-l)^{-\frac{1}{\varphi}}}{c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi}}$$

$$\Updownarrow$$

$$\phi_0 = \frac{c^{\chi_0(\chi-1)}w\left(c-bc\mu_z^{-1}\right)^{-\chi}}{(1-l)^{-\frac{1}{\varphi}}}$$

And

$$\lambda = \frac{wu_{11}+u_{12}}{u_{22}+wu_{12}} = \frac{w\left[-\chi c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi-1}\right]}{-\frac{1}{\varphi}\phi_0(1-l)^{-\frac{1}{\varphi}-1}}$$

$$= \frac{w\chi c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi-1}}{\frac{1}{\varphi}\frac{c^{\chi_0(\chi-1)}w\left(c-bc\mu_z^{-1}\right)^{-\chi}}{(1-l)^{-\frac{1}{\varphi}}}(1-l)^{-\frac{1}{\varphi}-1}}$$

$$= \frac{\chi\left(c-bc\mu_z^{-1}\right)^{-1}}{\frac{1}{\varphi}(1-l)^{-1}}$$

$$= \frac{\chi(1-l)}{\frac{1}{\varphi}\left(c-bc\mu_z^{-1}\right)}$$

Note also that

$$w\lambda = \frac{\phi_0(1-l)^{-\frac{1}{\varphi}}}{c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi}} \frac{\chi(1-l)}{\frac{1}{\varphi}\left(c-bc\mu_z^{-1}\right)} = \frac{\chi\phi_0(1-l)^{1-\frac{1}{\varphi}}}{\frac{1}{\varphi}c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{1-\chi}}$$

Hence, the first measure of relative risk-aversion is:
$RRA^c = c\left(\frac{-u_{11}}{u_1}\frac{1}{1+w\lambda}+\alpha\frac{u_1}{u}\right)$

$$= c\frac{\chi c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi-1}}{c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi}} \frac{1}{1+\frac{\chi\phi_0(1-l)^{1-\frac{1}{\varphi}}}{\frac{1}{\varphi}c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{1-\chi}}} + c\alpha\frac{c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi}}{\left(u_0^d+u_0\right)+\frac{c^{\chi_0(\chi-1)}}{1-\chi}\left(c-bc\mu_z^{-1}\right)^{1-\chi}+\varphi_0\frac{(1-l)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}}}$$

$$= c\left[\chi\left(c-bc\mu_z^{-1}\right)^{-1}\right] \frac{1}{1+\frac{\chi\phi_0(1-l)^{1-\frac{1}{\varphi}}}{\frac{1}{\varphi}c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{1-\chi}}} + c\alpha\frac{c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi}}{\left(u_0^d+u_0\right)+\frac{c^{\chi_0(\chi-1)}}{1-\chi}\left(c-bc\mu_z^{-1}\right)^{1-\chi}+\varphi_0\frac{(1-l)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}}}$$

$$= \frac{c\chi}{c-bc\mu_z^{-1}} \frac{1}{1+\frac{\chi\phi_0(1-l)^{1-\frac{1}{\varphi}}}{\frac{1}{\varphi}c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{1-\chi}}} + c\alpha\frac{c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi}}{\left(u_0^d+u_0\right)+\frac{c^{\chi_0(\chi-1)}}{1-\chi}\left(c-bc\mu_z^{-1}\right)^{1-\chi}+\varphi_0\frac{(1-l)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}}}$$

$$= \frac{c\chi}{c-bc\mu_z^{-1}+\frac{\chi\phi_0(1-l)^{1-\frac{1}{\varphi}}}{\frac{1}{\varphi}c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi}}} + c\alpha\frac{c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi}}{\left(u_0^d+u_0\right)+\frac{c^{\chi_0(\chi-1)}}{1-\chi}\left(c-bc\mu_z^{-1}\right)^{1-\chi}+\varphi_0\frac{(1-l)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}}}$$

$$= \frac{c\chi\left(c-bc\mu_z^{-1}\right)^{-\chi}}{\left(c-bc\mu_z^{-1}\right)^{1-\chi}+\frac{\chi\phi_0(1-l)^{1-\frac{1}{\varphi}}}{c^{\chi_0(\chi-1)}\frac{1}{\varphi}}} + c\alpha\frac{c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi}}{\left(u_0^d+u_0\right)+\frac{c^{\chi_0(\chi-1)}}{1-\chi}\left(c-bc\mu_z^{-1}\right)^{1-\chi}+\varphi_0\frac{(1-l)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}}}$$

$$= \frac{c\chi\left(c-bc\mu_z^{-1}\right)^{-\chi}}{\left(c-bc\mu_z^{-1}\right)^{1-\chi}+\frac{wc^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi}}{(1-l)^{-\frac{1}{\varphi}}}\frac{\chi(1-l)^{1-\frac{1}{\varphi}}}{c^{\chi_0(\chi-1)}\frac{1}{\varphi}}} + c\alpha\frac{c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi}}{\left(u_0^d+u_0\right)+\frac{c^{\chi_0(\chi-1)}}{1-\chi}\left(c-bc\mu_z^{-1}\right)^{1-\chi}+\frac{wc^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi}}{(1-l)^{-\frac{1}{\varphi}}}\frac{(1-l)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}}}$$

using $\phi_0 = \frac{wc^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi}}{(1-l)^{-\frac{1}{\varphi}}}$

$$= \frac{c\chi}{\left(c-bc\mu_z^{-1}\right)+\frac{w}{1}\frac{\chi(1-l)}{\frac{1}{\varphi}}} + c\alpha\frac{c^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi}}{\left(u_0^d+u_0\right)+\frac{c^{\chi_0(\chi-1)}}{1-\chi}\left(c-bc\mu_z^{-1}\right)^{1-\chi}+\frac{wc^{\chi_0(\chi-1)}\left(c-bc\mu_z^{-1}\right)^{-\chi}}{1}\frac{(1-l)}{1-\frac{1}{\varphi}}}$$

$$= \frac{c\chi}{\left(c-bc\mu_z^{-1}\right)+\frac{w}{1}\frac{\chi(1-l)}{\frac{1}{\varphi}}} + c\alpha\frac{\left(c-bc\mu_z^{-1}\right)^{-\chi}}{\frac{\left(u_0^d+u_0\right)}{c^{\chi_0(\chi-1)}}+\frac{1}{1-\chi}\left(c-bc\mu_z^{-1}\right)^{1-\chi}+\frac{w\left(c-bc\mu_z^{-1}\right)^{-\chi}}{1}\frac{(1-l)}{1-\frac{1}{\varphi}}}$$

$$= \frac{\chi}{\left(1-b\mu_z^{-1}\right)+\frac{w}{c}\frac{\chi(1-l)}{\frac{1}{\varphi}}} + c\alpha\frac{1}{\frac{\left(u_0^d+u_0\right)}{c^{\chi_0(\chi-1)}}\left(c-bc\mu_z^{-1}\right)^{\chi}+\frac{1}{1-\chi}\left(c-bc\mu_z^{-1}\right)+\frac{w}{1}\frac{(1-l)}{1-\frac{1}{\varphi}}}$$

$$= \frac{\chi}{\left(1-b\mu_z^{-1}\right)+\frac{w}{c}\frac{\chi(1-l)}{\frac{1}{\varphi}}} + \alpha\frac{1}{\frac{\left(u_0^d+u_0\right)}{c^{\chi_0(\chi-1)}c}\left(c-bc\mu_z^{-1}\right)^{\chi}+\frac{1}{1-\chi}\left(1-b\mu_z^{-1}\right)+\frac{w}{c}\frac{(1-l)}{1-\frac{1}{\varphi}}}$$

$$= \frac{\chi}{\left(1-b\mu_z^{-1}\right)+\frac{w}{c}\frac{\chi(1-l)}{\frac{1}{\varphi}}} + \alpha\frac{1-\chi}{(1-\chi)\frac{\left(u_0^d+u_0\right)}{c^{\chi_0(\chi-1)+1}}\left(c-bc\mu_z^{-1}\right)^{\chi}+\left(1-b\mu_z^{-1}\right)+\frac{w(1-l)}{c}\frac{(1-\chi)}{1-\frac{1}{\varphi}}}$$

Note when $b = 0, \chi_0 = 0$, and $u_0 = 0$, we get

$$RRA^c = \frac{\chi}{1+\frac{w}{c}\frac{\chi(1-l)}{\frac{1}{\varphi}}} + \alpha\frac{1-\chi}{(1-\chi)\frac{u_0^d}{c^{\chi_0(\chi-1)+1}}c^{\chi}+1+\frac{w(1-l)}{c}\frac{(1-\chi)}{1-\frac{1}{\varphi}}}$$

$$= \frac{\chi}{1+\chi\varphi\frac{w(1-l)}{c}} + \alpha\frac{1-\chi}{(1-\chi)u_0^d c^{\chi-1}+1+\frac{(1-\chi)}{1-\frac{1}{\varphi}}\frac{w(1-l)}{c}}$$

Also, when $u_0^d = u_0 = 0$ (standard EZ preferences) we get

$$RRA^c = \frac{\chi}{\left(1-b\mu_z^{-1}\right)+\frac{w}{c}\frac{\chi(1-l)}{\frac{1}{\varphi}}} + \alpha\frac{1-\chi}{(1-\chi)\frac{\left(u_0^d+u_0\right)}{c^{\chi_0(\chi-1)+1}}\left(c-bc\mu_z^{-1}\right)^{\chi}+\left(1-b\mu_z^{-1}\right)+\frac{w(1-l)}{c}\frac{(1-\chi)}{1-\frac{1}{\varphi}}}$$

$$= \frac{\chi}{1-b\mu_z^{-1}+\frac{w}{c}\frac{\chi(1-l)}{\frac{1}{\varphi}}} + \alpha\frac{1-\chi}{1-b\mu_z^{-1}+\frac{w(1-l)}{c}\frac{(1-\chi)}{1-\frac{1}{\varphi}}}$$

Thus, to back out α from RRA^c we have

$$RRA^c = \frac{\chi}{\left(1-b\mu_z^{-1}\right)+\frac{w}{c}\frac{\chi(1-l)}{\frac{1}{\varphi}}} + \alpha\frac{1-\chi}{(1-\chi)\frac{\left(u_0^d+u_0\right)}{c^{\chi_0(\chi-1)+1}}\left(c-bc\mu_z^{-1}\right)^{\chi}+\left(1-b\mu_z^{-1}\right)+\frac{w(1-l)}{c}\frac{(1-\chi)}{1-\frac{1}{\varphi}}}$$

$\Updownarrow$

$$\alpha = \frac{RRA^c - \frac{\chi}{\left(1-b\mu_z^{-1}\right)+\frac{w}{c}\frac{\chi(1-l)}{\frac{1}{\varphi}}}}{\frac{1-\chi}{(1-\chi)\frac{\left(u_0^d+u_0\right)}{c^{\chi_0(\chi-1)+1}}\left(c-bc\mu_z^{-1}\right)^{\chi}+\left(1-b\mu_z^{-1}\right)+\frac{w(1-l)}{c}\frac{(1-\chi)}{1-\frac{1}{\varphi}}}}$$

Note also that

$$RRA^c = c\left(\frac{-u_{11}}{u_1}\frac{1}{1+w\lambda}+\alpha\frac{u_1}{u}\right)$$

$\Updownarrow$

$$\frac{RRA^c}{c}+\frac{u_{11}}{u_1}\frac{1}{1+w\lambda} = \alpha\frac{u_1}{u}$$

$\Updownarrow$

$$\alpha = \left(\frac{RRA^c}{c}+\frac{u_{11}}{u_1}\frac{1}{1+w\lambda}\right)\frac{u}{u_1}$$

And therefore

$$RRA^{cl} = \left(1+\frac{w}{c}(1-l)\right)RRA^c = \left(1+\frac{w}{c}(1-l)\right)RRA^c$$

If we condition on a given value of α (for instance, based on reasons motivated by accuracy of the model solution), then we can alternatively back out the value of u_0 to get a given RRA^c. That is, we get

$$RRA^c = \frac{\chi}{\left(1-b\mu_z^{-1}\right)+\frac{w}{c}\frac{\chi(1-l)}{\frac{1}{\varphi}}} + \alpha\frac{1-\chi}{(1-\chi)\frac{\left(u_0^d+u_0\right)}{c^{\chi_0(\chi-1)+1}}\left(c-bc\mu_z^{-1}\right)^{\chi}+\left(1-b\mu_z^{-1}\right)+\frac{w(1-l)}{c}\frac{(1-\chi)}{1-\frac{1}{\varphi}}}$$

$\Updownarrow$

$$\frac{1}{\alpha(1-\chi)}\left[RRA^c - \frac{\chi}{\left(1-b\mu_z^{-1}\right)+\frac{w}{c}\frac{\chi(1-l)}{\frac{1}{\varphi}}}\right] = \frac{1}{(1-\chi)\frac{\left(u_0^d+u_0\right)}{c^{\chi_0(\chi-1)+1}}\left(c-bc\mu_z^{-1}\right)^{\chi}+\left(1-b\mu_z^{-1}\right)+\frac{w(1-l)}{c}\frac{(1-\chi)}{1-\frac{1}{\varphi}}}$$

$\Updownarrow$

$$(1-\chi)\frac{\left(u_0^d+u_0\right)}{c^{\chi_0(\chi-1)+1}}\left(c-bc\mu_z^{-1}\right)^{\chi}+\left(1-b\mu_z^{-1}\right)+\frac{w\,(1-l)}{c}\frac{(1-\chi)}{1-\frac{1}{\varphi}}=\frac{\alpha\,(1-\chi)}{\left(RRA^c-\frac{\chi}{\left(1-b\mu_z^{-1}\right)+\frac{w}{c}\frac{\chi(1-l)}{\frac{1}{\varphi}}}\right)}$$

$\Updownarrow$

$$(1-\chi)\frac{\left(u_0^d+u_0\right)}{c^{\chi_0(\chi-1)+1}}\left(c-bc\mu_z^{-1}\right)^{\chi}=\frac{\alpha\,(1-\chi)}{\left(RRA^c-\frac{\chi}{\left(1-b\mu_z^{-1}\right)+\frac{w}{c}\frac{\chi(1-l)}{\frac{1}{\varphi}}}\right)}-\left(1-b\mu_z^{-1}\right)-\frac{w\,(1-l)}{c}\frac{(1-\chi)}{1-\frac{1}{\varphi}}$$

$\Updownarrow$

$$u_0^d+u_0=\frac{c^{\chi_0(\chi-1)+1}}{\left(c-bc\mu_z^{-1}\right)^{\chi}(1-\chi)}\left[\frac{\alpha\,(1-\chi)}{\left(RRA^c-\frac{\chi}{\left(1-b\mu_z^{-1}\right)+\frac{w}{c}\frac{\chi(1-l)}{\frac{1}{\varphi}}}\right)}-\left(1-b\mu_z^{-1}\right)-\frac{w\,(1-l)}{c}\frac{(1-\chi)}{1-\frac{1}{\varphi}}\right]$$

where either u_0^d or u_0 are zero.

2.2 The Frisch Labor Supply Elasticity

Recall that this elasticity is given by

$elas_F=\frac{u_l}{l\left(u_{ll}-\frac{u_{cl}}{u_{cc}}\right)}$

$\varphi_0\frac{(1-l_t)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}}$

In our case (given that the utility of leisure is $\varphi_0\frac{(1-l_t)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}}$)

$u_l=-\varphi_0\,(1-l_t)^{-\frac{1}{\varphi}}$

$u_{ll}=-\varphi_0\left(-\frac{1}{\varphi}\right)(1-l_t)^{-\frac{1}{\varphi}-1}(-1)=-\varphi_0\frac{1}{\varphi}\,(1-l_t)^{-\frac{1}{\varphi}-1}$

$u_{cl}=0$

So, in the steady state we have

$elas_F=\frac{u_l}{l_{ss}\left(u_{ll}-\frac{u_{cl}}{u_{cc}}\right)}$

$=\frac{-\varphi_0(1-l_{ss})^{-\frac{1}{\varphi}}}{l_{ss}\left(-\varphi_0\frac{1}{\varphi}(1-l_{ss})^{-\frac{1}{\varphi}-1}-0\right)}$

$=\frac{\varphi}{l_{ss}(1-l_{ss})^{-1}}$

$=\varphi\frac{1-l_{ss}}{l_{ss}}$

$=\varphi\left(\frac{1}{l_{ss}}-1\right)$

When $l_{ss}=\frac{1}{3}$, then

$elas_F=\varphi\left(\frac{3}{1}-1\right)=2\varphi$

2.3 Intertemporal Elasticity of Substitution

The intertemporal elasticity of substitution (IES) is given by

$$
\begin{aligned}
IES_t &= -\frac{\mathcal{U}_c}{\mathcal{U}_{cc}c_t} \\
&= \frac{-\widetilde{c}_{ss}^{\chi_0(\chi-1)}\left(\frac{c_t-bc_{t-1}}{z_t^{\chi_0}}\right)^{-\chi}\frac{1}{\widetilde{c}_{ss}^{\chi_0}z_t^{\chi_0}}}{-\chi\widetilde{c}_{ss}^{\chi_0(\chi-1)}\left(\frac{c_t-bc_{t-1}}{z_t^{\chi_0}}\right)^{-\chi-1}\frac{c_t}{z_t^{2\chi_0}}} \\
&= \frac{-\left(\frac{c_t-bc_{t-1}}{z_t^{\chi_0}}\frac{z_t^{1-\chi_0}}{z_t^{1-\chi_0}}\right)^{-\chi}}{-\chi\left(\frac{c_t-bc_{t-1}}{z_t^{\chi_0}}\frac{z_t^{1-\chi_0}}{z_t^{1-\chi_0}}\right)^{-\chi-1}\frac{c_t}{z_t^{\chi_0}}} \\
&= \frac{-\left(\frac{c_t-bc_{t-1}}{z_t}\frac{z_t^{1-\chi_0}}{1}\right)^{-\chi}}{-\chi\left(\frac{c_t-bc_{t-1}}{z_t}\frac{z_t^{1-\chi_0}}{1}\right)^{-\chi-1}\frac{c_t}{z_t^{\chi_0}}} \\
&= \frac{\left(\tilde{c}_t-b\tilde{c}_{t-1}\mu_{z,t}^{-1}\right)^{-\chi}}{\chi\left(\tilde{c}_t-b\tilde{c}_{t-1}\mu_{z,t}^{-1}\right)^{-\chi-1}\frac{c_t}{z_t^{\chi_0}}}\frac{1}{z_t^{-(1-\chi_0)}} \\
&= \frac{\left(\tilde{c}_t-b\tilde{c}_{t-1}\mu_{z,t}^{-1}\right)^{-\chi}}{\chi\left(\tilde{c}_t-b\tilde{c}_{t-1}\mu_{z,t}^{-1}\right)^{-\chi-1}c_tz_t^{-\chi_0}}\frac{1}{z_t^{-1+\chi_0}} \\
&= \frac{\tilde{c}_t-b\tilde{c}_{t-1}\mu_{z,t}^{-1}}{\chi\tilde{c}_t}
\end{aligned}
$$

In the steady state we get

$$IES_{ss} = \frac{1}{\chi}\left(1-b\mu_{z,ss}^{-1}\right)$$

2.4 The Equity Price

We have

$$P_t^m = \mathbb{E}_t\left[M_{t,t+1}^{\text{real}}\left(Div_{t+1}+P_{t+1}^m\right)\right],$$

where $Div_t = c_t$, i.e. a claim on consumption.

Inducing stationarity

$$\frac{P_t^m}{z_t} = \mathbb{E}_t\left[M_{t,t+1}^{\text{real}}\left(\left(\frac{Div_{t+1}}{z_{t+1}}\right)+\left(\frac{P_{t+1}^m}{z_{t+1}}\right)\right)\left(\frac{z_{t+1}}{z_t}\right)\right]$$

$\Updownarrow$

$$\tilde{P}_t^m = \mathbb{E}_t\left[M_{t,t+1}^{\text{real}}\left(\widetilde{Div}_{t+1}+\tilde{P}_{t+1}^m\right)\mu_{z,t+1}\right]$$

Hence, in the steady state:

$$\tilde{P}_{ss}^m = M_{ss,ss+1}^{\text{real}}\left(\widetilde{Div}_{ss}+\tilde{P}_{ss}^m\right)\mu_{z,ss}$$

$\Updownarrow$

$$\tilde{P}_{ss}^m = M_{ss,ss+1}^{\text{real}}\mu_{z,ss}\widetilde{Div}_{ss}+M_{ss}^{\text{real}}\mu_{z,ss}\tilde{P}_{ss}^m$$

$\Updownarrow$

$$\tilde{P}_{ss}^m = \frac{M_{ss,ss+1}^{\text{real}}\mu_{z,ss}\widetilde{Div}_{ss}}{1-M_{ss}^{\text{real}}\mu_{z,ss}}$$

Recall that $M_{ss,ss} = \beta \frac{\mu_{z,ss}^{-\chi(1-\chi_0)-\chi_0}}{\pi_{ss}}$, so $M_{ss,ss}^{\text{real}} = \beta \mu_{z,ss}^{-\chi(1-\chi_0)-\chi_0}$, and we therefore get

$$\tilde{P}_{ss}^{m} = \frac{\beta \mu_{z,ss}^{1-\chi(1-\chi_0)-\chi_0} \widetilde{Div}_{ss}}{1 - \beta \mu_{z,ss}^{1-\chi(1-\chi_0)-\chi_0}}$$

where $\widetilde{Div}_{ss} = \tilde{c}_{ss}$. Note also that

$$\begin{aligned} P_t^m &= \tilde{P}_t^m z_t \\ &= \exp\left\{\log \tilde{P}_t^m + \log z_t\right\} \end{aligned}$$

and similarly for dividends.

The requity return is

$$\begin{aligned} \exp\left\{r_{t+1}^m\right\} &\equiv \frac{\left(Div_{t+1} + P_{t+1}^m\right)}{\tilde{P}_t^m} \\ &= \frac{\left(\widetilde{Div}_{t+1}^{\omega} + \tilde{P}_{t+1}^m\right)\mu_{z,t+1}}{\tilde{P}_t^m} \end{aligned}$$

$\Updownarrow$

$$r_{t+1}^m = \log\left(\widetilde{Div}_{t+1} + \tilde{P}_{t+1}^m\right) - \log\left(\tilde{P}_t^m\right) + \log \mu_{z,t+1}.$$

2.5 The Timing Premium

We extend the definition of the timing premium to accommodate leisure in the utility function. Here, we apply the approach suggested in Andreasen & Jørgensen (2020). The basic idea is to impose that the household is on the equilibrium path for the leisure and consumption trade-off. That is, we use the first order condition

$$z_t^{(1-\chi)(1-\chi_0)} n_t \varphi_0 \left(1 - l_t\right)^{-\frac{1}{\varphi}} = \left(\frac{c_t - bc_{t-1}}{\left(\widetilde{c}_{ss} z_t\right)^{\chi_0}}\right)^{-\chi} \frac{w_t^*}{\left(\widetilde{c}_{ss} z_t\right)^{\chi_0}}$$

$\Updownarrow$

$$\frac{1}{\left(1 - l_t\right)^{\frac{1}{\varphi}}} = \frac{\left(\frac{c_t - bc_{t-1}}{\left(\widetilde{c}_{ss} z_t\right)^{\chi_0}}\right)^{-\chi} \frac{w_t^*}{\left(\widetilde{c}_{ss} z_t\right)^{\chi_0}}}{z_t^{(1-\chi)(1-\chi_0)} n_t \varphi_0}$$

$\Updownarrow$

$$\left(1 - l_t\right)^{\frac{1}{\varphi}} = \frac{z_t^{(1-\chi)(1-\chi_0)} n_t \varphi_0}{\left(\frac{c_t - bc_{t-1}}{\left(\widetilde{c}_{ss} z_t\right)^{\chi_0}}\right)^{-\chi} \frac{w_t^*}{\left(\widetilde{c}_{ss} z_t\right)^{\chi_0}}}$$

$\Updownarrow$

$$\left(1 - l_t\right) = \left[\frac{z_t^{(1-\chi)(1-\chi_0)} n_t \varphi_0}{\left(\frac{c_t - bc_{t-1}}{\left(\widetilde{c}_{ss} z_t\right)^{\chi_0}}\right)^{-\chi} \frac{w_t^*}{\left(\widetilde{c}_{ss} z_t\right)^{\chi_0}}}\right]^{\varphi}$$

This implies that

$$
\begin{aligned}
n_t\varphi_0 \frac{(1-l_t)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}} &= n_t\varphi_0 \frac{\left[\frac{z_t^{(1-\chi)(1-\chi_0)} n_t\varphi_0}{\left(\frac{c_t-bc_{t-1}}{(\widetilde{c}_{ss}z_t)^{\chi_0}}\right)^{-\chi}\frac{w_t^*}{(\widetilde{c}_{ss}z_t)^{\chi_0}}}\right]^{\varphi\left(1-\frac{1}{\varphi}\right)}}{1-\frac{1}{\varphi}} \\
&= n_t^{1+\varphi\left(1-\frac{1}{\varphi}\right)} \frac{\varphi_0^{1+\varphi\left(1-\frac{1}{\varphi}\right)}}{1-\frac{1}{\varphi}} \frac{z_t^{(1-\chi)(1-\chi_0)\varphi\left(1-\frac{1}{\varphi}\right)}}{\left(\frac{c_t-bc_{t-1}}{(\widetilde{c}_{ss}z_t)^{\chi_0}}\right)^{-\chi\varphi\left(1-\frac{1}{\varphi}\right)} \frac{(w_t^*)^{\varphi\left(1-\frac{1}{\varphi}\right)}}{(\widetilde{c}_{ss}z_t)^{\chi_0\varphi\left(1-\frac{1}{\varphi}\right)}}} \\
&= n_t^{1+\varphi-1} \frac{\varphi_0^{1+\varphi-1}}{1-\frac{1}{\varphi}} \frac{z_t^{(1-\chi)(1-\chi_0)\varphi\left(1-\frac{1}{\varphi}\right)}}{\left(\frac{c_t-bc_{t-1}}{(\widetilde{c}_{ss}z_t)^{\chi_0}}\right)^{-\chi\varphi\left(1-\frac{1}{\varphi}\right)} \frac{(w_t^*)^{\varphi\left(1-\frac{1}{\varphi}\right)}}{(\widetilde{c}_{ss}z_t)^{\chi_0\varphi\left(1-\frac{1}{\varphi}\right)}}} \\
&= \frac{(n_t\varphi_0)^{\varphi}}{1-\frac{1}{\varphi}} \frac{z_t^{(1-\chi)(1-\chi_0)\varphi\left(1-\frac{1}{\varphi}\right)} (\widetilde{c}_{ss}z_t)^{\chi_0\varphi\left(1-\frac{1}{\varphi}\right)} \left(\frac{c_t-bc_{t-1}}{(\widetilde{c}_{ss}z_t)^{\chi_0}}\right)^{\chi\varphi\left(1-\frac{1}{\varphi}\right)}}{(w_t^*)_t^{\varphi\left(1-\frac{1}{\varphi}\right)}} \\
&= \frac{(n_t\varphi_0)^{\varphi}}{1-\frac{1}{\varphi}} \frac{z_t^{(1-\chi)(1-\chi_0)(\varphi-1)} (c_t-bc_{t-1})^{\chi(\varphi-1)} (\widetilde{c}_{ss}z_t)^{\chi_0(\varphi-1)(1-\chi)}}{(w_t^*)^{(\varphi-1)}}
\end{aligned}
$$

Hence, we can write the value function as (for negative value function) as

$$
\begin{aligned}
V_t &= u_0 z_t^{(1-\chi)(1-\chi_0)} + d_t\left[\frac{1}{1-\chi}\left(\frac{c_t-bc_{t-1}}{(\widetilde{c}_{ss}z_t)^{\chi_0}}\right)^{1-\chi} + u_0^d \times z_t^{(1-\chi)(1-\chi_0)} + z_t^{(1-\chi)(1-\chi_0)} n_t\varphi_0 \frac{(1-l_t)^{1-\frac{1}{\varphi}}}{1-\frac{1}{\varphi}}\right] \\
&\quad -\beta\left(\mathbb{E}_t\left[(-V_{t+1})^{1-\alpha}\right]\right)^{\frac{1}{1-\alpha}}
\end{aligned}
$$

$\Updownarrow$

$$
\begin{aligned}
V_t &= u_0 z_t^{(1-\chi)(1-\chi_0)} + d_t\left[\begin{array}{c} \frac{1}{1-\chi}\left(\frac{c_t-bc_{t-1}}{(\widetilde{c}_{ss}z_t)^{\chi_0}}\right)^{1-\chi} + u_0^d \times z_t^{(1-\chi)(1-\chi_0)} \\ +z_t^{(1-\chi)(1-\chi_0)} \frac{(n_t\varphi_0)^{\varphi}}{1-\frac{1}{\varphi}} \frac{z_t^{(1-\chi)(1-\chi_0)(\varphi-1)} (c_t-bc_{t-1})^{\chi(\varphi-1)} (\widetilde{c}_{ss}z_t)^{\chi_0(\varphi-1)(1-\chi)}}{(w_t^*)^{(\varphi-1)}} \end{array}\right] \\
&\quad -\beta\left(\mathbb{E}_t\left[(-V_{t+1})^{1-\alpha}\right]\right)^{\frac{1}{1-\alpha}}
\end{aligned}
$$

$\Updownarrow$

$$
\begin{aligned}
V_t &= u_0 z_t^{(1-\chi)(1-\chi_0)} + d_t\left[\begin{array}{c} \frac{1}{1-\chi}\left(\frac{c_t-bc_{t-1}}{(\widetilde{c}_{ss}z_t)^{\chi_0}}\right)^{1-\chi} + u_0^d \times z_t^{(1-\chi)(1-\chi_0)} \\ +\frac{(n_t\varphi_0)^{\varphi}}{1-\frac{1}{\varphi}} \frac{z_t^{(1-\chi)(1-\chi_0)\varphi} (c_t-bc_{t-1})^{\chi(\varphi-1)} (\widetilde{c}_{ss}z_t)^{\chi_0(\varphi-1)(1-\chi)}}{(w_t^*)^{(\varphi-1)}} \end{array}\right] \\
&\quad -\beta\left(\mathbb{E}_t\left[(-V_{t+1})^{1-\alpha}\right]\right)^{\frac{1}{1-\alpha}}
\end{aligned}
$$

We define the timing premium Π_t implicitely as

$$\begin{aligned} V_t &= u_0 \left(z_t \left(1-\Pi_t\right)\right)^{(1-\chi)(1-\chi_0)} \\ &+d_t \left[\frac{1}{1-\chi}\left(\frac{c_t - bc_{t-1}}{\left(\widetilde{c}_{ss} z_t\right)^{\chi_0}}\left(1-\Pi_t\right)\right)^{1-\chi} + u_0^d \times \left(z_t\left(1-\Pi_t\right)\right)^{(1-\chi)(1-\chi_0)}\right] \\ &+d_t \frac{\left(n_t \varphi_0\right)^{\varphi}}{1-\frac{1}{\varphi}} \frac{z_t^{(1-\chi)(1-\chi_0)\varphi}\left(\left(c_t - bc_{t-1}\right)\left(1-\Pi_t\right)\right)^{\chi(\varphi-1)}}{\left(w_t^*\right)^{(\varphi-1)}} \\ &-\beta\left(\mathbb{E}_t\left[\left(\begin{array}{c} -N_{0,t+1}\left(1-\Pi_t\right)^{(1-\chi)(1-\chi_0)} - N_{c,t+1}\left(1-\Pi_t\right)^{(1-\chi)} \\ -N_{d,t+1}\left(1-\Pi_t\right)^{(1-\chi)(1-\chi_0)} - N_{l,t+1}\left(1-\Pi_t\right)^{\chi(\varphi-1)} \end{array}\right)^{1-\alpha}\right]\right)^{\frac{1}{1-\alpha}} \end{aligned}$$

$\Updownarrow$

$$\begin{aligned} V_t &= u_0 \left(z_t \left(1-\Pi_t\right)\right)^{(1-\chi)(1-\chi_0)} \\ &+\left[\frac{d_t}{1-\chi}\left(\frac{\left(c_t - bc_{t-1}\right)}{\left(\widetilde{c}_{ss} z_t\right)^{\chi_0}}\left(1-\Pi_t\right)\right)^{1-\chi} + d_t u_0^d \times \left(z_t\left(1-\Pi_t\right)\right)^{(1-\chi)(1-\chi_0)}\right] \\ &+d_t \frac{\left(n_t \varphi_0\right)^{\varphi}}{1-\frac{1}{\varphi}} \frac{z_t^{(1-\chi)(1-\chi_0)\varphi}\left(\left(c_t - bc_{t-1}\right)\left(1-\Pi_t\right)\right)^{\chi(\varphi-1)}\left(\widetilde{c}_{ss} z_t\right)^{\chi_0(\varphi-1)(1-\chi)}}{\left(w_t^*\right)^{(\varphi-1)}} \\ &-\beta\left(\mathbb{E}_t\left[\left(\begin{array}{c} -N_{0,t+1}\left(1-\Pi_t\right)^{(1-\chi)(1-\chi_0)} - N_{c,t+1}\left(1-\Pi_t\right)^{(1-\chi)} \\ -N_{d,t+1}\left(1-\Pi_t\right)^{(1-\chi)(1-\chi_0)} - N_{l,t+1}\left(1-\Pi_t\right)^{\chi(\varphi-1)} \end{array}\right)^{1-\alpha}\right]\right)^{\frac{1}{1-\alpha}} \end{aligned}$$

Here, $N_{0,t}$ is the continuation value of $u_0 z_t^{(1-\chi)(1-\chi_0)}$ if uncertainty is resolved in the next period. The variable $N_{c,t}$ measures the continuation value of $\frac{d_t}{1-\chi}\left(\frac{c_t - bc_{t-1}}{(\widetilde{c}_{ss} z_t)^{\chi_0}}\right)^{1-\chi}$, while $N_{d,t}$ measures the continuation value of $d_t u_0^d z_t^{(1-\chi)(1-\chi_0)}$. Finally, $N_{l,t}$ is the continuation value of leisure. Note that

$$\begin{aligned} N_{0,t+1} &= u_0 z_{t+1}^{(1-\chi)(1-\chi_0)} + \beta u_0 z_{t+2}^{(1-\chi)(1-\chi_0)} + \beta^2 u_0 z_{t+3}^{(1-\chi)(1-\chi_0)} + ... \\ &= u_0 \sum_{i=1}^{\infty} \beta^{i-1} z_{t+i}^{(1-\chi)(1-\chi_0)} \end{aligned}$$

$$N_{c,t+1} = \sum_{i=1}^{\infty} \beta^{i-1} \frac{d_{t+i}}{1-\chi}\left(\frac{\left(c_{t+i} - bc_{t-1+i}\right)}{\left(\widetilde{c}_{ss} z_{t+i}\right)^{\chi_0}}\right)^{1-\chi}$$

$$N_{d,t+1} = u_0^d \sum_{i=1}^{\infty} \beta^{i-1} d_{t+i} z_{t+i}^{(1-\chi)(1-\chi_0)}$$

and

$$N_{l,t+1} = \sum_{i=1}^{\infty} \beta^{i-1} d_{t+i} \frac{\left(n_{t+i}\varphi_0\right)^{\varphi}}{1-\frac{1}{\varphi}} \frac{z_{t+i}^{(1-\chi)(1-\chi_0)\varphi}\left(c_{t+i} - bc_{t-1+i}\right)^{\chi(\varphi-1)}\left(\widetilde{c}_{ss} z_{t+i}\right)^{\chi_0(\varphi-1)(1-\chi)}}{\left(w_{t+i}^*\right)^{\varphi-1}}$$

where we recall that

$$\log\left(\frac{\mu_{z,t+1}}{\mu_{z,ss}}\right) = \rho_{\mu_z} \log\left(\frac{\mu_{z,t}}{\mu_{z,ss}}\right) + \sigma_{\mu_z}\epsilon_{\mu_z,t+1}$$

$$\log d_{t+1} = \rho_d \log d_t + \sigma_d \epsilon_{d,t+1}$$

$$\log n_{t+1} = \rho_n \log n_t + \sigma_n \epsilon_{n,t+1}$$

$$z_t = z_{t-1}\mu_{z,t}$$

and

$$w_t = \kappa_w \left(\widetilde{w}_{ss} z_t\right) + \left(1 - \kappa_w\right) w_t^*$$

$\Updownarrow$

$$w_t^* = \frac{w_t - \kappa_w \left(\widetilde{w}_{ss} z_t\right)}{1 - \kappa_w}$$

We start the simulations of all these sums at the steady state, i.e. where $z_t = 1$, $n_t = 1$, and $d_t = 1$. Also the timing premium is computed for the economy at the steady state.

3 Data

3.1 Hours

Hours is calculated based on total nonfarm payrolls, which is detrended using the regression filter of Hamilton (2018).

3.2 Wages

For real wages, we use the real hourly compensation for all employed persons in the nonfarm business sector, Index 2012=100, and seasonally adjusted. This data series is then detrended using the regression filter of Hamilton (2018).

3.3 Consumption

Quarterly consumption is for real per capita non-durables and service expenditures, which are available from the Federal Reserve Bank of St. Louis.

3.4 Inflation

Inflation is calculated as the year-on-year growth rate in the seasonal adjusted Consumer Price Index (excluding food and energy) for all urban consumers.

3.5 Bond Yields

We use the 3-month nominal risk-free rate r_t from the secondary market. The 1-year, 3-year, 5-year, 7-year, and 10-year bond yields are from the Gurkaynak, Sack and Wright dataset. This data series is available from 1961Q2.

3.6 Short Rate Expectations from Surveys

Expected future short rates 1,2,3, and 4 quarters ahead are taken from the survey of Professionel Forecasts for the 3-month T bill. Here, we use the median from panel of forecasts in the survey Professionel Forecasts. These survey data are not used in the paper, but they are included in the matlab codes linked to the paper.

3.7 Stock Market Data

Dividend and market return series are constructed from two CRSP series, the value-weighted index including distributions (VWRD) and the value-weighted index excluding distributions (VWRX). A price index is constructed as

$$P_{s+1} = P_s \left(1 + VWRX_{s+1}\right), \tag{11}$$

with $P_0 = 1$, where s denotes a monthly time index. The related dividends are then calculated as

$$D_{s+1} = P_s \left(VWRD_{s+1} - VWRX_{s+1}\right). \tag{12}$$

To remove seasonality effects in monthly dividends, we construct aggregate dividends $D_s^a = \sum_{i=0}^{11} D_{s-i}$. Using this series, dividend growth is calculated as

$$\Delta d_s = \log D_s^a - \log D_{s-1}^a, \quad (13)$$

and then summing these monthly observations over the particular quarter to get quarterly dividends. The market return each month is

$$r_s^m = \log (P_s + D_s) - \log P_{s-1}, \quad (14)$$

which we annualize. Finally the price-dividend ratio is

$$p_s - d_s = \log P_s - \log D_s^a. \quad (15)$$

We then obtain quarterly time series of the price-dividend ratio by using values at the end of each quarter.

3.8 Related Model Variables

1. Hours

We have

$$\hat{l}_t^{Data} = \hat{l}_t^{Model},$$

where we use the standard notation that a "hat" denotes deviation from the steady state, i.e. $\hat{l}_t = \log \left(\frac{l_t}{l_{ss}} \right)$.

2. Wages

We have

$$\hat{w}_t^{Data} = \widehat{\widetilde{w}}_t^{Model}$$

3. Consumption growth

The expressions for real quarterly consumption growth is given by

$$\begin{aligned} \Delta c_t &\equiv \log \frac{c_t}{c_{t-1}} = \log \frac{\widetilde{c}_t z_t}{\widetilde{c}_{t-1} z_{t-1}} = \log \frac{\widetilde{c}_t}{\widetilde{c}_{ss}} - \log \frac{\widetilde{c}_{t-1}}{\widetilde{c}_{ss}} + \log \frac{z_t}{z_{t-1}} \\ &= \widehat{\widetilde{c}}_t - \widehat{\widetilde{c}}_{t-1} + \log \left(\mu_{z,t} \right) \end{aligned}$$

because $\widehat{\widetilde{c}}_t \equiv \log \frac{\widetilde{c}_t}{\widetilde{c}_{ss}}$.

4. Inflation

The quarterly inflation rate is given by

$$\log \pi_t = \log \pi_{ss} + \hat{\pi}_t$$

where $\hat{\pi}_t = \log \left(\frac{\pi_t}{\pi_{ss}} \right)$

5. The Nominal Yield Curve

All yields are given by

$$r_t^{(k)} = r_{ss} + \hat{r}_t^{(k)},$$

where k denotes the maturity. Note that $r_t^{(1)} \equiv r_t$.

6. Surveys of Expected future short rates

$$\mathbb{E}_t \left[r_{t+i} \right] = r_{ss} + \widehat{\mathbb{E}_t \left[r_{t+i} \right]}$$

where $i = \{1, 2, 3, 4\}$.

Note finally, that all variables are annualized through a multiplication by 4, except for hours and wages.

4 Estimation Methodology: Filtering with Shrinkage

This section presents the estimation methodology used in the paper.

4.1 The Estimator

To describe the considered estimator, let $\boldsymbol{\theta}$ denote the structural parameters of dimension $n_\theta \times 1$. We would like to estimate $\boldsymbol{\theta}$ using the observables $\mathbf{y}_t^{obs}$ with dimension $n_y \times 1$. Given the presence of latent variables $\mathbf{x}_t$ of dimension $n_x \times 1$ in the model, a log-likelihood function $\sum_{t=1}^{T} \mathcal{L}_t^{CDKF}$ is evaluated by Kalman filtering. Due to the nonlinear structure of our state space system, we evaluate a quasi-log likelihood function by the central difference Kalman filter of Norgaard, Poulsen & Ravn (2000) as studied in Andreasen (2013) within the context of nonlinear DSGE models.

To robustify the estimation of $\boldsymbol{\theta}$ from a quasi log-likelihood function, we propose to shrink the estimation of $\boldsymbol{\theta}$ towards a set of unconditonal moments. The empirical moments are denoted by $\frac{1}{T}\sum_{t=1}^{T}\mathbf{m}_t$ and the model-implied moments by $\mathbb{E}\left[\mathbf{m}\left(\boldsymbol{\theta}\right)\right]$, which both have dimension $n_m \times 1$. Hence, the contribution from these shrinkage moments are given by

$$
\begin{aligned}
Q &= \left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{m}_t - \mathbb{E}\left[\mathbf{m}\left(\boldsymbol{\theta}\right)\right]\right)' \mathbf{W} \left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{m}_t - \mathbb{E}\left[\mathbf{m}\left(\boldsymbol{\theta}\right)\right]\right) \\
&= \mathbf{g}_{1:T}\left(\boldsymbol{\theta}\right)' \mathbf{W} \mathbf{g}_{1:T}\left(\boldsymbol{\theta}\right)
\end{aligned}
$$

where $\mathbf{g}_{1:T}\left(\boldsymbol{\theta}\right) \equiv \frac{1}{T}\sum_{t=1}^{T}\mathbf{g}_t\left(\boldsymbol{\theta}\right)$ with $\mathbf{g}_t\left(\boldsymbol{\theta}\right) \equiv \left(\mathbf{m}_t - \mathbb{E}\left[\mathbf{m}\left(\boldsymbol{\theta}\right)\right]\right)$ and $\mathbf{W}$ is a weigthing matrix. Here, we consider the case where $\mathbf{W}$ is diagonal and its elements are given by the inverse of the standard error attached to each of the moments in $\mathbf{m}_T$.

Hence, the considered estimator is given by

$$
\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{Max} \ \frac{1}{T}\sum_{t=1}^{T}\mathcal{L}_t^{CDKF} - \lambda \mathbf{g}_{1:T}\left(\boldsymbol{\theta}\right)' \mathbf{W} \mathbf{g}_{1:T}\left(\boldsymbol{\theta}\right) \tag{16}
$$

where $\lambda \in \mathbb{R}_+$ controls the weight given to the skrinkage moments. For the implemention in Matlab, where we use minimization routines, we use the equivalent formulation

$$
\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{Min} \ -\frac{1}{T}\mathcal{L}_t^{CDKF} + \lambda \mathbf{g}_{1:T}\left(\boldsymbol{\theta}\right)' \mathbf{W} \mathbf{g}_{1:T}\left(\boldsymbol{\theta}\right).
$$

4.2 The Unconditional Shrinkage Moments

To describe the unconditional moments included in the estimation, consider the following nine variables:

$$
\underbrace{\begin{bmatrix} \hat{l}_t \\ \hat{w}_t \\ \Delta c_t \\ \log \pi_t \\ r_t^{(i)} \end{bmatrix}}_{\mathbf{y}_t^{mom}}
$$

where $i = \{1, 4, 12, 20, 28, 40\}$. The first set of moments we include contains the first and second uncentered moments for $\mathbf{y}_t^{mom}$, that is

$$
\frac{1}{T}\sum_{t=1}^{T}\mathbf{m}_{1,t} = \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T}\mathbf{y}_t^{mom} \\ \frac{1}{T}\sum_{t=1}^{T}\left(\mathbf{y}_t^{mom}\right)^2 \end{bmatrix}.
$$

To describe the second set of moments, consider the Campbell-Shiller regression

$$
r_{t+m}^{(k-m)} - r_t^{(k)} = \alpha_k + \beta_k \frac{m}{k-m}\left(r_t^{(k)} - r_t^{(m)}\right) + u_t^{(k)}
$$

where m indicates the forecast horizon. Here, we apply $m = 4$ for forecasting four quarters ahead. The population value for β_k is given by

$$\begin{aligned}\beta_k &= \frac{Cov\left(r_{t+m}^{(k-m)} - r_t^{(k)}, \frac{m}{k-m}\left(r_t^{(k)} - r_t^{(m)}\right)\right)}{Var\left(\frac{m}{k-m}\left(r_t^{(k)} - r_t^{(m)}\right)\right)} \\ &= \frac{\frac{m}{k-m}Cov\left(r_{t+m}^{(k-m)} - r_t^{(k)}, r_t^{(k)} - r_t^{(m)}\right)}{\left(\frac{m}{k-m}\right)^2 Var\left(r_t^{(k)} - r_t^{(m)}\right)} \\ &= \frac{Cov\left(r_{t+m}^{(k-m)} - r_t^{(k)}, r_t^{(k)} - r_t^{(m)}\right)}{\frac{m}{k-m}Var\left(r_t^{(k)} - r_t^{(m)}\right)}\end{aligned}$$

We represent β_k by including moments for the covariance and the variance determining β_k. That is, we target

$$\frac{1}{T}\sum_{t=1}^T \mathbf{m}_{2,t} = \begin{bmatrix} \frac{1}{T}\sum_{t=1}^T \mathbf{m}_{2,t}^{Cov} \\ \frac{1}{T}\sum_{t=1}^T \mathbf{m}_{2,t}^{VAR} \end{bmatrix}$$

where

$$\frac{1}{T}\sum_{t=1}^T \mathbf{m}_{2,t}^{Cov} = \left\{\frac{1}{T}\sum_{t=1}^{T-m}\left(r_{t+m}^{(k-m)} - r_t^{(k)}\right)\left(\left(r_t^{(k)} - r_t^{(m)}\right) - \overline{slope_t}^{(k)}\right)\right\}_{k=\{8,16,...,40\}}'$$

and

$$\frac{1}{T}\sum_{t=1}^T \mathbf{m}_{2,t}^{VAR} = \left\{\frac{1}{T}\sum_{t=1}^{T-m}\left(r_t^{(k)} - r_t^{(m)}\right)\left(\left(r_t^{(k)} - r_t^{(m)}\right) - \overline{slope_t}^{(k)}\right)\right\}_{k=\{8,16,...,40\}}',$$

with $\overline{slope_t}^{(k)} \equiv \frac{1}{T}\sum_{t=1}^T\left(r_t^{(k)} - r_t^{(m)}\right)$ for $k = \{8, 16, ..., 40\}$.

Thus, we consider the following unconditional moments for the skrinkage part of our estimator

$$\frac{1}{T}\sum_{t=1}^T \mathbf{m}_t \equiv \begin{bmatrix} \frac{1}{T}\sum_{t=1}^T \mathbf{m}_{1,t} \\ \frac{1}{T}\sum_{t=1}^T \mathbf{m}_{2,t} \end{bmatrix}.$$

4.3 Computing the Shrinkage Moments

In general, $\mathbb{E}\left[\mathbf{m}\left(\boldsymbol{\theta}\right)\right]$ and hence Q must be computed by simulation. However, for the pruned approximation (and a linearized approximation), it is possible to compute $\mathbb{E}\left[\mathbf{m}\left(\boldsymbol{\theta}\right)\right]$ in closed form. When the standard perturbation approximation is used, we can easily compute the closed-form expression for $\mathbb{E}\left[\mathbf{m}\left(\boldsymbol{\theta}\right)\right]$ using the results in Andreasen et al. (2018). For the second-order projection approximation applied in the paper, we extend the results of Andreasen et al. (2018) to this approximation, such that it is also possible to obtain a closed form expression for $\mathbb{E}\left[\mathbf{m}\left(\boldsymbol{\theta}\right)\right]$ in this case.

The value of $\mathbb{E}\left[\mathbf{m}_1\left(\boldsymbol{\theta}\right)\right]$ follows directly from the unconditional mean of a control variable $\mathbf{y}_t$ and its covariance matrix. For $\mathbb{E}\left[\mathbf{m}_2\left(\boldsymbol{\theta}\right)^{Cov}\right]$, we first observe for the kth element of $\mathbf{m}_2\left(\boldsymbol{\theta}\right)^{Cov}$ that

$$\begin{aligned}\mathbb{E}\left[m_2\left(\boldsymbol{\theta}\right)_k^{Cov}\right] &= \mathbb{E}\left[\left(r_{t+m}^{(k-m)} - r_t^{(k)}\right)\left(\left(r_t^{(k)} - r_t^{(m)}\right) - \mathbb{E}\left[slope_t^{(k)}\right]\right)\right] \\ &= \mathbb{E}\left[\left(\left(r_{t+m}^{(k-m)} - r_t^{(k)}\right) - \mathbb{E}\left[r_{t+m}^{(k-m)} - r_t^{(k)}\right]\right)\left(slope_t^{(k)} - \mathbb{E}\left[slope_t^{(k)}\right]\right)\right] \\ &= Cov\left[r_{t+m}^{(k-m)} - r_t^{(k)}, slope_t^{(k)}\right] \\ &= Cov\left[r_{t+m}^{(k-m)}, slope_t^{(k)}\right] - Cov\left[r_t^{(k)}, slope_t^{(k)}\right]\end{aligned}$$

Hence, $Cov\left[r_t^{(k)}, slope_t^{(k)}\right]$ can be computed directly using the expression for unconditional second moments for control variables, and $Cov\left[r_{t+m}^{(k-m)}, slope_t^{(k)}\right]$ follows from the auto-covariance matrix for control variables. As for $\mathbb{E}\left[\mathbf{m}_2\left(\boldsymbol{\theta}\right)^{VAR}\right]$ we have

$$\begin{aligned}\mathbb{E}\left[m_2\left(\boldsymbol{\theta}\right)_k^{VAR}\right] &= \mathbb{E}\left[\left(r_t^{(k)} - r_t^{(m)}\right)\left(\left(r_t^{(k)} - r_t^{(m)}\right) - \mathbb{E}\left[slope_t^{(k)}\right]\right)\right] \\ &= \mathbb{E}\left[\left(\left(r_t^{(k)} - r_t^{(m)}\right) - \mathbb{E}\left[slope_t^{(k)}\right]\right)\left(\left(r_t^{(k)} - r_t^{(m)}\right) - \mathbb{E}\left[slope_t^{(k)}\right]\right)\right] \\ &= Var\left[\left(r_t^{(k)} - r_t^{(m)}\right)\right],\end{aligned}$$

which also follows directly using the expression for unconditional second moments for control variables when applied to the slope of the yield curve.

4.4 The Link Between Campbell-Shiller and Return Regressions

We first introduce some notation. Let $B_t^{(k)}$ denote the price of a zero-coupon bond with maturity k at time t. The m-period holding period return on a k-period bond is

$$\begin{aligned}\widetilde{hpr}_{t+m}^{(k)} &\equiv \log\left(\frac{B_{t+m}^{(k-m)}}{B_t^{(k)}}\right) \\ &= \log B_{t+m}^{(k-m)} - \log B_t^{(k)} \\ &= -\left(k-m\right) r_{t+m}^{(k-m)} + k r_t^{(k)},\end{aligned}$$

because $r_{t,}^{(k)} = -\frac{1}{k}\log B_t^{(k)}$. One period in our model corresponds to one quarter, so it is natural to express everything in quarterly returns. That is, we get quarterly returns as

$$hpr_{t+m}^{(k)} = -\frac{k-m}{4} r_{t+m}^{(k-m)} + \frac{k}{4} r_t^{(k)}$$

The Campbell-Shiller regression is given by

$$r_{t+m}^{(k-m)} - r_t^{(k)} = \alpha_k + \beta_k \frac{m}{k-m}\left(r_t^{(k)} - r_t^{(m)}\right) + u_{t+m,k},$$

The classic return regression is given by

$$hpr_{t+m}^{(k)} - \frac{m}{4} r_t^{(m)} = \tilde{\alpha}_k + \tilde{\beta}_k\left(r_t^{(k)} - r_t^{(m)}\right) + \varepsilon_{t+m,k}$$

To see the link between these two regressions, consider the quarterly holding period returns

$$hpr_{t+m}^{(k)} = -\frac{k-m}{4} r_{t+m}^{(k-m)} + \frac{k}{4} r_t^{(k)}.$$

Now subtract $\frac{m}{4} r_t^{(m)}$ on both sides to obtain

$$hpr_{t+m}^{(k)} - \frac{m}{4} r_t^{(m)} = -\frac{k-m}{4} r_{t+m}^{(k-m)} + \frac{k}{4} r_t^{(k)} - \frac{m}{4} r_t^{(m)}.$$

Then add and subtract $\frac{m}{4} r_t^{(k)}$ on the right hand side

$$\begin{aligned}hpr_{t+m}^{(k)} - \frac{m}{4} r_t^{(m)} &= -\frac{k-m}{4} r_{t+m}^{(k-m)} + \frac{k}{4} r_t^{(k)} - \frac{m}{4} r_t^{(m)} + \frac{m}{4} r_t^{(k)} - \frac{m}{4} r_t^{(k)} \\ &= -\frac{k-m}{4} r_{t+m}^{(k-m)} + \frac{k}{4} r_t^{(k)} + \frac{m}{4}\left(r_t^{(k)} - r_t^{(m)}\right) - \frac{m}{4} r_t^{(k)} \\ &= -\frac{k-m}{4} r_{t+m}^{(k-m)} + \frac{k-m}{4} r_t^{(k)} + \frac{m}{4}\left(r_t^{(k)} - r_t^{(m)}\right) \\ &= -\frac{k-m}{4}\left(r_{t+m}^{(k-m)} - r_t^{(k)}\right) + \frac{m}{4}\left(r_t^{(k)} - r_t^{(m)}\right).\end{aligned}$$

Multiplying both sides by $\frac{4}{k-m}$ and rearranging, we get

$$\frac{4}{k-m}\left(hpr_{t+m}^{(k)}-\frac{m}{4}r_t^{(m)}\right)=\frac{4}{k-m}\left(-\frac{k-m}{4}\left(r_{t+m}^{(k-m)}-r_t^{(k)}\right)+\frac{m}{4}\left(r_t^{(k)}-r_t^{(m)}\right)\right)$$

$\Updownarrow$

$$\frac{4}{k-m}\left(hpr_{t+m}^{(k)}-\frac{m}{4}r_t^{(m)}\right)=\left(-\frac{k-m}{k-m}\left(r_{t+m}^{(k-m)}-r_t^{(k)}\right)+\frac{m}{k-m}\left(r_t^{(k)}-r_t^{(m)}\right)\right)$$

$\Updownarrow$

$$\frac{4}{k-m}\left(hpr_{t+m}^{(k)}-\frac{m}{4}r_t^{(m)}\right)=\frac{m}{k-m}\left(r_t^{(k)}-r_t^{(m)}\right)-\left(r_{t+m}^{(k-m)}-r_t^{(k)}\right). \tag{17}$$

Regressing each of the three terms in (17) on a constant and $\frac{m}{k-m}\left(r_t^{(k)}-r_t^{(m)}\right)$ we get:

i) $\left(hpr_{t+m}^{(k)}-\frac{m}{4}r_t^{(m)}\right)=\tilde{\alpha}_k+\left(\tilde{\beta}_k\frac{k-m}{m}\right)\frac{m}{k-m}\left(r_t^{(k)}-r_t^{(m)}\right)$

ii) $\left(r_t^{(k)}-r_t^{(m)}\right)=0+\left(1\frac{k-m}{m}\right)\times\frac{m}{k-m}\left(r_t^{(k)}-r_t^{(m)}\right)$

iii) $\left(r_{t+m}^{(k-m)}-r_t^{(k)}\right)=\alpha_k+\beta_k\frac{m}{k-m}\left(r_t^{(k)}-r_t^{(m)}\right)$

Thus, for the constants, we simply insert the intercepts from i) to iii) into (17). That is

$$\frac{4}{k-m}\times\text{intercept from i)}=\frac{m}{k-m}\times\text{intercept from ii)}\ -\text{intercept from iii)}$$

$\Updownarrow$

$$\frac{4}{k-m}\times\tilde{\alpha}_k=\frac{m}{k-m}\times 0-\alpha_k$$

$\Updownarrow$

$$\tilde{\alpha}_k=-\alpha_k\frac{k-m}{4}$$

For the loadings, we get

$$\frac{4}{k-m}\times\text{slope from i)}=\frac{m}{k-m}\times\text{slope from ii)}\ -\text{slope from iii)}$$

$$\frac{4}{k-m}\tilde{\beta}_k\frac{k-m}{m}=\frac{m}{k-m}\times\frac{k-m}{m}-\beta_k$$

$\Updownarrow$

$$\frac{4}{m}\tilde{\beta}_k=1-\beta_k$$

Thus the two regressions contain the same information.

For the case of $m=4$, we thus have the simple mappings $\tilde{\alpha}_k=-\alpha_k\frac{k-4}{4}=-\alpha_k\left(\frac{k}{4}-1\right)$ and $\tilde{\beta}_k=1-\beta_k$.

4.5 Asymptotic Distribution

The asymptotic analysis is carried out based on the assumption that any filtering errors caused solely due to the adopted integration approximation in the CDKF are small and none essential for the estimates. The validity of this assumption is explored in the Monte Carlo study presented in this Online Appendix. Given this assumption, the standard QML estimator based on the CDKF is consistent (see Andreasen (2013)). We also have that GMM is consistent, given standard regularity conditions (see Hansen (1982)), and therefore, the proposed estimator with shrinkage is also consistent with respect to $\boldsymbol{\theta}$.

To derive the asymptotic distribution of $\hat{\boldsymbol{\theta}}$ for $T \longrightarrow \infty$, let

$$\mathcal{L} = \frac{1}{T}\sum_{t=1}^{T} \mathcal{L}_t^{CDKF} - \lambda Q,$$

and note that

$$\begin{aligned}
\underset{n_\theta \times 1}{\frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}}} &= \frac{1}{T}\sum_{t=1}^{T} \underset{n_\theta \times 1}{\frac{\partial \mathcal{L}_t^{CDKF}}{\partial \boldsymbol{\theta}}} - 2\lambda \left(\underset{n_m \times n_\theta}{\frac{\mathbb{E}\left[\mathbf{g}_{1:T}(\boldsymbol{\theta})\right]}{\partial \boldsymbol{\theta}}} \right)' \underset{n_m \times n_m}{\mathbf{W}} \underset{n_m \times 1}{\mathbf{g}_{1:T}(\boldsymbol{\theta})} \\
&= \frac{1}{T}\sum_{t=1}^{T} \left(\mathbf{s}_t(\boldsymbol{\theta}) - 2\lambda \mathbf{G}(\boldsymbol{\theta})' \mathbf{W} \mathbf{g}_t(\boldsymbol{\theta}) \right)
\end{aligned}$$

where $\mathbf{s}_t(\boldsymbol{\theta}) \equiv \frac{\partial \mathcal{L}_t^{CDKF}}{\partial \boldsymbol{\theta}}$ is the score function for observation t related to the Central Difference Kalman filter and $\mathbf{G}(\boldsymbol{\theta}) \equiv \frac{\partial \mathbf{g}_{1:T}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} = \frac{1}{T}\sum_{t=1}^{T} \frac{\mathbf{g}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}$ is the Jacobian related to the shrinkage moments. The first-order conditions are

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}} = \frac{1}{T}\sum_{t=1}^{T} \left(\mathbf{s}_t\left(\hat{\boldsymbol{\theta}}\right) - 2\lambda \mathbf{G}\left(\hat{\boldsymbol{\theta}}\right)' \mathbf{W} \mathbf{g}_t\left(\hat{\boldsymbol{\theta}}\right) \right) = 0$$

A mean-value expansion of $\mathbf{s}_t$ and $\mathbf{g}_t$ around the true value $\boldsymbol{\theta}_o$ gives

$$\frac{1}{T}\sum_{t=1}^{T} \left\{ \mathbf{s}_t(\boldsymbol{\theta}_o) + \tilde{\mathbf{H}}_t \left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o\right) \right\} - 2\lambda \mathbf{G}\left(\hat{\boldsymbol{\theta}}\right)' \mathbf{W} \left\{ \mathbf{g}_{1:T}(\boldsymbol{\theta}_o) + \tilde{\mathbf{G}} \left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o\right) \right\} = \mathbf{0}.$$

Here, $\tilde{\mathbf{H}}_t$ is the $n_\theta \times n_\theta$ Hessian matrix of observation t related to the Kalman filtering. The tilde on the Hessian matrix indicates that each row of $\tilde{\mathbf{H}}_t$ is evaluated at a different mean value, which is on the line segment between $\boldsymbol{\theta}_o$ and $\hat{\boldsymbol{\theta}}$. Similarly, $\tilde{\mathbf{G}}$ is the Jacobian related to the shrinkage moments where each row of $\tilde{\mathbf{G}}$ is evaluated at a different mean value, which is on the line segment between $\boldsymbol{\theta}_o$ and $\hat{\boldsymbol{\theta}}$. Thus, we get

$$\frac{1}{T}\sum_{t=1}^{T} \left(\mathbf{s}_t(\boldsymbol{\theta}_o) + \tilde{\mathbf{H}}_t \left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o\right) \right) - 2\lambda \mathbf{G}\left(\hat{\boldsymbol{\theta}}\right)' \mathbf{W} \left\{ \mathbf{g}_{1:T}(\boldsymbol{\theta}_o) + \tilde{\mathbf{G}} \left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o\right) \right\} = \mathbf{0}$$

$\Updownarrow$

$$\frac{1}{T}\sum_{t=1}^{T} \mathbf{s}_t(\boldsymbol{\theta}_o) - 2\lambda \mathbf{G}\left(\hat{\boldsymbol{\theta}}\right)' \mathbf{W} \left(\frac{1}{T}\sum_{t=1}^{T} \mathbf{g}_t(\boldsymbol{\theta}_o) \right) + \left(\frac{1}{T}\sum_{t=1}^{T} \tilde{\mathbf{H}}_t - 2\lambda \mathbf{G}\left(\hat{\boldsymbol{\theta}}\right)' \mathbf{W} \tilde{\mathbf{G}} \right) \left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o\right) = \mathbf{0}$$

because $\mathbf{g}_{1:T}(\boldsymbol{\theta}) \equiv \frac{1}{T}\sum_{t=1}^{T} \mathbf{g}_t(\boldsymbol{\theta})$. Thus

$$\frac{1}{T}\sum_{t=1}^{T} \left(-\mathbf{s}_t(\boldsymbol{\theta}_o) + 2\lambda \mathbf{G}\left(\hat{\boldsymbol{\theta}}\right)' \mathbf{W} \mathbf{g}_t(\boldsymbol{\theta}_o) \right) = \left(\frac{1}{T}\sum_{t=1}^{T} \tilde{\mathbf{H}}_t - 2\lambda \mathbf{G}\left(\hat{\boldsymbol{\theta}}\right)' \mathbf{W} \tilde{\mathbf{G}} \right) \left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o\right)$$

$\Updownarrow$

$$\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o\right) = \left(\frac{1}{T}\sum_{t=1}^{T} \tilde{\mathbf{H}}_t - 2\lambda \mathbf{G}\left(\hat{\boldsymbol{\theta}}\right)' \mathbf{W} \tilde{\mathbf{G}} \right)^{-1} \frac{1}{T}\sum_{t=1}^{T} \left(-\mathbf{s}_t(\boldsymbol{\theta}_o)' + 2\lambda \mathbf{G}\left(\hat{\boldsymbol{\theta}}\right)' \mathbf{W} \mathbf{g}_t(\boldsymbol{\theta}_o) \right)$$

For sufficiently large T, we have

$$\frac{1}{T}\sum_{t=1}^{T} \tilde{\mathbf{H}}_t - 2\lambda \mathbf{G}\left(\hat{\boldsymbol{\theta}}\right)' \mathbf{W} \tilde{\mathbf{G}} \xrightarrow{p} \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} \mathbf{H}_t(\boldsymbol{\theta}_o) - 2\lambda \mathbf{G}(\boldsymbol{\theta}_o)' \mathbf{W} \mathbf{G}(\boldsymbol{\theta}_o) \right] \equiv \mathbf{A}_o.$$

Also, let $\mathbf{q}_t(\boldsymbol{\theta}) \equiv -\mathbf{s}_t(\boldsymbol{\theta}) + 2\lambda \mathbf{G}(\boldsymbol{\theta})' \mathbf{W} \mathbf{g}_{\mathbf{t}}(\boldsymbol{\theta})$, then given sufficient regularity conditions, we have that $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \mathbf{q}_t(\boldsymbol{\theta})$ converges to a multivariate normal distribution for $T \longrightarrow \infty$ (see for instance Hansen (1982)). That is,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \mathbf{q}_t\left(\hat{\boldsymbol{\theta}}\right) \xrightarrow{d} \mathcal{N}\left(\mathbb{E}\left[\mathbf{q}_t(\boldsymbol{\theta}_o)\right], Var\left(\mathbf{q}_t(\boldsymbol{\theta}_o)\right)\right)$$

where $\mathbb{E}\left[\mathbf{q}_t(\boldsymbol{\theta}_o)\right] = \mathbf{0}$. To realize that $\mathbb{E}\left[\mathbf{q}_t(\boldsymbol{\theta}_o)\right] = \mathbf{0}$ recall that $\boldsymbol{\theta}_o$ solves the population problem

$$\underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}}{Max} \ \ \mathbb{E}\left[\frac{1}{T} \sum_{t=1}^{T} \mathcal{L}_t^{CDKF} - \lambda \mathbf{g}_{1:T}(\boldsymbol{\theta})' \mathbf{W} \mathbf{g}_{1:T}(\boldsymbol{\theta})\right],$$

which implies the first-order conditions

$$\frac{\partial}{\partial \boldsymbol{\theta}} \mathbb{E}\left[\frac{1}{T} \sum_{t=1}^{T} \mathcal{L}_t^{CDKF} - \lambda \mathbf{g}_{1:T}(\boldsymbol{\theta})' \mathbf{W} \mathbf{g}_{1:T}(\boldsymbol{\theta})\right]\Bigg|_{\boldsymbol{\theta} = \boldsymbol{\theta}_o} = \mathbf{0}.$$

Given sufficient regularity conditions such that the derivative and the expectation operator can be interchanged, we have

$$\mathbb{E}\left[\frac{1}{T} \sum_{t=1}^{T} \left(\mathbf{s}_t(\boldsymbol{\theta}) - 2\lambda \mathbf{G}(\boldsymbol{\theta})' \mathbf{W} \mathbf{g}_t(\boldsymbol{\theta})\right)\right]\Bigg|_{\boldsymbol{\theta} = \boldsymbol{\theta}_o} = \mathbf{0}.$$

$\Updownarrow$

$$\mathbb{E}\left[\mathbf{q}_t(\boldsymbol{\theta})\right]\big|_{\boldsymbol{\theta} = \boldsymbol{\theta}_o} = \mathbf{0}$$

as desired. Thus, given standard regularity conditions, as stated in Hayashi (2000), we have

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \mathbf{A}_o^{-1} Var\left(\mathbf{q}_t(\boldsymbol{\theta}_o)\right) \mathbf{A}_o^{-1}\right)$$

where $\mathbf{q}_t(\boldsymbol{\theta}) \equiv -\mathbf{s}_t(\boldsymbol{\theta}) + 2\lambda \mathbf{G}(\boldsymbol{\theta})' \mathbf{W} \mathbf{g}_t(\boldsymbol{\theta})$. Thus, the asymptotic covariance matric is given by

$$Var\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o\right) = \frac{1}{T} \mathbf{A}_o^{-1} Var\left(\mathbf{q}_t(\boldsymbol{\theta}_o)\right) \mathbf{A}_o^{-1}.$$

We can consistently estimate $\mathbf{A}_o$ by

$$\hat{\mathbf{A}} = \frac{1}{T} \sum_{t=1}^{T} \mathbf{H}_t\left(\hat{\boldsymbol{\theta}}\right) - 2\lambda \mathbf{G}\left(\hat{\boldsymbol{\theta}}\right)' \mathbf{W} \mathbf{G}\left(\hat{\boldsymbol{\theta}}\right)$$

As for $Var\left(\mathbf{q}_t(\boldsymbol{\theta}_o)\right)$, we can use standard estimators to account for autocorrelation and heteroskedasticity in a time series, which is generally needed because of autocorrelation and heteroskedasticity in $\mathbf{g}_t(\boldsymbol{\theta}_o)$. We use the estimator of Newey & West (1987), i.e.

$$\widehat{Var}\left(\mathbf{q}_t\left(\hat{\boldsymbol{\theta}}\right)\right) = \hat{\boldsymbol{\Gamma}}_0 + \sum_{\nu=1}^{k} \left(\hat{\boldsymbol{\Gamma}}_\nu + \hat{\boldsymbol{\Gamma}}_\nu'\right)$$

where

$$\hat{\boldsymbol{\Gamma}}_\nu = \frac{1}{T} \sum_{t=\nu+1}^{T} \left(\mathbf{q}_t\left(\hat{\boldsymbol{\theta}}\right) - \bar{\mathbf{q}}_t\left(\hat{\boldsymbol{\theta}}\right)\right)\left(\mathbf{q}_{t-\nu}\left(\hat{\boldsymbol{\theta}}\right) - \bar{\mathbf{q}}_t\left(\hat{\boldsymbol{\theta}}\right)\right)'$$

and k is a tuning parameter.

In terms of the specific value of λ, we suggest simply to let $\lambda = T$, i.e. the sample size.

4.6 Alternative Interpretation of the Proposed Estimator

One may alternatively consider the proposed estimator as belonging to the class of Laplace type estimators (LTE) or quasi-Bayesian estimators in Chernozhukov & Hong (2003) with the endogenours prior specification of Christiano, Trabandt & Walentin (2011). The use of LTE implies that a potentially misspecified log-likelihood function (as considered in our case) may be used within a Bayesian setting, whereas such a misspecification is generally not accommodated within a standard Bayesian framework.

To realize this, let $L_T(\boldsymbol{\theta}) \equiv \frac{1}{T}\sum_{t=1}^{T}\mathcal{L}_t^{CDKF}$ be the considered extremum statistic and let the priors be denoted by $\pi(\boldsymbol{\theta})$. The quasi-posterior density $p_T(\boldsymbol{\theta})$ for the estimator in Chernozhukov & Hong (2003) is proportional to

$$p_T(\boldsymbol{\theta}) \propto \exp\{L_T(\boldsymbol{\theta})\}\pi(\boldsymbol{\theta})$$

$\Updownarrow$

$$\log p_T(\boldsymbol{\theta}) \propto L_T(\boldsymbol{\theta}) + \log\pi(\boldsymbol{\theta}).$$

For the priors we use the endogenours prior specifiction in Christiano et al. (2011) based on sample moments $\mathbf{m}_t$. For a sufficiently large pre-sample size T, it follows that the density of the empirical sample moments $\mathbf{m}_T = \frac{1}{T}\sum_{t=1}^{T}\mathbf{m}_t$ is given by (see Appendix B (page 38) related to Christiano et al. (2011))

$$p(\mathbf{m}_T|\boldsymbol{\theta}) = \left(\frac{T}{2\pi}\right)^{T/2}\left|\hat{\mathbf{S}}\right|^{-1/2}\exp\left\{-T/2\left(\mathbf{m}_T - \mathbb{E}[\mathbf{m}(\boldsymbol{\theta})]\right)'\hat{\mathbf{S}}^{-1}\left(\mathbf{m}_T - \mathbb{E}[\mathbf{m}(\boldsymbol{\theta})]\right)\right\},$$

where $\hat{\mathbf{S}}$ is the estimated co-variance matrix of the sample moments. Letting $p(\boldsymbol{\theta})$ denote the primitive priors before observing $\mathbf{m}_T$, then

$$\pi(\boldsymbol{\theta}) = p(\mathbf{m}_T|\boldsymbol{\theta})\,p(\boldsymbol{\theta}),$$

or simply

$$\pi(\boldsymbol{\theta}) = p(\mathbf{m}_T|\boldsymbol{\theta}),$$

when $p(\boldsymbol{\theta})$ is set to flat priors. Accordingly

$$\begin{aligned}\log p_T(\boldsymbol{\theta}) &\propto L_T(\boldsymbol{\theta}) + \log\pi(\boldsymbol{\theta}) \\ &\propto L_T(\boldsymbol{\theta}) - T/2\left(\mathbf{m}_T - \mathbb{E}[\mathbf{m}(\boldsymbol{\theta})]\right)'\hat{\mathbf{S}}^{-1}\left(\mathbf{m}_T - \mathbb{E}[\mathbf{m}(\boldsymbol{\theta})]\right) \\ &= L_T(\boldsymbol{\theta}) - T/2 \times Q\end{aligned}$$

when $\mathbf{W} = \hat{\mathbf{S}}^{-1}$, which is identical to (16) when $\lambda = T/2$.

4.7 Monte Carlo Evidence

This section presents a small simulation study to illustrate the benefit of introducing shrinkage when the New Keynesian model is misspecified. To make the simulation study computational feasible, we consider a reduced version of the New Keynesian model. That is, we omit preference shocks, labor supply shocks, and shocks to the inflation target. Also, we omit consumption habits ($b = 0$) and we do not include wage stickiness ($\kappa_w = 0$), and therefore we also omit wages as an observable in the estimation. The model is solved by third-order perturbation, as the issue related to the accuracy of the solution is not essential in this context. That is, the data generating process is given by a third-order perturbation approximation, which is also used for the estimation. We estimate the parameters listed below, while the remaining paramters take the values $a = -100$, $l_{ss} = 0.33$, $\beta = 0.9925$, $\phi = 0.075$, $\eta = 6$, $(K/Y)_{ss} = 2.5$, $\delta = 0.025$, $\mu_{z,ss} = 1.0055$. We simulate 1,000 samples, each with $T = 250$ observations and estimate the model using the proposed estimation for various values of λ. We study the performance of QML (i.e., $\lambda = 0$) and shrinkage towards the first and second unconditional moments as described in the paper.

Without any misspecification in Panel A in Table 1, we see that QML (i.e., $\lambda = 0$) gives nearly unbiased estimates with low degree of variability as measured by the standard deviation of the sampling distribution. Letting $\lambda = T$ or $\lambda = 10^6$ only worsen the performance of the estimator, as it increases the biases and generates less efficient estimates.

Panel B in Table 1 considers the case, where the data generating model is solved using the true structural parameters (to get the **g**- and the **h**-functions), but when simulating data we use the value $\rho_a = 0.8$ for the persistence in the stationary technology shocks. That is, we introduce a clear misspecification in the model. To ensure that the unconditional variance of the stationary technology shocks are unaffected by this misspecification, we let $\sigma_a = \sqrt{\frac{0.01^2}{1-0.98^2}\left(1-0.80^2\right)}$. Hence, for this case with misspecification, there is no true value of ρ_a and σ_a, and this explains why we do not report their estimates in Panel B. We then estimate the model without accounting for this misspecification, i.e., with the same value of ρ_a being used to solve and simulate the model. Panel B shows that QML with shrinkage ($\lambda = T$) gives parameter estimates that are less biased when compared to QML without shrinkage (i.e., $\lambda = 0$). This is seen clearly from the overall root measure squared biases

$$RMSB^{\theta} = \sqrt{\frac{1}{n_\theta}\sum_{i=1}^{n_\theta}\left(\theta_i - \bar{\theta}_i\right)^2},$$

where θ_i is the true value of the structural parameter and $\bar{\theta}_i$ is the mean estimate of the ith parameter in the Monte Carlo study. Here, n_θ denotes the number of estimated structural parameters. We find $RMSB^{\theta} = 0.116$ with $\lambda = 0$ but only $RMSB^{\theta} = 0.087$ with $\lambda = T$. The cost of robustifying the QML estimator in this way is that we find less efficient estimate with $\lambda = T$ compared to $\lambda = 0$.

We benchmark these results to using an extreme degree of shrinkage with $\lambda = 10^6$, which corresponds to estimating the structural parameters by GMM and obtaining the states afterwards by the CDKF. The Monte Carlo study shows that these GMM estimates display notable biases in finite samples and are clearly less efficient compared to the standard QML estimator ($\lambda = 0$) and the proposed estimator (with $\lambda = T$), both with and without model misspecification. These imprecise GMM estimates of the structural parameters also imply less accurate state estimates when compared to the proposed estimator with $\lambda = T$. This is seen from the following measure related to the estimated states

$$RMSE^{States} = \frac{1}{1,000}\sum_{s=1}^{1,000}\sqrt{\frac{1}{n_x}\sum_{t=1}^{T}\sum_{i=1}^{n_x}\left(x_{i,t}^{(s)} - \hat{x}_{i,t}^{(s)}\right)^2},$$

which is $RMSE^{States} = 0.0223$ with $\lambda = 10^6$ but only $RMSE^{States} = 0.0159$ with $\lambda = T$.

5 Projection Approximation

The considered second order projection approximation reads (unpruned)

$$\mathbf{y}_t = \mathbf{g}_0 + \mathbf{g}_\mathbf{x}\mathbf{x}_t + \mathbf{G}_{\mathbf{xx}}\left(\mathbf{x}_t \otimes \mathbf{x}_t\right)$$

$$\mathbf{x}_{t+1} = \mathbf{h}_0 + \mathbf{h}_\mathbf{x}\mathbf{x}_t + \mathbf{H}_{\mathbf{xx}}\left(\mathbf{x}_t \otimes \mathbf{x}_t\right) + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}$$

where $\mathbf{G}_{\mathbf{xx}} \equiv reshape\left(\mathbf{g}_{\mathbf{xx}}, n_y, n_x^2\right)$ and $\mathbf{H}_{\mathbf{xx}} \equiv reshape\left(\mathbf{h}_{\mathbf{xx}}, n_x, n_x^2\right)$, whereas the pruned version reads

$$\mathbf{y}_t = \mathbf{g}_0 + \mathbf{g}_\mathbf{x}\left(\mathbf{x}_t^f + \mathbf{x}_t^s\right) + \mathbf{G}_{\mathbf{xx}}\left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right)$$

$$\mathbf{x}_{t+1}^f = \mathbf{h}_0 + \mathbf{h}_\mathbf{x}\mathbf{x}_t^f + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}$$

$$\mathbf{x}_{t+1}^s = \mathbf{h}_\mathbf{x}\mathbf{x}_t^s + \mathbf{H}_{\mathbf{xx}}\left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right)$$

Thus, the only differences with respect to the pruned state space system for the standard perturbation approximation are i) the presence of $\mathbf{h}_0$ in the law of motion for $\mathbf{x}_t^f$ and ii) the absence of $\frac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2$ in the law of motion for $\mathbf{x}_t^s$. We include $\mathbf{h}_0$ in the law of motion for $\mathbf{x}_t^f$ because it is presence even in a first-order projection solution as it captures a risk-adjustment. Thus, to get the closed-form moments for the pruned projection solution, all we need to do is to extend the results in Andreasen et al. (2018) to account for $\mathbf{h}_0$. This is done below.

For the pruned version, we can show stability and compute the first and second unconditional moments.

Table 1: A Monte Carlo Study

This table reports the biases and standard deviation for the estimated parameters in a simulation study using 1,000 replications and $T = 250$.

	True value	Parameter bias			Standard deviation		
		$\lambda = 0$	$\lambda = T$	$\lambda = 10^6$	$\lambda = 0$	$\lambda = T$	$\lambda = 10^6$
A: No Misspecification							
χ	2.000	0.005	0.076	-0.009	0.102	0.230	0.226
ξ_{Calvo}	0.750	0.010	0.030	0.013	0.040	0.053	0.057
ϕ_{π}	1.250	0.026	0.092	0.128	0.073	0.147	0.147
$\phi_{\Delta c}$	0.250	-0.001	0.019	0.053	0.045	0.143	0.145
π_{ss}	1.005	0.000	0.001	-0.001	0.002	0.005	0.004
ρ_{μ_z}	0.500	0.051	-0.141	-0.271	0.106	0.191	0.212
σ_{μ_z}	0.001	0.000	0.001	0.000	0.000	0.001	0.001
ρ_a	0.980	-0.001	0.005	0.002	0.004	0.006	0.005
σ_a	0.010	0.000	-0.001	-0.001	0.001	0.001	0.001
RMSB^{θ}	-	0.041	0.212	0.214	-	-	-
RMSE^{States}	-	0.0066	0.0269	0.0327	-	-	-
B: With misspecification							
χ	2.000	-0.339	-0.084	-0.181	0.347	0.192	0.132
ξ_{Calvo}	0.750	-0.010	-0.079	-0.043	0.079	0.193	0.141
ϕ_{π}	1.250	0.004	0.175	0.316	0.100	0.217	0.287
$\phi_{\Delta c}$	0.250	-0.015	0.023	0.148	0.091	0.184	0.279
π_{ss}	1.005	0.000	-0.001	-0.002	0.004	0.002	0.002
ρ_{μ_z}	0.500	0.120	0.035	-0.239	0.151	0.213	0.313
σ_{μ_z}	0.001	0.000	0.000	0.000	0.000	0.001	0.001
RMSB^{θ}	-	0.116	0.087	0.350	-	-	-
RMSE^{States}	-	0.0346	0.0159	0.0223	-	-	-

5.1 Covariance-stationary

Proposition 1:

The pruned second-order approximation for $\mathbf{x}_t^f, \mathbf{x}_t^s$, and $\mathbf{y}_t^s$ is covariance-stationary if

1. If all eigenvalue of $\mathbf{h}_\mathbf{x}$ have modulus less than one
2. $\boldsymbol{\epsilon}_{t+1}$ has finite fourth moment

Proof

We now form the extended state vector

$$\mathbf{z}_t \equiv \begin{bmatrix} \mathbf{x}_t^f \\ \mathbf{x}_t^s \\ \mathbf{x}_t^f \otimes \mathbf{x}_t^f \end{bmatrix}$$

We know the law of motion for $\mathbf{x}_t^f$ and $\mathbf{x}_t^s$, so we only need to find the law of motion for $\mathbf{x}_t^f \otimes \mathbf{x}_t^f$. Hence consider

$$\mathbf{x}_{t+1}^f \otimes \mathbf{x}_{t+1}^f = \left(\mathbf{h}_0 + \mathbf{h}_\mathbf{x}\mathbf{x}_t^f + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right) \otimes \left(\mathbf{h}_0 + \mathbf{h}_\mathbf{x}\mathbf{x}_t^f + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right)$$

$$= \mathbf{h}_0 \otimes \left(\mathbf{h}_0 + \mathbf{h}_\mathbf{x}\mathbf{x}_t^f + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right)$$
$$+\mathbf{h}_\mathbf{x}\mathbf{x}_t^f \otimes \left(\mathbf{h}_0 + \mathbf{h}_\mathbf{x}\mathbf{x}_t^f + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right)$$

$$+\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} \otimes \left(\mathbf{h}_0 + \mathbf{h}_{\mathbf{x}}\mathbf{x}_t^f + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right)$$

using $(\mathbf{A}+\mathbf{B}) \otimes (\mathbf{C}+\mathbf{D}) = \mathbf{A}\otimes\mathbf{C} + \mathbf{A}\otimes\mathbf{D} + \mathbf{B}\otimes\mathbf{C} + \mathbf{B}\otimes\mathbf{D}$

$$\begin{aligned}
&= \mathbf{h}_0 \otimes \mathbf{h}_0 + \mathbf{h}_0 \otimes \mathbf{h}_{\mathbf{x}}\mathbf{x}_t^f + \mathbf{h}_0 \otimes \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} \\
&+\mathbf{h}_{\mathbf{x}}\mathbf{x}_t^f \otimes \mathbf{h}_0 + \mathbf{h}_{\mathbf{x}}\mathbf{x}_t^f \otimes \mathbf{h}_{\mathbf{x}}\mathbf{x}_t^f + \mathbf{h}_{\mathbf{x}}\mathbf{x}_t^f \otimes \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} \\
&+\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{h}_0 + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{h}_{\mathbf{x}}\mathbf{x}_t^f + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}
\end{aligned}$$

$$\begin{aligned}
&= \mathbf{h}_0 \otimes \mathbf{h}_0 + (\mathbf{h}_0 \otimes \mathbf{h}_{\mathbf{x}})\,\mathbf{x}_t^f + (\mathbf{h}_0 \otimes \boldsymbol{\eta})\,\boldsymbol{\epsilon}_{t+1} \\
&+\mathbf{h}_{\mathbf{x}}\mathbf{x}_t^f \otimes \mathbf{h}_0 1 + (\mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_{\mathbf{x}})\left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) + (\mathbf{h}_{\mathbf{x}} \otimes \sigma\boldsymbol{\eta})\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right) \\
&+\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{h}_0 1 + (\boldsymbol{\eta} \otimes \mathbf{h}_{\mathbf{x}})\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right) + (\boldsymbol{\eta} \otimes \boldsymbol{\eta})\,(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1})
\end{aligned}$$

using $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = \mathbf{AC}\otimes\mathbf{BD}$

$$\begin{aligned}
&= \mathbf{h}_0 \otimes \mathbf{h}_0 + (\mathbf{h}_0 \otimes \mathbf{h}_{\mathbf{x}})\,\mathbf{x}_t^f + (\mathbf{h}_0 \otimes \boldsymbol{\eta})\,\boldsymbol{\epsilon}_{t+1} \\
&+ (\mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_0)\left(\mathbf{x}_t^f \otimes 1\right) + (\mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_{\mathbf{x}})\left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) + (\mathbf{h}_{\mathbf{x}} \otimes \boldsymbol{\eta})\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right) \\
&+ (\boldsymbol{\eta} \otimes \mathbf{h}_{\mathbf{0}})\,(\boldsymbol{\epsilon}_{t+1} \otimes 1) + (\boldsymbol{\eta} \otimes \mathbf{h}_{\mathbf{x}})\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right) + (\boldsymbol{\eta} \otimes \boldsymbol{\eta})\,(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1})
\end{aligned}$$

Note that $E\left[(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1})\right] = vec\left(\mathbf{I}_{n_e}\right)$. Thus, we get

$$\begin{aligned}
\mathbf{x}_{t+1}^f \otimes \mathbf{x}_{t+1}^f &= \mathbf{h}_0 \otimes \mathbf{h}_0 + (\mathbf{h}_0 \otimes \mathbf{h}_{\mathbf{x}})\,\mathbf{x}_t^f + (\mathbf{h}_0 \otimes \boldsymbol{\eta})\,\boldsymbol{\epsilon}_{t+1} + (\boldsymbol{\eta} \otimes \boldsymbol{\eta})\,vec\left(\mathbf{I}_{n_e}\right) \\
&+ (\mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_0)\,\mathbf{x}_t^f + (\mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_{\mathbf{x}})\left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) + (\mathbf{h}_{\mathbf{x}} \otimes \boldsymbol{\eta})\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right) \\
&+ (\boldsymbol{\eta} \otimes \mathbf{h}_{\mathbf{0}})\,\boldsymbol{\epsilon}_{t+1} + (\boldsymbol{\eta} \otimes \mathbf{h}_{\mathbf{x}})\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right) + (\boldsymbol{\eta} \otimes \boldsymbol{\eta})\left((\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}) - vec\left(\mathbf{I}_{n_e}\right)\right)
\end{aligned}$$

Accordingly

$$\begin{bmatrix} \mathbf{x}_{t+1}^f \\ \mathbf{x}_{t+1}^s \\ \mathbf{x}_{t+1}^f \otimes \mathbf{x}_{t+1}^f \end{bmatrix} = \begin{bmatrix} \mathbf{h}_{\mathbf{x}} & \mathbf{0}_{n_x \times n_x} & \mathbf{0}_{n_x \times n_x^2} \\ \mathbf{0}_{n_x \times n_x} & \mathbf{h}_{\mathbf{x}} & \mathbf{H}_{\mathbf{xx}} \\ \mathbf{h}_0 \otimes \mathbf{h}_{\mathbf{x}} + \mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_0 & \mathbf{0}_{n_x^2 \times n_x} & \mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_{\mathbf{x}} \end{bmatrix} \begin{bmatrix} \mathbf{x}_t^f \\ \mathbf{x}_t^s \\ \mathbf{x}_t^f \otimes \mathbf{x}_t^f \end{bmatrix} + \begin{bmatrix} \mathbf{h}_0 \\ \mathbf{0} \\ \mathbf{h}_0 \otimes \mathbf{h}_0 + (\boldsymbol{\eta} \otimes \boldsymbol{\eta})\,vec\left(\mathbf{I}_{n_e}\right) \end{bmatrix}$$

$$+ \begin{bmatrix} \boldsymbol{\eta} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{h}_0 \otimes \boldsymbol{\eta} + \boldsymbol{\eta} \otimes \mathbf{h}_{\mathbf{0}} & \boldsymbol{\eta} \otimes \boldsymbol{\eta} & \boldsymbol{\eta} \otimes \mathbf{h}_{\mathbf{x}} & \mathbf{h}_{\mathbf{x}} \otimes \boldsymbol{\eta} \end{bmatrix} \begin{bmatrix} \boldsymbol{\epsilon}_{t+1} \\ \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec\left(\mathbf{I}_{n_e}\right) \\ \boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f \\ \mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1} \end{bmatrix}$$

$\Updownarrow$

$$\mathbf{z}_{t+1} = \mathbf{A}\mathbf{z}_t + \mathbf{c} + \mathbf{B}\boldsymbol{\xi}_{t+1} \tag{18}$$

where $Cov\left(\boldsymbol{\xi}_{t+1}, \boldsymbol{\xi}_{t-s}\right) = \mathbf{0}$ for $s = 1, 2, 3, ...$ because $\boldsymbol{\epsilon}_{t+1}$ is independent across time. This follows from the observation that $\boldsymbol{\xi}_{t+1}$ is identical to the value stated for $\boldsymbol{\xi}_{t+1}$ in Andreasen et al. (2018), which show the claimed result..

The absolute value of the eigenvalues in $\mathbf{h}_{\mathbf{x}}$ are all strictly less then one by assumption. Accordingly, all eigenvalues of $\mathbf{A}$ are also strictly less than one. To see this note first that

$p(\lambda) = \left|\mathbf{A} - \lambda\mathbf{I}_{2n_x + n_x^2}\right|$

$$= \left|\begin{bmatrix} \mathbf{h}_{\mathbf{x}} - \lambda\mathbf{I}_{n_x} & \mathbf{0}_{n_x \times n_x} & \mathbf{0}_{n_x \times n_x^2} \\ \mathbf{0}_{n_x \times n_x} & \mathbf{h}_{\mathbf{x}} - \lambda\mathbf{I}_{n_x} & \mathbf{H}_{\mathbf{xx}} \\ \mathbf{h}_0 \otimes \mathbf{h}_{\mathbf{x}} + \mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_0 & \mathbf{0}_{n_x^2 \times n_x} & \mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_{\mathbf{x}} - \lambda\mathbf{I}_{n_x^2} \end{bmatrix}\right|$$

$$= \begin{vmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{vmatrix}$$

where we let

$\mathbf{B}_{11} \equiv \begin{bmatrix} \mathbf{h}_{\mathbf{x}} - \lambda \mathbf{I}_{n_x} & \mathbf{0}_{n_x \times n_x} \\ \mathbf{0}_{n_x \times n_x} & \mathbf{h}_{\mathbf{x}} - \lambda \mathbf{I}_{n_x} \end{bmatrix}$ which is $2n_x \times 2n_x$

$\mathbf{B}_{12} \equiv \begin{bmatrix} \mathbf{0}_{n_x \times n_x^2} \\ \mathbf{H}_{\mathbf{xx}} \end{bmatrix}$ which is $2n_x \times n_x^2$

$\mathbf{B}_{21} \equiv \begin{bmatrix} \mathbf{h}_0 \otimes \mathbf{h}_{\mathbf{x}} + \mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_0 & \mathbf{0}_{n_x^2 \times n_x} \end{bmatrix}$ which is $n_x^2 \times 2n_x$

$\mathbf{B}_{22} \equiv \mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_{\mathbf{x}} - \lambda \mathbf{I}_{n_x^2}$ which is $n_x^2 \times n_x^2$

$$= |\mathbf{B}_{11}| \, |\mathbf{B}_{22}|$$

using $\begin{vmatrix} \mathbf{U} & \mathbf{C} \\ \mathbf{0} & \mathbf{Y} \end{vmatrix} = |\mathbf{U}| \, |\mathbf{Y}|$ where $\mathbf{U}$ is $m \times m$ and $\mathbf{Y}$ is $n \times n$

$$= \left| \begin{bmatrix} \mathbf{h}_{\mathbf{x}} - \lambda \mathbf{I}_{n_x} & \mathbf{0}_{n_x \times n_x} \\ \mathbf{0}_{n_x \times n_x} & \mathbf{h}_{\mathbf{x}} - \lambda \mathbf{I}_{n_x} \end{bmatrix} \right| \left| \mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_{\mathbf{x}} - \lambda \mathbf{I}_{n_x^2} \right|$$

$$= |\mathbf{h}_{\mathbf{x}} - \lambda \mathbf{I}_{n_x}| \, |\mathbf{h}_{\mathbf{x}} - \lambda \mathbf{I}_{n_x}| \, \left| \mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_{\mathbf{x}} - \lambda \mathbf{I}_{n_x^2} \right|$$

Hence, the eigenvalue λ solves the problem

$p(\lambda) = 0$

$\Updownarrow$

$|\mathbf{h}_{\mathbf{x}} - \lambda \mathbf{I}_{n_x}| \, |\mathbf{h}_{\mathbf{x}} - \lambda \mathbf{I}_{n_x}| \, \left| \mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_{\mathbf{x}} - \lambda \mathbf{I}_{n_x^2} \right| = 0$

$\Updownarrow$

$|\mathbf{h}_{\mathbf{x}} - \lambda \mathbf{I}_{n_x}| = 0$ or $\left| \mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_{\mathbf{x}} - \lambda \mathbf{I}_{n_x^2} \right| = 0$

The absolute value of all eigenvalues to the first problem are strictly less than one. That is $|\lambda_i| < 1$ $i = 1, 2, ..., n_x$. This is also the case for the second problem because the eigenvalues to $\mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_{\mathbf{x}}$ are $\lambda_i \lambda_j$ for $i = 1, 2, ..., n_x$ and $j = 1, 2, ..., n_x$

Thus, the system in (18) is covariance stationary if $\boldsymbol{\xi}_{t+1}$ has finite first and second moment. It follows directly that $E[\boldsymbol{\xi}_{t+1}] = \mathbf{0}$ and $\boldsymbol{\xi}_{t+1}$ has finite second moments if $\boldsymbol{\epsilon}_{t+1}$ has a finite fourth moment. The latter holds by assumption.

For the control variables we have $\mathbf{y}_t = \mathbf{D}\mathbf{z}_t + \mathbf{g}_0$ where $\mathbf{D} \equiv \begin{bmatrix} \mathbf{g}_{\mathbf{x}} & \mathbf{g}_{\mathbf{x}} & \mathbf{G}_{\mathbf{xx}} \end{bmatrix}$. That is $\mathbf{y}_t$ is linear function of $\mathbf{z}_t$ and $\mathbf{y}_t$ is therefore also covariance-stationary.

Q.E.D.

5.2 Formulas for the first and second moments

This section computes first and second unconditional moments using the representation of the second-order system stated above. Note first that

$$E\left[\mathbf{x}_{t+1}^f\right] = \mathbf{h}_0 + \mathbf{h}_{\mathbf{x}} E\left[\mathbf{x}_t^f\right]$$

$\Downarrow$

$$E\left[\mathbf{x}_t^f\right] = (\mathbf{I} - \mathbf{h}_{\mathbf{x}})^{-1} \mathbf{h}_0$$

And

$$Var\left[\mathbf{x}_{t+1}^f\right] = \mathbf{h}_{\mathbf{x}} Var\left[\mathbf{x}_t^f\right] \mathbf{h}_{\mathbf{x}}' + \boldsymbol{\eta}\boldsymbol{\eta}'$$

$\Updownarrow$

$$vec\left(Var\left[\mathbf{x}_{t+1}^f\right]\right) = (\mathbf{h}_{\mathbf{x}} \otimes \mathbf{h}_{\mathbf{x}}) \, vec\left(Var\left[\mathbf{x}_t^f\right]\right) + vec\left(\boldsymbol{\eta}\boldsymbol{\eta}'\right)$$

$\Downarrow$

$$vec\left(Var\left[\mathbf{x}_t^f\right]\right) = \left(\mathbf{I} - \mathbf{h_x} \otimes \mathbf{h_x}\right)^{-1} vec\left(\boldsymbol{\eta}\boldsymbol{\eta}'\right)$$

Hence,

$$\begin{aligned} Var\left[\mathbf{x}_t^f\right] &= E\left[\left(\mathbf{x}_t^f - E\left[\mathbf{x}_t^f\right]\right)\left(\mathbf{x}_t^f - E\left[\mathbf{x}_t^f\right]\right)'\right] \\ &= E\left[\left(\mathbf{x}_t^f - E\left[\mathbf{x}_t^f\right]\right)\left(\mathbf{x}_t^f\right)'\right] \\ &= E\left[\mathbf{x}_t^f\left(\mathbf{x}_t^f\right)'\right] - E\left[\mathbf{x}_t^f\right]E\left[\left(\mathbf{x}_t^f\right)'\right] \end{aligned}$$

$\Updownarrow$

$$E\left[\mathbf{x}_t^f\left(\mathbf{x}_t^f\right)'\right] = Var\left[\mathbf{x}_t^f\right] + E\left[\mathbf{x}_t^f\right]E\left[\left(\mathbf{x}_t^f\right)'\right].$$

The system

$$\mathbf{z}_{t+1} = \mathbf{c} + \mathbf{A}\mathbf{z}_t + \mathbf{B}\boldsymbol{\xi}_{t+1}$$

$$\mathbf{y}_t = \mathbf{D}\mathbf{z}_t + \mathbf{g}_0$$

The mean values are

$$E\left[\mathbf{z}_t\right] = \left(\mathbf{I}_{2n_x+n_x^2} - \mathbf{A}\right)^{-1}\mathbf{c}.$$

$$E\left[\mathbf{y}_t\right] = \mathbf{D}E\left[\mathbf{z}_t\right] + \mathbf{g}_0$$

Then note that $E\left[\mathbf{x}_t^f\right]$ is at the top of $E\left[\mathbf{z}_t\right]$ and $E\left[\mathbf{x}_t^f\left(\mathbf{x}_t^f\right)'\right]$ is at the bottom of $E\left[\mathbf{z}_t\right]$, which is an easy way to get these moments.

For the variances we first have that

$$E\left[\mathbf{z}_{t+1}\mathbf{z}_{t+1}'\right] = E\left[\left(\mathbf{c} + \mathbf{A}\mathbf{z}_t + \mathbf{B}\boldsymbol{\xi}_{t+1}\right)\left(\mathbf{c} + \mathbf{A}\mathbf{z}_t + \mathbf{B}\boldsymbol{\xi}_{t+1}\right)'\right]$$

$$= E\left[\left(\mathbf{c} + \mathbf{A}\mathbf{z}_t + \mathbf{B}\boldsymbol{\xi}_{t+1}\right)\left(\mathbf{c}' + \mathbf{z}_t'\mathbf{A}' + \boldsymbol{\xi}_{t+1}'\mathbf{B}'\right)\right]$$

$$\begin{aligned} &= E\left[\mathbf{c}\left(\mathbf{c}' + \mathbf{z}_t'\mathbf{A}' + \boldsymbol{\xi}_{t+1}'\mathbf{B}'\right)\right] \\ &+ E\left[\mathbf{A}\mathbf{z}_t\left(\mathbf{c}' + \mathbf{z}_t'\mathbf{A}' + \boldsymbol{\xi}_{t+1}'\mathbf{B}'\right)\right] \\ &+ E\left[\mathbf{B}\boldsymbol{\xi}_{t+1}\left(\mathbf{c}' + \mathbf{z}_t'\mathbf{A}' + \boldsymbol{\xi}_{t+1}'\mathbf{B}'\right)\right] \end{aligned}$$

$$\begin{aligned} &= E\left[\mathbf{c}\mathbf{c}' + \mathbf{c}\mathbf{z}_t'\mathbf{A}' + \mathbf{c}\boldsymbol{\xi}_{t+1}'\mathbf{B}'\right] \\ &+ E\left[\mathbf{A}\mathbf{z}_t\mathbf{c}' + \mathbf{A}\mathbf{z}_t\mathbf{z}_t'\mathbf{A}' + \mathbf{A}\mathbf{z}_t\boldsymbol{\xi}_{t+1}'\mathbf{B}'\right] \\ &+ E\left[\mathbf{B}\boldsymbol{\xi}_{t+1}\mathbf{c}' + \mathbf{B}\boldsymbol{\xi}_{t+1}\mathbf{z}_t'\mathbf{A}' + \mathbf{B}\boldsymbol{\xi}_{t+1}\boldsymbol{\xi}_{t+1}'\mathbf{B}'\right] \end{aligned}$$

$$\begin{aligned} &= \mathbf{c}\mathbf{c}' + \mathbf{c}E\left[\mathbf{z}_t'\right]\mathbf{A}' \\ &+ \mathbf{A}E\left[\mathbf{z}_t\right]\mathbf{c}' + \mathbf{A}E\left[\mathbf{z}_t\mathbf{z}_t'\right]\mathbf{A}' + \mathbf{A}E\left[\mathbf{z}_t\boldsymbol{\xi}_{t+1}'\right]\mathbf{B}' \\ &+ \mathbf{B}E\left[\boldsymbol{\xi}_{t+1}\mathbf{z}_t'\right]\mathbf{A}' + \mathbf{B}E\left[\boldsymbol{\xi}_{t+1}\boldsymbol{\xi}_{t+1}'\right]\mathbf{B}' \end{aligned}$$

We then note that

$$E\left[\mathbf{z}_t\boldsymbol{\xi}_{t+1}'\right] = E\left[\begin{bmatrix} \mathbf{x}_t^f \\ \mathbf{x}_t^s \\ \mathbf{x}_t^f \otimes \mathbf{x}_t^f \end{bmatrix}\begin{bmatrix} \boldsymbol{\epsilon}_{t+1}' & \left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec\left(\mathbf{I}_{ne}\right)\right)' & \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' & \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)' \end{bmatrix}\right]$$

$$= E\begin{bmatrix} \mathbf{x}_t^f \boldsymbol{\epsilon}_{t+1}' & \mathbf{x}_t^f \left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec\left(\mathbf{I}_{ne}\right)\right)' & \mathbf{x}_t^f \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' & \mathbf{x}_t^f \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)' \\ \mathbf{x}_t^s \boldsymbol{\epsilon}_{t+1}' & \mathbf{x}_t^s \left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec\left(\mathbf{I}_{ne}\right)\right)' & \mathbf{x}_t^s \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' & \mathbf{x}_t^s \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)' \\ \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) \boldsymbol{\epsilon}_{t+1}' & \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) \left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec\left(\mathbf{I}_{ne}\right)\right)' & \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' & \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)' \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$$

Thus

$$E\left[\mathbf{z}_{t+1}\mathbf{z}_{t+1}'\right] = \mathbf{c}\mathbf{c}' + \mathbf{c}E\left[\mathbf{z}_t'\right]\mathbf{A}' + \mathbf{A}E\left[\mathbf{z}_t\right]\mathbf{c}' + \mathbf{A}E\left[\mathbf{z}_t\mathbf{z}_t'\right]\mathbf{A}' + \mathbf{B}E\left[\boldsymbol{\xi}_{t+1}\boldsymbol{\xi}_{t+1}'\right]\mathbf{B}'$$

$$= \mathbf{c}E\left[\mathbf{z}_t'\right]\mathbf{A}' + \left(\mathbf{c} + \mathbf{A}E\left[\mathbf{z}_t\right]\right)\mathbf{c}' + \mathbf{A}E\left[\mathbf{z}_t\mathbf{z}_t'\right]\mathbf{A}' + \mathbf{B}E\left[\boldsymbol{\xi}_{t+1}\boldsymbol{\xi}_{t+1}'\right]\mathbf{B}'$$

Note also that

$$E\left[\mathbf{z}_t\right]E\left[\mathbf{z}_t\right]' = \left(\mathbf{c} + \mathbf{A}E\left[\mathbf{z}_t\right]\right)\left(\mathbf{c} + \mathbf{A}E\left[\mathbf{z}_t\right]\right)'$$

$$= \left(\mathbf{c} + \mathbf{A}E\left[\mathbf{z}_t\right]\right)\mathbf{c}' + \left(\mathbf{c} + \mathbf{A}E\left[\mathbf{z}_t\right]\right)E\left[\mathbf{z}_t'\right]\mathbf{A}'$$

$$= \left(\mathbf{c} + \mathbf{A}E\left[\mathbf{z}_t\right]\right)\mathbf{c}' + \mathbf{c}E\left[\mathbf{z}_t'\right]\mathbf{A}' + \mathbf{A}E\left[\mathbf{z}_t\right]E\left[\mathbf{z}_t'\right]\mathbf{A}'$$

So

$$E\left[\mathbf{z}_{t+1}\mathbf{z}_{t+1}'\right] - E\left[\mathbf{z}_t\right]E\left[\mathbf{z}_t\right]' = \mathbf{c}E\left[\mathbf{z}_t'\right]\mathbf{A}' + \left(\mathbf{c} + \mathbf{A}E\left[\mathbf{z}_t\right]\right)\mathbf{c}' + \mathbf{A}E\left[\mathbf{z}_t\mathbf{z}_t'\right]\mathbf{A}' + \mathbf{B}E\left[\boldsymbol{\xi}_{t+1}\boldsymbol{\xi}_{t+1}'\right]\mathbf{B}' - \left(\mathbf{c} + \mathbf{A}E\left[\mathbf{z}_t\right]\right)\mathbf{c}' - \mathbf{c}E\left[\mathbf{z}_t'\right]\mathbf{A}' - \mathbf{A}E\left[\mathbf{z}_t\right]E\left[\mathbf{z}_t'\right]\mathbf{A}'$$

$$= \mathbf{A}E\left[\mathbf{z}_t\mathbf{z}_t'\right]\mathbf{A}' + \mathbf{B}E\left[\boldsymbol{\xi}_{t+1}\boldsymbol{\xi}_{t+1}'\right]\mathbf{B}' - \mathbf{A}E\left[\mathbf{z}_t\right]E\left[\mathbf{z}_t'\right]\mathbf{A}'$$

$$= \mathbf{A}\left(E\left[\mathbf{z}_t\mathbf{z}_t'\right] - E\left[\mathbf{z}_t\right]E\left[\mathbf{z}_t'\right]\right)\mathbf{A}' + \mathbf{B}E\left[\boldsymbol{\xi}_{t+1}\boldsymbol{\xi}_{t+1}'\right]\mathbf{B}'$$

$\Updownarrow$

$$Var\left(\mathbf{z}_{t+1}\right) = \mathbf{A}Var\left(\mathbf{z}_t\right)\mathbf{A}' + \mathbf{B}Var\left(\boldsymbol{\xi}_{t+1}\right)\mathbf{B}'$$

$\Updownarrow$

$$vec\left(Var\left(\mathbf{z}_{t+1}\right)\right) = vec\left(\mathbf{A}Var\left(\mathbf{z}_t\right)\mathbf{A}'\right) + vec\left(\mathbf{B}Var\left(\boldsymbol{\xi}_{t+1}\right)\mathbf{B}'\right)$$

$\Updownarrow$

$$vec\left(Var\left(\mathbf{z}_{t+1}\right)\right) = \left(\mathbf{A} \otimes \mathbf{A}\right)vec\left(Var\left(\mathbf{z}_t\right)\right) + vec\left(\mathbf{B}Var\left(\boldsymbol{\xi}_{t+1}\right)\mathbf{B}'\right)$$

$\Updownarrow$

$$vec\left(Var\left(\mathbf{z}_{t+1}\right)\right)\left(\mathbf{I}_{(2n_x+n_x^2)^2} - \left(\mathbf{A} \otimes \mathbf{A}\right)\right) = vec\left(\mathbf{B}Var\left(\boldsymbol{\xi}_{t+1}\right)\mathbf{B}'\right)$$

$\Updownarrow$

$$vec\left(Var\left(\mathbf{z}_{t+1}\right)\right) = \left(\mathbf{I}_{(2n_x+n_x^2)^2} - \left(\mathbf{A} \otimes \mathbf{A}\right)\right)^{-1} vec\left(\mathbf{B}Var\left(\boldsymbol{\xi}_{t+1}\right)\mathbf{B}'\right)$$

Hence we only need to compute $Var\left(\boldsymbol{\xi}_{t+1}\right)$. Given that the expression for $\boldsymbol{\xi}_{t+1}$ is identifical to the one for the pruned standard perturbation approximation, the matrix $Var\left(\boldsymbol{\xi}_{t+1}\right)$ can be computed as outlined in Andreasen et al. (2018). For completeness, we reproduce how this is done below.

$$Var\left(\boldsymbol{\xi}_{t+1}\right) = E\left[\left[\begin{array}{c}\boldsymbol{\epsilon}_{t+1}\\ \boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\\ \boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\\ \mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\end{array}\right]\left[\begin{array}{c}\boldsymbol{\epsilon}_{t+1}\\ \boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\\ \boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\\ \mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\end{array}\right]'\right]$$

$$= E\left[\left[\begin{array}{c}\boldsymbol{\epsilon}_{t+1}\\ \boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\\ \boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\\ \mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\end{array}\right]\begin{array}{cccc}\boldsymbol{\epsilon}'_{t+1} & \left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)' & \left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)' & \left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)'\end{array}\right]$$

$$= E\left[\begin{array}{cc}\boldsymbol{\epsilon}_{t+1}\boldsymbol{\epsilon}'_{t+1} & \boldsymbol{\epsilon}_{t+1}\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)'\\ \left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)\boldsymbol{\epsilon}'_{t+1} & \left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)'\\ \left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)\boldsymbol{\epsilon}'_{t+1} & \left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)'\\ \left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)\boldsymbol{\epsilon}'_{t+1} & \left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)'\end{array}\right.$$

$$\left.\begin{array}{cc}\boldsymbol{\epsilon}_{t+1}\left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)' & \boldsymbol{\epsilon}_{t+1}\left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)'\\ \left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)\left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)' & \left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)\left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)'\\ \left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)\left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)' & \left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)\left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)'\\ \left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)\left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)' & \left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)\left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)'\end{array}\right]$$

We next evaluate each of these terms.
The first row:

$E\left[\boldsymbol{\epsilon}_{t+1}\boldsymbol{\epsilon}'_{t+1}\right] = \mathbf{I}_{n_\varepsilon}$

$E\left[\boldsymbol{\epsilon}_{t+1}\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)'\right] = E\left[\boldsymbol{\epsilon}_{t+1}\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)'-\boldsymbol{\epsilon}_{t+1}vec\left(\mathbf{I}_{n_e}\right)'\right]$

$\quad = E\left[\boldsymbol{\epsilon}_{t+1}\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)'\right]$

$E\left[\boldsymbol{\epsilon}_{t+1}\left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)'\right] = E\left[\left(\boldsymbol{\epsilon}_{t+1}\otimes 1\right)\left(\boldsymbol{\epsilon}'_{t+1}\otimes\left(\mathbf{x}_t^f\right)'\right)\right]$

$\quad = E\left[\boldsymbol{\epsilon}_{t+1}\boldsymbol{\epsilon}'_{t+1}\otimes\left(\mathbf{x}_t^f\right)'\right]$

$\quad = E\left[E_t\left[\boldsymbol{\epsilon}_{t+1}\boldsymbol{\epsilon}'_{t+1}\right]\otimes\left(\mathbf{x}_t^f\right)'\right]$

$\quad = E\left[\mathbf{I}_{n_\varepsilon}\otimes\left(\mathbf{x}_t^f\right)'\right]$

$\quad = \mathbf{I}_{n_\varepsilon}\otimes E\left[\left(\mathbf{x}_t^f\right)'\right]$

$E\left[\boldsymbol{\epsilon}_{t+1}\left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)'\right] = E\left[\left(1\otimes\boldsymbol{\epsilon}_{t+1}\right)\left(\left(\mathbf{x}_t^f\right)'\otimes\boldsymbol{\epsilon}'_{t+1}\right)\right]$

$\quad = E\left[\left(\mathbf{x}_t^f\right)'\otimes\boldsymbol{\epsilon}_{t+1}\boldsymbol{\epsilon}'_{t+1}\right]$

$\quad = E\left[\left(\mathbf{x}_t^f\right)'\right]\otimes\mathbf{I}_{n_\varepsilon}$

The second row

$E\left[\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)\boldsymbol{\epsilon}'_{t+1}\right] = E\left[\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)\boldsymbol{\epsilon}'_{t+1}\right]$

$E\left[\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)'\right] = E\left[\left(\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)-vec\left(\mathbf{I}_{n_e}\right)\right)\left(\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)'-vec\left(\mathbf{I}_{n_e}\right)'\right)\right]$

$\quad = E\left[\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)'-\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)vec\left(\mathbf{I}_{n_e}\right)'-vec\left(\mathbf{I}_{n_e}\right)\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)'+vec\left(\mathbf{I}_{n_e}\right)vec\left(\mathbf{I}_{n_e}\right)'\right]$

$$= E\left[(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}) (\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1})' - vec(\mathbf{I}_{n_e}) vec(\mathbf{I}_{n_e})' - vec(\mathbf{I}_{n_e}) vec(\mathbf{I}_{n_e})' + vec(\mathbf{I}_{n_e}) vec(\mathbf{I}_{n_e})' \right]$$
$$= E\left[(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}) (\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1})' - vec(\mathbf{I}_{n_e}) vec(\mathbf{I}_{n_e})' \right]$$
$$E\left[(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec(\mathbf{I}_{n_e})) \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' \right] = E\left[(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}) \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' - vec(\mathbf{I}_{n_e}) \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' \right]$$
$$= E\left[(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}) \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' \right]$$
$$E\left[(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec(\mathbf{I}_{n_e})) \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)' \right] = E\left[(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}) \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)' \right]$$

Third row

$$E\left[\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right) \boldsymbol{\epsilon}'_{t+1} \right] = E\left[\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right) \left(\boldsymbol{\epsilon}'_{t+1} \otimes 1\right) \right]$$
$$= E\left[\boldsymbol{\epsilon}_{t+1}\boldsymbol{\epsilon}'_{t+1} \otimes \mathbf{x}_t^f \right]$$
$$= \mathbf{I}_{n_\varepsilon} \otimes E\left[\mathbf{x}_t^f \right]$$
$$E\left[\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right) (\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec(\mathbf{I}_{n_e}))' \right] = E\left[\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right) (\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1})' \right]$$
$$E\left[\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right) \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' \right] = E\left[\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right) \left(\boldsymbol{\epsilon}'_{t+1} \otimes \left(\mathbf{x}_t^f\right)'\right) \right]$$
$$= E\left[(\boldsymbol{\epsilon}_{t+1}\boldsymbol{\epsilon}'_{t+1}) \otimes \left(\mathbf{x}_t^f \left(\mathbf{x}_t^f\right)'\right) \right]$$
$$= \mathbf{I}_{n_e} \otimes E\left[\mathbf{x}_t^f \left(\mathbf{x}_t^f\right)' \right]$$

using $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$

$$E\left[\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right) \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)' \right] = E\left[\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right) \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)' \right]$$

Fourth row

$$E\left[\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right) \boldsymbol{\epsilon}'_{t+1} \right] = E\left[\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right) \left(1 \otimes \boldsymbol{\epsilon}'_{t+1}\right) \right] = E\left[\mathbf{x}_t^f \otimes (\boldsymbol{\epsilon}_{t+1}\boldsymbol{\epsilon}'_{t+1}) \right] = E\left[\mathbf{x}_t^f \right] \otimes \mathbf{I}_{n_e}$$
$$E\left[\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right) (\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec(\mathbf{I}_{n_e}))' \right] = E\left[\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right) (\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1})' \right]$$
$$E\left[\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right) \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' \right] = E\left[\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right) \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' \right]$$
$$E\left[\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right) \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)' \right] = E\left[\left(\mathbf{x}_t^f \left(\mathbf{x}_t^f\right)'\right) \otimes (\boldsymbol{\epsilon}_{t+1}\boldsymbol{\epsilon}'_{t+1}) \right] = E\left[\mathbf{x}_t^f \left(\mathbf{x}_t^f\right)' \right] \otimes \mathbf{I}_{n_e}$$

Exploiting that all third moments of $\boldsymbol{\epsilon}_{t+1}$ are zero for the normal distribution, we get

$$Var(\boldsymbol{\xi}_{t+1}) = \begin{bmatrix} \mathbf{I}_{n_e} & \mathbf{0}_{n_\varepsilon \times n_\varepsilon^2} & \mathbf{I}_{n_\varepsilon} \otimes E\left[\left(\mathbf{x}_t^f\right)'\right] & E\left[\left(\mathbf{x}_t^f\right)'\right] \otimes \mathbf{I}_{n_\varepsilon} \\ \mathbf{0}_{n_\varepsilon^2 \times n_\varepsilon} & E\left[(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}) (\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1})' - vec(\mathbf{I}_{n_e}) vec(\mathbf{I}_{n_e})' \right] & \mathbf{0}_{n_\varepsilon^2 \times n_\varepsilon n_x} & \mathbf{0}_{n_\varepsilon^2 \times n_\varepsilon n_x} \\ \mathbf{I}_{n_\varepsilon} \otimes E\left[\mathbf{x}_t^f\right] & \mathbf{0}_{n_\varepsilon n_x \times n_\varepsilon^2} & \mathbf{I}_{n_e} \otimes E\left[\mathbf{x}_t^f \left(\mathbf{x}_t^f\right)'\right] & E\left[\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right) \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)' \right] \\ E\left[\mathbf{x}_t^f\right] \otimes \mathbf{I}_{n_e} & \mathbf{0}_{n_\varepsilon n_x \times n_\varepsilon^2} & E\left[\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right) \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' \right] & E\left[\mathbf{x}_t^f \left(\mathbf{x}_t^f\right)'\right] \otimes \mathbf{I}_{n_e} \end{bmatrix}$$

All elements in this matrix can be computed (and coded) directly as shown below. The variance of the control variables is then given by

$$Var\left[\mathbf{y}_t^s\right] = \mathbf{D} Var\left[\mathbf{z}_t\right] \mathbf{D}'$$

5.2.1 Computing the variance of the innovations

1) for $E\left[\boldsymbol{\epsilon}_{t+1}\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)'\right]$

$$E\left[\boldsymbol{\epsilon}_{t+1}\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)'\right]=E\left[\left\{\epsilon_{t+1}\left(\phi_1,1\right)\right\}_{\phi_1=1}^{n_e}\left(\left\{\epsilon_{t+1}\left(\phi_2,1\right)\left\{\epsilon_{t+1}\left(\phi_3,1\right)\right\}_{\phi_3=1}^{n_e}\right\}_{\phi_2=1}^{n_e}\right)'\right]$$

Hence the quasi MATLAB codes are :

```
E_eps_eps2 = zeros(ne, (ne)^2)
for phi1 = 1 : ne
    index2 = 0
    for phi2 = 1 : ne
        for phi3 = 1 : ne
            index2 = index2 + 1
            if (phi1 = phi2 = phi3)
                E_eps_eps2(phi1, index2) = m^3 (ϵ_{t+1} (phi1))
            end
        end
    end
end
```

Note also that $E\left[\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)\boldsymbol{\epsilon}_{t+1}'\right]=\left(E\left[\boldsymbol{\epsilon}_{t+1}\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)'\right]\right)'$

2) $E\left[\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)'\right]-vec\left(\mathbf{I}_{n_e}\right)vec\left(\mathbf{I}_{n_e}\right)'$

Here

$E\left[\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)'\right]$

$$=E\left[\left\{\epsilon_{t+1}\left(\phi_1,1\right)\left\{\epsilon_{t+1}\left(\phi_2,1\right)\right\}_{\phi_2=1}^{n_e}\right\}_{\phi_1=1}^{n_e}\left(\left\{\epsilon_{t+1}\left(\phi_3,1\right)\left\{\epsilon_{t+1}\left(\phi_4,1\right)\right\}_{\phi_4=1}^{n_e}\right\}_{\phi_3=1}^{n_e}\right)'\right]$$

Hence the quasi MATLAB codes are

```
E_eps2_eps2 = zeros(n_e^2, n_e^2)
index1 = 0
for phi1 = 1 : n_e
    for phi2 = 1 : n_e
        index1 = index1 + 1
        index2 = 0
        for phi3 = 1 : n_e
            for phi4 = 1 : n_e
                index2 = index2 + 1
                % second moments
                if (phi1 == phi2 && phi3 == phi4 && phi1~ = phi4)
                    E_eps2_eps2(index1, index2) = 1
                elseif (phi1 == phi3 && phi2 == phi4 && phi1~ = phi2)
                    E_eps2_eps2(index1, index2) = 1
                elseif (phi1 == phi4 && phi2 == phi3 && phi1~ = phi2)
                    E_eps2_eps2(index1, index2) = 1
                % fourth moments
                elseif(phi1 == phi2 && phi1 == phi3 && phi1 == phi4)
                    E_eps2_eps2(index1, index2) = m^4 (ϵ_{t+1} (phi1))
                end
            end
        end
    end
```

end

3) $E\left[\left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)\left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)'\right]$

Here

$E\left[\left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)\left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)'\right]$

$$= E\left[\left\{\epsilon_{t+1}(\phi_1,1)\left\{x_t^f(\gamma_1,1)\right\}_{\gamma_1=1}^{n_x}\right\}_{\phi_1=1}^{n_e}\left\{\epsilon_{t+1}(\phi_2,1)\left\{x_t^f(\gamma_2,1)\right\}_{\gamma_2=1}^{n_x}\right\}_{\phi_2=1}^{n_e}\right]$$

Hence the quasi MATLAB codes are

```
E_epsxf_epsxf = zeros(n_e n_x, n_x n_e)
index1 = 0
for phi1 = 1 : ne
    for gama1 = 1 : nx
        index1 = index1 + 1
        index2 = 0
        for phi2 = 1 : ne
            for gama2 = 1 : nx
                index2 = index2 + 1
                if phi1 = phi2
                    E_epsxf_epsxf (index1, index2) = E_xf_xf(gama1, gama2)
                end
            end
        end
    end
end
```

where $E\_xf\_xf = reshape(E\left[\mathbf{x}_t^f\otimes\mathbf{x}_t^f\right], nx, nx)$

4) $E\left[\left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)\left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)'\right]$

Here

$E\left[\left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)\left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)'\right]$

$$= E\left[\left\{\epsilon_{t+1}(\phi_1,1)\left\{x_t^f(\gamma_1,1)\right\}_{\gamma_1=1}^{n_x}\right\}_{\phi_1=1}^{n_e}\left\{x_t^f(\gamma_2,1)\left\{\epsilon_{t+1}(\phi_2,1)\right\}_{\phi_2=1}^{n_e}\right\}_{\gamma_2=1}^{n_x}\right]$$

Hence the quasi MATLAB codes are

```
E_epsxf_xfeps = zeros(n_e n_x, n_e n_x)
index1 = 0
for phi1 = 1 : ne
    for gama1 = 1 : nx
        index1 = index1 + 1
        index2 = 0
        for gama2 = 1 : nx
            for phi2 = 1 : ne
                index2 = index2 + 1
                if phi1 = phi2
                    E_epsxf_xfeps (index1, index2) = E_xf_xf(gama1, gama2)
                end
            end
```

end
end
end

5) $E\left[\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)'\right]$
Here
$E\left[\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)'\right] = \left[E\left[\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)'\right]\right]'$

so $E\_xfeps\_epsxf = E\_epsxf\_xfeps'$

6) $E\left[\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)'\right]$
Here
$E\left[\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)\left(\boldsymbol{\epsilon}_{t+1}' \otimes \left(\mathbf{x}_t^f\right)'\right)\right]$

$$= E\left[\left\{x_t^f(\gamma_1,1)\left\{\epsilon_{t+1}(\phi_1,1)\right\}_{\phi_1=1}^{n_e}\right\}_{\gamma_1=1}^{n_x}\left(\left\{\epsilon_{t+1}(\phi_2,1)\left\{x_t^f(\gamma_2,1)\right\}_{\gamma_2=1}^{n_x}\right\}_{\phi_2=1}^{n_e}\right)'\right]$$

Thus the quasi Matlab codes are

```
E_xfeps_epsxf = zeros(n_x n_e, n_x n_e)
index1 = 0
for gama1 = 1 : nx
    for phi1 = 1 : ne
        index1 = index1 + 1
        index2 = 0
        for phi2 = 1 : ne
            for gama2 = 1 : nx
                index2 = index2 + 1
                if phi1 = phi2
                    E_xfeps_epsxf (index1, index2) = E_xf_xf(gama1, gama2)
                end
            end
        end
    end
end
```

where $E\_xf\_xf = reshape(E\left[\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right], nx, nx)$

5.3 The auto-correlations

This section derives the auto-correlations for the states and the control variables.

5.3.1 The innovations

We start by showing that $\boldsymbol{\xi}_{t+1}$ and $\boldsymbol{\xi}_{t+1+s}$ are uncorrelated for $s = 1, 2, ...$ To see this note that
$E\left[\boldsymbol{\xi}_{t+1}\boldsymbol{\xi}_{t+1+s}'\right] =$

$$E\left[\left[\begin{array}{c}\boldsymbol{\epsilon}_{t+1}\\ \boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\\ \boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\\ \mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\end{array}\right]\begin{array}{cccc}\boldsymbol{\epsilon}'_{t+1+s} & \left(\boldsymbol{\epsilon}_{t+1+s}\otimes\boldsymbol{\epsilon}_{t+1+s}-vec\left(\mathbf{I}_{n_e}\right)\right)' & \left(\boldsymbol{\epsilon}_{t+1+s}\otimes\mathbf{x}_{t+s}^f\right)' & \left(\mathbf{x}_{t+s}^f\otimes\boldsymbol{\epsilon}_{t+1+s}\right)'\end{array}\right]$$

$$=E\left[\begin{array}{cc}\boldsymbol{\epsilon}_{t+1}\boldsymbol{\epsilon}'_{t+1+s} & \boldsymbol{\epsilon}_{t+1}\left(\boldsymbol{\epsilon}_{t+1+s}\otimes\boldsymbol{\epsilon}_{t+1+s}-vec\left(\mathbf{I}_{n_e}\right)\right)'\\ \left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)\boldsymbol{\epsilon}'_{t+1+s} & \left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)\left(\boldsymbol{\epsilon}_{t+1+s}\otimes\boldsymbol{\epsilon}_{t+1+s}-vec\left(\mathbf{I}_{n_e}\right)\right)'\\ \left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)\boldsymbol{\epsilon}'_{t+1+s} & \left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)\left(\boldsymbol{\epsilon}_{t+1+s}\otimes\boldsymbol{\epsilon}_{t+1+s}-vec\left(\mathbf{I}_{n_e}\right)\right)'\\ \left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)\boldsymbol{\epsilon}'_{t+1+s} & \left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)\left(\boldsymbol{\epsilon}_{t+1+s}\otimes\boldsymbol{\epsilon}_{t+1+s}-vec\left(\mathbf{I}_{n_e}\right)\right)'\end{array}\right.$$

$$\left.\begin{array}{cc}\boldsymbol{\epsilon}_{t+1}\left(\boldsymbol{\epsilon}_{t+1+s}\otimes\mathbf{x}_{t+s}^f\right)' & \boldsymbol{\epsilon}_{t+1}\left(\mathbf{x}_{t+s}^f\otimes\boldsymbol{\epsilon}_{t+1+s}\right)'\\ \left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)\left(\boldsymbol{\epsilon}_{t+1+s}\otimes\mathbf{x}_{t+s}^f\right)' & \left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)\left(\mathbf{x}_{t+s}^f\otimes\boldsymbol{\epsilon}_{t+1+s}\right)'\\ \left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)\left(\boldsymbol{\epsilon}_{t+1+s}\otimes\mathbf{x}_{t+s}^f\right)' & \left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)\left(\mathbf{x}_{t+s}^f\otimes\boldsymbol{\epsilon}_{t+1+s}\right)'\\ \left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)\left(\boldsymbol{\epsilon}_{t+1+s}\otimes\mathbf{x}_{t+s}^f\right)' & \left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)\left(\mathbf{x}_{t+s}^f\otimes\boldsymbol{\epsilon}_{t+1+s}\right)'\end{array}\right]$$

$$=\left[\begin{array}{cccc}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{array}\right]$$

5.3.2 The auto-covariances

Recall that we have

$$\mathbf{z}_t=\left[\begin{array}{c}\mathbf{x}_t^f\\ \mathbf{x}_t^s\\ \mathbf{x}_t^f\otimes\mathbf{x}_t^f\end{array}\right]$$

$$\mathbf{z}_{t+1}=\mathbf{c}+\mathbf{A}\mathbf{z}_t+\mathbf{B}\boldsymbol{\xi}_{t+1}$$

$$\mathbf{y}_t^s=\mathbf{D}\mathbf{z}_t+\mathbf{g}_0$$

To find the one period auto-correlation, i.e. $Cov\left(\mathbf{z}_{t+1},\mathbf{z}_t\right)$, we have

$$Cov\left(\mathbf{z}_{t+1},\mathbf{z}_t\right)=Cov\left(\mathbf{c}+\mathbf{A}\mathbf{z}_t+\mathbf{B}\boldsymbol{\xi}_{t+1},\mathbf{z}_t\right)=\mathbf{A}Cov\left(\mathbf{z}_t,\mathbf{z}_t\right)=\mathbf{A}Var\left(\mathbf{z}_t\right)$$

because $Cov\left(\mathbf{z}_t,\boldsymbol{\xi}_{t+1}\right)=0$ as shown above. And for two periods

$Cov\left(\mathbf{z}_{t+2},\mathbf{z}_t\right)=Cov\left(\mathbf{c}+\mathbf{A}\mathbf{z}_{t+1}+\mathbf{B}\boldsymbol{\xi}_{t+2},\mathbf{z}_t\right)$
$=Cov\left(\mathbf{A}\left(\mathbf{c}+\mathbf{A}\mathbf{z}_t+\mathbf{B}\boldsymbol{\xi}_{t+1}\right)+\mathbf{B}\boldsymbol{\xi}_{t+2},\mathbf{z}_t\right)$
$=Cov\left(\mathbf{A}^2\mathbf{z}_t+\mathbf{B}\boldsymbol{\xi}_{t+2},\mathbf{z}_t\right)$
$=Cov\left(\mathbf{A}^2\mathbf{z}_t,\mathbf{z}_t\right)$
$=\mathbf{A}^2Cov\left(\mathbf{z}_t,\mathbf{z}_t\right)$
$=\mathbf{A}^2Var\left(\mathbf{z}_t\right)$

Here, we use the fact that $Cov\left(\boldsymbol{\xi}_{t+2},\mathbf{z}_t\right)=0$. This follows from the same arguements as above, that is consider

$$E\left[\mathbf{z}_t\boldsymbol{\xi}'_{t+2}\right]=E\left[\left[\begin{array}{c}\mathbf{x}_t^f\\ \mathbf{x}_t^s\\ \mathbf{x}_t^f\otimes\mathbf{x}_t^f\end{array}\right]\left[\begin{array}{cccc}\boldsymbol{\epsilon}'_{t+2} & \left(\boldsymbol{\epsilon}_{t+2}\otimes\boldsymbol{\epsilon}_{t+2}-vec\left(\mathbf{I}_{ne}\right)\right)' & \left(\boldsymbol{\epsilon}_{t+2}\otimes\mathbf{x}_{t+1}^f\right)' & \left(\mathbf{x}_{t+1}^f\otimes\boldsymbol{\epsilon}_{t+2}\right)'\end{array}\right]\right]$$

$$= E\begin{bmatrix} \mathbf{x}_t^f \boldsymbol{\epsilon}_{t+2}' & \mathbf{x}_t^f \left(\boldsymbol{\epsilon}_{t+2} \otimes \boldsymbol{\epsilon}_{t+2} - vec\left(\mathbf{I}_{ne}\right)\right)' & \mathbf{x}_t^f \left(\boldsymbol{\epsilon}_{t+2} \otimes \mathbf{x}_{t+1}^f\right)' & \mathbf{x}_t^f \left(\mathbf{x}_{t+1}^f \otimes \boldsymbol{\epsilon}_{t+2}\right)' \\ \mathbf{x}_t^s \boldsymbol{\epsilon}_{t+2}' & \mathbf{x}_t^s \left(\boldsymbol{\epsilon}_{t+2} \otimes \boldsymbol{\epsilon}_{t+2} - vec\left(\mathbf{I}_{ne}\right)\right)' & \mathbf{x}_t^s \left(\boldsymbol{\epsilon}_{t+2} \otimes \mathbf{x}_{t+1}^f\right)' & \mathbf{x}_t^s \left(\mathbf{x}_{t+1}^f \otimes \boldsymbol{\epsilon}_{t+2}\right)' \\ \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) \boldsymbol{\epsilon}_{t+1}' & \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) \left(\boldsymbol{\epsilon}_{t+2} \otimes \boldsymbol{\epsilon}_{t+2} - vec\left(\mathbf{I}_{ne}\right)\right)' & \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) \left(\boldsymbol{\epsilon}_{t+2} \otimes \mathbf{x}_{t+1}^f\right)' & \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) \left(\mathbf{x}_{t+1}^f \otimes \boldsymbol{\epsilon}_{t+2}\right)' \end{bmatrix}$$

$$= 0$$

Hence, in the general case
$Cov\left(\mathbf{z}_{t+l}, \mathbf{z}_t\right) = \mathbf{A}^l Var\left(\mathbf{z}_t\right)$

For the control variables:
$Cov\left(\mathbf{y}_{t+l}^s, \mathbf{y}_t^s\right) = Cov\left(\mathbf{D}\mathbf{z}_{t+l} + \frac{1}{2}\mathbf{g}_{\sigma\sigma}\sigma^2, \mathbf{D}\mathbf{z}_t + \mathbf{g}_0\right)$

$$= Cov\left(\mathbf{D}\mathbf{z}_{t+l}, \mathbf{D}\mathbf{z}_t\right)$$

$$= \mathbf{D}Cov\left(\mathbf{z}_{t+l}, \mathbf{z}_t\right)\mathbf{D}'$$

$$= \mathbf{D}\mathbf{A}^l Var\left(\mathbf{z}_t\right)\mathbf{D}'$$

5.4 Impulse Response Functions

This section derives closed-form expressions for the generalized impulse response functions in a non-linear DSGE model approximated by the pruned second-order projection solution. The generalized impulse response functions are defined as

$$GIRF_{\mathbf{var}}\left(l, \nu_i, \mathbf{w}_t\right) = E_t\left[\mathbf{var}_{t+l} | \nu_i\right] - E_t\left[\mathbf{var}_{t+l}\right]$$

for a disturbance to innovation i. To reduce the notational burden in the derivations below, we adopt the parsimonious notation

$$IRF_{\mathbf{var}}\left(l, \nu_i, \mathbf{w}_t\right) = E_t\left[\widetilde{\mathbf{var}}_{t+l}\right] - E_t\left[\mathbf{var}_{t+l}\right]$$

in relation to the conditional expectation operators. Note that the formulas we derive below also apply even if we want to explore the joint effects of more than one shock - for instance when simultaneous shocking disturbances i and j, i.e. $GIRF_{\mathbf{var}}\left(l, \nu_i, \nu_j, \mathbf{w}_t\right) = E_t\left[\mathbf{var}_{t+l} | \nu_i, \nu_j\right] - E_t\left[\mathbf{var}_{t+l}\right]$.

5.4.1 The specification for the conditional information

This subsection explains how we will compute conditional expectations by conditioning on ν_i - and possible more disturbances. Let $\mathbf{S}$ be $n_\epsilon \times n_\epsilon$ diagonal selection matrix with either 1 or zeros on the diagonal, and let the shock sizes appear in the vector $\boldsymbol{\nu}$ of dimension $n_\epsilon \times 1$. For shocks which are not hit by a disturbance, we simply put them to zero.

As an example, consider an economy with three shocks and we want to condition our expectations on the first shock. Hence, we need the vector

$$\begin{bmatrix} \nu_1 \\ \epsilon_{2,t+1} \\ \epsilon_{3,t+1} \end{bmatrix}$$

We can form this vector by letting

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and

$$\boldsymbol{\nu} = \begin{bmatrix} \nu_1 \\ 0 \\ 0 \end{bmatrix}.$$

Then we have

$$\begin{aligned} \mathbf{S}\boldsymbol{\nu} + (\mathbf{I} - \mathbf{S})\,\boldsymbol{\epsilon}_{t+1} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \nu_1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \epsilon_{1,t+1} \\ \epsilon_{2,t+1} \\ \epsilon_{3,t+1} \end{bmatrix} \\ &= \begin{bmatrix} \nu_1 \\ \epsilon_{2,t+1} \\ \epsilon_{3,t+1} \end{bmatrix} \end{aligned}$$

Similarly, if we want to condition on the first two shocks, then we let

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and

$$\boldsymbol{\nu} = \begin{bmatrix} \nu_1 \\ \nu_2 \\ 0 \end{bmatrix}.$$

meaning that

$$\mathbf{S}\boldsymbol{\nu} + (\mathbf{I} - \mathbf{S})\,\boldsymbol{\epsilon}_{t+1} = \begin{bmatrix} \nu_1 \\ \nu_2 \\ \epsilon_{3,t+1} \end{bmatrix}$$

5.4.2 At first order

Recall that we have:

$$\mathbf{x}^f_{t+1} = \mathbf{h}_0 + \mathbf{h}_\mathbf{x}\mathbf{x}^f_t + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}$$

and

$$\begin{aligned} \mathbf{x}^f_{t+2} &= \mathbf{h}_0 + \mathbf{h}_\mathbf{x}\mathbf{x}^f_{t+1} + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+2} \\ &= \mathbf{h}_0 + \mathbf{h}_\mathbf{x}\left(\mathbf{h}_0 + \mathbf{h}_\mathbf{x}\mathbf{x}^f_t + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right) + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+2} \\ &= \mathbf{h}_0 + \mathbf{h}_\mathbf{x}\mathbf{h}_0 + \mathbf{h}^2_\mathbf{x}\mathbf{x}^f_t + \mathbf{h}_\mathbf{x}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+2} \end{aligned}$$

and

$$\begin{aligned} \mathbf{x}^f_{t+3} &= \mathbf{h}_0 + \mathbf{h}_\mathbf{x}\mathbf{x}^f_{t+2} + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+3} \\ &= \mathbf{h}_0 + \mathbf{h}_\mathbf{x}\left(\mathbf{h}_0 + \mathbf{h}_\mathbf{x}\mathbf{h}_0 + \mathbf{h}^2_\mathbf{x}\mathbf{x}^f_t + \mathbf{h}_\mathbf{x}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+2}\right) + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+3} \\ &= \mathbf{h}_0 + \mathbf{h}_\mathbf{x}\mathbf{h}_0 + \mathbf{h}^2_\mathbf{x}\mathbf{h}_0 + \mathbf{h}^3_\mathbf{x}\mathbf{x}^f_t + \mathbf{h}^2_\mathbf{x}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} + \mathbf{h}_\mathbf{x}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+2} + \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+3} \\ &= \mathbf{h}^3_\mathbf{x}\mathbf{x}^f_t + \sum_{j=1}^{3} \mathbf{h}^{3-j}_\mathbf{x}\mathbf{h}_0 + \sum_{j=1}^{3} \mathbf{h}^{3-j}_\mathbf{x}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \end{aligned}$$

In general

$$\mathbf{x}^f_{t+l} = \mathbf{h}^l_\mathbf{x}\mathbf{x}^f_t + \sum_{j=1}^{l} \mathbf{h}^{l-j}_\mathbf{x}\mathbf{h}_0 + \sum_{j=1}^{l} \mathbf{h}^{l-j}_\mathbf{x}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}$$

With a shock of $\boldsymbol{\nu}$ in period $t+1$, we have

$$\tilde{\mathbf{x}}_{t+l}^f = \mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j} \mathbf{h}_0 + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j} \boldsymbol{\eta} \boldsymbol{\delta}_{t+j}$$

where we define $\boldsymbol{\delta}_t$ such that

$$\boldsymbol{\delta}_{t+j} = \begin{cases} \mathbf{S}\boldsymbol{\nu} + (\mathbf{I} - \mathbf{S}) \boldsymbol{\epsilon}_{t+1} & \text{for } j = 1 \\ \boldsymbol{\epsilon}_{t+j} & \text{for } j \neq 1 \end{cases}$$

Agents know the size of the shock $\boldsymbol{\nu}$ at time $t+1$, and it is therefore in agents' information set. I.e. $\boldsymbol{\nu}$ is non-stochastic. So

$$E_t \left[\tilde{\mathbf{x}}_{t+l}^f - \mathbf{x}_{t+l}^f \right] = E_t \left[\sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j} \boldsymbol{\eta} \boldsymbol{\delta}_{t+j} - \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right]$$

$$= E_t \left[\mathbf{h}_\mathbf{x}^{l-1} \boldsymbol{\eta} \left(\mathbf{S}\boldsymbol{\nu} + (\mathbf{I} - \mathbf{S}) \boldsymbol{\epsilon}_{t+1} \right) \right]$$

$$= \mathbf{h}_\mathbf{x}^{l-1} \boldsymbol{\eta} \mathbf{S} \boldsymbol{\nu}$$

and

$$E_t \left[\tilde{\mathbf{y}}_{t+l}^f - \mathbf{y}_{t+l}^f \right] = \mathbf{g}_\mathbf{x} E_t \left[\tilde{\mathbf{x}}_{t+l}^f - \mathbf{x}_{t+l}^f \right]$$

5.4.3 At second order

We need to consider:

$$\mathbf{x}_{t+1}^s = \mathbf{h}_\mathbf{x} \mathbf{x}_t^s + \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f \right)$$

$$\begin{aligned}
\mathbf{x}_{t+2}^s &= \mathbf{h}_\mathbf{x} \mathbf{x}_{t+1}^s + \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_{t+1}^f \otimes \mathbf{x}_{t+1}^f \right) \\
&= \mathbf{h}_\mathbf{x} \left(\mathbf{h}_\mathbf{x} \mathbf{x}_t^s + \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f \right) \right) + \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_{t+1}^f \otimes \mathbf{x}_{t+1}^f \right) \\
&= \mathbf{h}_\mathbf{x}^2 \mathbf{x}_t^s + \mathbf{h}_\mathbf{x} \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f \right) + \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_{t+1}^f \otimes \mathbf{x}_{t+1}^f \right)
\end{aligned}$$

$$\begin{aligned}
\mathbf{x}_{t+3}^s &= \mathbf{h}_\mathbf{x} \mathbf{x}_{t+2}^s + \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_{t+2}^f \otimes \mathbf{x}_{t+2}^f \right) \\
&= \mathbf{h}_\mathbf{x} \left(\mathbf{h}_\mathbf{x}^2 \mathbf{x}_t^s + \mathbf{h}_\mathbf{x} \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f \right) + \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_{t+1}^f \otimes \mathbf{x}_{t+1}^f \right) \right) \\
&\quad + \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_{t+2}^f \otimes \mathbf{x}_{t+2}^f \right) \\
&= \mathbf{h}_\mathbf{x}^3 \mathbf{x}_t^s + \mathbf{h}_\mathbf{x}^2 \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f \right) + \mathbf{h}_\mathbf{x} \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_{t+1}^f \otimes \mathbf{x}_{t+1}^f \right) \\
&\quad + \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_{t+2}^f \otimes \mathbf{x}_{t+2}^f \right) \\
&= \mathbf{h}_\mathbf{x}^3 \mathbf{x}_t^s + \mathbf{h}_\mathbf{x}^2 \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f \right) + \mathbf{h}_\mathbf{x} \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_{t+1}^f \otimes \mathbf{x}_{t+1}^f \right) + \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_{t+2}^f \otimes \mathbf{x}_{t+2}^f \right) \\
&= \mathbf{h}_\mathbf{x}^3 \mathbf{x}_t^s + \sum_{j=0}^{2} \mathbf{h}_\mathbf{x}^{2-j} \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_{t+j}^f \otimes \mathbf{x}_{t+j}^f \right)
\end{aligned}$$

and in general

$$\mathbf{x}_{t+l}^s = \mathbf{h}_\mathbf{x}^l \mathbf{x}_t^s + \sum_{j=0}^{l-1} \mathbf{h}_\mathbf{x}^{(l-1)-j} \mathbf{H}_{\mathbf{xx}} \left(\mathbf{x}_{t+j}^f \otimes \mathbf{x}_{t+j}^f \right)$$

for $l = 1, 2, 3, \ldots$

Thus, to compute $E_t \left[\tilde{\mathbf{x}}_{t+l}^s - \mathbf{x}_{t+l}^s \right]$, we need to find $E_t \left[\tilde{\mathbf{x}}_{t+l}^f \otimes \tilde{\mathbf{x}}_{t+l}^f - \mathbf{x}_{t+l}^f \otimes \mathbf{x}_{t+l}^f \right]$. Hence, consider:

$$\mathbf{x}_{t+l}^f \otimes \mathbf{x}_{t+l}^f = \left(\mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j} \mathbf{h}_0 + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right) \otimes \left(\mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j} \mathbf{h}_0 + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right)$$

$$
\begin{aligned}
&= \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} \otimes \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} + \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} + \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \\
&+ \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} \otimes \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} + \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} + \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \\
&+ \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} + \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} + \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}
\end{aligned}
$$

and

$$
\begin{aligned}
\tilde{\mathbf{x}}_{t+l}^{f} \otimes \tilde{\mathbf{x}}_{t+l}^{f} &= \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} \otimes \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} + \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} + \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \\
&+ \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} \otimes \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} + \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} + \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \\
&+ \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \otimes \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} + \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} + \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j}
\end{aligned}
$$

where we define $\boldsymbol{\delta}_t$ such that:

$$
\boldsymbol{\delta}_{t+j} = \begin{cases} \mathbf{S}\boldsymbol{\nu} + (\mathbf{I}-\mathbf{S})\,\boldsymbol{\epsilon}_{t+1} & \text{for } j=1 \\ \boldsymbol{\epsilon}_{t+j} & \text{for } j \neq 1 \end{cases}
$$

This means that:

$$
E_t\left[\tilde{\mathbf{x}}_{t+l}^{f} \otimes \tilde{\mathbf{x}}_{t+l}^{f} - \mathbf{x}_{t+l}^{f} \otimes \mathbf{x}_{t+l}^{f}\right]
$$

$$
\begin{aligned}
&= E_t[\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} \otimes \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} + \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} + \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \\
&+ \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} \otimes \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} + \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} + \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \\
&+ \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \otimes \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} + \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} + \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \\
&- \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} \otimes \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} - \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} - \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \\
&- \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} \otimes \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} - \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} - \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \\
&- \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} - \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} - \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}]
\end{aligned}
$$

$$
\begin{aligned}
&= E_t[\mathbf{0} + \mathbf{0} + \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \\
&+ 0 + 0 + \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \\
&+ \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \otimes \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} + \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} + \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \\
&- \mathbf{0} - \mathbf{0} - \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \\
&- 0 - 0 - \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \\
&- \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_{t}^{f} - \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_{0} - \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}]
\end{aligned}
$$

$$= E_t[\mathbf{h}_{\mathbf{x}}^l \mathbf{x}_t^f \otimes \sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\delta}_{t+j}$$
$$\sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \mathbf{h}_0 \otimes \sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\delta}_{t+j}$$
$$+ \sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\delta}_{t+j} \otimes \mathbf{h}_{\mathbf{x}}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\delta}_{t+j} \otimes \sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \mathbf{h}_0 + \sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\delta}_{t+j} \otimes \sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\delta}_{t+j}$$
$$-0$$
$$-0$$
$$-0 - 0 - \sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \otimes \sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j}]$$

as $E_t \left[\boldsymbol{\epsilon}_{t+j}\right] = \mathbf{0}$

$$= E_t[\mathbf{h}_{\mathbf{x}}^l \mathbf{x}_t^f \otimes \mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \mathbf{S} \boldsymbol{\nu}$$
$$\sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \mathbf{h}_0 \otimes \mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \mathbf{S} \boldsymbol{\nu}$$
$$+ \mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \mathbf{S} \boldsymbol{\nu} \otimes \mathbf{h}_{\mathbf{x}}^l \mathbf{x}_t^f + \mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \mathbf{S} \boldsymbol{\nu} \otimes \sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \mathbf{h}_0$$
$$+ \left(\mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \left(\mathbf{S} \boldsymbol{\nu} + (\mathbf{I} - \mathbf{S}) \boldsymbol{\epsilon}_{t+1} \right) + \sum_{j=2}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right) \otimes \left(\mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \left(\mathbf{S} \boldsymbol{\nu} + (\mathbf{I} - \mathbf{S}) \boldsymbol{\epsilon}_{t+1} \right) + \sum_{j=2}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right)$$
$$- \left(\mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} + \sum_{j=2}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right) \otimes \left(\mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} + \sum_{j=2}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right)]$$

using that $\boldsymbol{\delta}_{t+j} = \begin{cases} \mathbf{S} \boldsymbol{\nu} + (\mathbf{I} - \mathbf{S}) \boldsymbol{\epsilon}_{t+1} & \text{for } j = 1 \\ \boldsymbol{\epsilon}_{t+j} & \text{for } j \neq 1 \end{cases}$ and $E_t \left[\boldsymbol{\epsilon}_{t+j}\right] = \mathbf{0}$

$$= E_t[\mathbf{h}_{\mathbf{x}}^l \mathbf{x}_t^f \otimes \mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \mathbf{S} \boldsymbol{\nu} + \sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \mathbf{h}_0 \otimes \mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \mathbf{S} \boldsymbol{\nu} + \mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \mathbf{S} \boldsymbol{\nu} \otimes \mathbf{h}_{\mathbf{x}}^l \mathbf{x}_t^f + \mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \mathbf{S} \boldsymbol{\nu} \otimes \sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \mathbf{h}_0$$
$$+ \left(\mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \left(\mathbf{S} \boldsymbol{\nu} + (\mathbf{I} - \mathbf{S}) \boldsymbol{\epsilon}_{t+1} \right) + \sum_{j=2}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right) \otimes \left(\mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \left(\mathbf{S} \boldsymbol{\nu} + (\mathbf{I} - \mathbf{S}) \boldsymbol{\epsilon}_{t+1} \right) + \sum_{j=2}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right)$$
$$- \left(\mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} + \sum_{j=2}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right) \otimes \left(\mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} + \sum_{j=2}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right)]$$

$$= E_t[\left(\mathbf{h}_{\mathbf{x}}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \mathbf{h}_0 \right) \otimes \mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \mathbf{S} \boldsymbol{\nu} + \mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \mathbf{S} \boldsymbol{\nu} \otimes \left(\mathbf{h}_{\mathbf{x}}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \mathbf{h}_0 \right)$$
$$+ \left(\mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \left(\mathbf{S} \boldsymbol{\nu} + (\mathbf{I} - \mathbf{S}) \boldsymbol{\epsilon}_{t+1} \right) + \sum_{j=2}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right) \otimes \left(\mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \left(\mathbf{S} \boldsymbol{\nu} + (\mathbf{I} - \mathbf{S}) \boldsymbol{\epsilon}_{t+1} \right) + \sum_{j=2}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right)$$
$$- \left(\mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} + \sum_{j=2}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right) \otimes \left(\mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} + \sum_{j=2}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right)]$$

$$= E_t[\left(\mathbf{h}_{\mathbf{x}}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \mathbf{h}_0 \right) \otimes \mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \mathbf{S} \boldsymbol{\nu} + \mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \mathbf{S} \boldsymbol{\nu} \otimes \left(\mathbf{h}_{\mathbf{x}}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \mathbf{h}_0 \right)$$
$$+ \left(\mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \mathbf{S} \boldsymbol{\nu} + \mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \left(\mathbf{I} - \mathbf{S} \right) \boldsymbol{\epsilon}_{t+1} \right) \otimes \left(\mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \left(\mathbf{S} \boldsymbol{\nu} + (\mathbf{I} - \mathbf{S}) \boldsymbol{\epsilon}_{t+1} \right) + \sum_{j=2}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right)$$
$$+ \sum_{j=2}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \otimes \left(\mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \mathbf{S} \boldsymbol{\nu} + \mathbf{h}_{\mathbf{x}}^{l-1} \boldsymbol{\eta} \left(\mathbf{I} - \mathbf{S} \right) \boldsymbol{\epsilon}_{t+1} + \sum_{j=2}^{l} \mathbf{h}_{\mathbf{x}}^{l-j} \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+j} \right)$$

$$-\left(\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right)\otimes\left(\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}+\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\right)$$
$$-\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\otimes\left(\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}+\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\right)]$$

$$=E_t[\left(\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f+\sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_0\right)\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}+\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}\otimes\left(\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f+\sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_0\right)$$
$$+\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}\otimes\left(\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\left(\mathbf{S}\boldsymbol{\nu}+\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1}\right)+\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\right)$$
$$+\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1}\otimes\left(\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\left(\mathbf{S}\boldsymbol{\nu}+\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1}\right)+\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\right)$$
$$+\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\otimes\left(\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}+\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1}\right)$$
$$+\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\otimes\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}$$
$$-\left(\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right)\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}$$
$$-\left(\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right)\otimes\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}$$
$$-\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}$$
$$-\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\otimes\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}]$$

$$=E_t[\left(\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f+\sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_0\right)\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}+\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}\otimes\left(\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f+\sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_0\right)$$
$$+\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}+\mathbf{0}+\mathbf{0}$$
$$+\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1}\otimes\left(\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1}\right)+0$$
$$+0$$
$$+\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\otimes\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}$$
$$-\left(\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right)\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}$$
$$-0$$
$$-0$$
$$-\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\otimes\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}]$$

using $E_t\left[\boldsymbol{\epsilon}_{t+j}\right]=\mathbf{0}$ and $\boldsymbol{\epsilon}_{t+j}$ are uncorrelated across time

$$=E_t[\left(\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f+\sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_0\right)\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}+\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}\otimes\left(\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f+\sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\mathbf{h}_0\right)$$
$$+\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}+\mathbf{0}+\mathbf{0}$$
$$+\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1}\otimes\left(\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1}\right)+0$$
$$-\left(\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right)\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}]$$

canceling terms $\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\otimes\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}$

$$= E_t[\left(\mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j}\mathbf{h}_0\right) \otimes \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} + \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} \otimes \left(\mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j}\mathbf{h}_0\right) + \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} \otimes \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}$$
$$+ \left(\mathbf{h}_\mathbf{x}^{l-1} \otimes \mathbf{h}_\mathbf{x}^{l-1}\right)\left(\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1}\right)$$
$$- \left(\mathbf{h}_\mathbf{x}^{l-1} \otimes \mathbf{h}_\mathbf{x}^{l-1}\right)\left(\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right)]$$

$$= E_t[\left(\mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j}\mathbf{h}_0\right) \otimes \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} + \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} \otimes \left(\mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j}\mathbf{h}_0\right) + \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} \otimes \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}$$
$$+ \left(\mathbf{h}_\mathbf{x}^{l-1} \otimes \mathbf{h}_\mathbf{x}^{l-1}\right)\left(\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1} - \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right)]$$

$$= \left(\mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j}\mathbf{h}_0\right) \otimes \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} + \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} \otimes \left(\mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j}\mathbf{h}_0\right) + \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} \otimes \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}$$
$$+ \left(\mathbf{h}_\mathbf{x}^{l-1} \otimes \mathbf{h}_\mathbf{x}^{l-1}\right)\left(E_t\left[\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1}\right] - E_t\left[\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right]\right)$$

Hence, we only need to compute $E_t\left[\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1}\right]$ and $E_t\left[\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right]$. We know
$E_t\left[\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right] = E_t\left[\left(\boldsymbol{\eta} \otimes \boldsymbol{\eta}\right)\left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}\right)\right] = \left(\boldsymbol{\eta} \otimes \boldsymbol{\eta}\right) vec\left(\mathbf{I}\right)$
using $\left(\mathbf{A} \otimes \mathbf{B}\right)\left(\mathbf{C} \otimes \mathbf{D}\right) = \mathbf{AC} \otimes \mathbf{BD}$ if $\mathbf{AC}$ and $\mathbf{BD}$ are defined

and
$E_t\left[\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\boldsymbol{\epsilon}_{t+1}\right] = E_t\left[\left(\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right) \otimes \boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\right)\left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}\right)\right] = \left(\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right) \otimes \boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\right) vec\left(\mathbf{I}\right)$

Thus, we have
$$E_t\left[\tilde{\mathbf{x}}_{t+l}^f \otimes \tilde{\mathbf{x}}_{t+l}^f - \mathbf{x}_{t+l}^f \otimes \mathbf{x}_{t+l}^f\right] = \left(\mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j}\mathbf{h}_0\right) \otimes \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} + \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} \otimes \left(\mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j}\mathbf{h}_0\right) + \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} \otimes \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}$$
$$+ \left(\mathbf{h}_\mathbf{x}^{l-1} \otimes \mathbf{h}_\mathbf{x}^{l-1}\right)\left(\left(\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right) \otimes \boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\right) vec\left(\mathbf{I}\right) - \left(\boldsymbol{\eta} \otimes \boldsymbol{\eta}\right) vec\left(\mathbf{I}\right)\right)$$

$$= \left(\mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j}\mathbf{h}_0\right) \otimes \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} + \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} \otimes \left(\mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j}\mathbf{h}_0\right) + \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} \otimes \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}$$
$$+ \left(\mathbf{h}_\mathbf{x}^{l-1} \otimes \mathbf{h}_\mathbf{x}^{l-1}\right)\left(\left(\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right) \otimes \boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\right) - \left(\boldsymbol{\eta} \otimes \boldsymbol{\eta}\right)\right) vec\left(\mathbf{I}\right)$$

$$= \left(\mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j}\mathbf{h}_0\right) \otimes \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} + \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} \otimes \left(\mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j}\mathbf{h}_0\right)$$
$$+ \left(\mathbf{h}_\mathbf{x}^{l-1} \otimes \mathbf{h}_\mathbf{x}^{l-1}\right)\left[\left(\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} \otimes \boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}\right) + \left(\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right) \otimes \boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\right) vec\left(\mathbf{I}\right) - \left(\boldsymbol{\eta} \otimes \boldsymbol{\eta}\right) vec\left(\mathbf{I}\right)\right]$$

$$= \left(\mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j}\mathbf{h}_0\right) \otimes \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} + \mathbf{h}_\mathbf{x}^{l-1}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} \otimes \left(\mathbf{h}_\mathbf{x}^l \mathbf{x}_t^f + \sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j}\mathbf{h}_0\right)$$
$$+ \left(\mathbf{h}_\mathbf{x}^{l-1} \otimes \mathbf{h}_\mathbf{x}^{l-1}\right)\left[\left(\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu} \otimes \boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}\right) + \boldsymbol{\Lambda}\right]$$
where
$$\boldsymbol{\Lambda} \equiv \left(\left(\boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right) \otimes \boldsymbol{\eta}\left(\mathbf{I}-\mathbf{S}\right)\right) - \left(\boldsymbol{\eta} \otimes \boldsymbol{\eta}\right)\right) vec\left(\mathbf{I}\right)$$

Then note that
$$\sum_{j=1}^{l} \mathbf{h}_\mathbf{x}^{l-j} = \mathbf{h}_\mathbf{x}^{l-1} + \mathbf{h}_\mathbf{x}^{l-2} + ... + \mathbf{h}_\mathbf{x}^{0} = \left(\mathbf{I}-\mathbf{h}_\mathbf{x}\right)^{-1}\left(\mathbf{I}-\mathbf{h}_\mathbf{x}^{l}\right)$$
as we recall the general result that
$$Q = \left(\mathbf{h}_\mathbf{x}\right)^0 + \left(\mathbf{h}_\mathbf{x}\right)^1 + ... + \left(\mathbf{h}_\mathbf{x}\right)^{11} = \left(\mathbf{I}-\mathbf{h}_\mathbf{x}\right)^{-1}\left(\mathbf{I}-\left(\mathbf{h}_\mathbf{x}\right)^{12}\right).$$

Hence, we have

$$E_t\left[\tilde{\mathbf{x}}^f_{t+l}\otimes\tilde{\mathbf{x}}^f_{t+l}-\mathbf{x}^f_{t+l}\otimes\mathbf{x}^f_{t+l}\right]=\left(\mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t+(\mathbf{I}-\mathbf{h}_{\mathbf{x}})^{-1}\left(\mathbf{I}-\mathbf{h}^l_{\mathbf{x}}\right)\mathbf{h}_0\right)\otimes\mathbf{h}^{l-1}_{\mathbf{x}}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}+\mathbf{h}^{l-1}_{\mathbf{x}}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}\otimes\left(\mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t+(\mathbf{I}-\mathbf{h}_{\mathbf{x}})^{-1}\left(\mathbf{I}-\mathbf{h}^l_{\mathbf{x}}\right)\mathbf{h}_0\right)$$
$$+\left(\mathbf{h}^{l-1}_{\mathbf{x}}\otimes\mathbf{h}^{l-1}_{\mathbf{x}}\right)\left[(\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}\otimes\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu})+\boldsymbol{\Lambda}\right]$$

Or (using another index)

$$E_t\left[\tilde{\mathbf{x}}^f_{t+j}\otimes\tilde{\mathbf{x}}^f_{t+j}-\mathbf{x}^f_{t+j}\otimes\mathbf{x}^f_{t+j}\right]=\left(\mathbf{h}^j_{\mathbf{x}}\mathbf{x}^f_t+(\mathbf{I}-\mathbf{h}_{\mathbf{x}})^{-1}\left(\mathbf{I}-\mathbf{h}^j_{\mathbf{x}}\right)\mathbf{h}_0\right)\otimes\mathbf{h}^{j-1}_{\mathbf{x}}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}+\mathbf{h}^{j-1}_{\mathbf{x}}\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}\otimes\left(\mathbf{h}^j_{\mathbf{x}}\mathbf{x}^f_t+(\mathbf{I}-\mathbf{h}_{\mathbf{x}})^{-1}\left(\mathbf{I}-\mathbf{h}^j_{\mathbf{x}}\right)\mathbf{h}_0\right)$$
$$+\left(\mathbf{h}^{j-1}_{\mathbf{x}}\otimes\mathbf{h}^{j-1}_{\mathbf{x}}\right)\left[(\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu}\otimes\boldsymbol{\eta}\mathbf{S}\boldsymbol{\nu})+\boldsymbol{\Lambda}\right]$$

for $j=1,2,3,...$

Thus, we have in general

$$E_t\left[\tilde{\mathbf{x}}^s_{t+l}-\mathbf{x}^s_{t+l}\right]=E_t\left[\sum_{j=0}^{l-1}\mathbf{h}^{l-1-j}_{\mathbf{x}}\mathbf{H}_{\mathbf{xx}}\left(\tilde{\mathbf{x}}^f_{t+j}\otimes\tilde{\mathbf{x}}^f_{t+j}\right)-\sum_{j=0}^{l-1}\mathbf{h}^{l-1-j}_{\mathbf{x}}\mathbf{H}_{\mathbf{xx}}\left(\mathbf{x}^f_{t+j}\otimes\mathbf{x}^f_{t+j}\right)\right]$$

$$=\sum_{j=1}^{l-1}\mathbf{h}^{l-1-j}_{\mathbf{x}}\mathbf{H}_{\mathbf{xx}}E_t\left[\tilde{\mathbf{x}}^f_{t+j}\otimes\tilde{\mathbf{x}}^f_{t+j}-\mathbf{x}^f_{t+j}\otimes\mathbf{x}^f_{t+j}\right]$$

the shock hits in period $t+1$, so $\left(\tilde{\mathbf{x}}^f_t\otimes\tilde{\mathbf{x}}^f_t\right)=\mathbf{x}^f_t\otimes\mathbf{x}^f_t$

When implementing the GIRF, it may be useful to have a recursive expression. Here, it is most convenient to use the general expression

$$E_t\left[\tilde{\mathbf{x}}^s_{t+l}-\mathbf{x}^s_{t+l}\right]=\sum_{j=1}^{l-1}\mathbf{h}^{l-1-j}_{\mathbf{x}}\mathbf{H}_{\mathbf{xx}}E_t\left[\tilde{\mathbf{x}}^f_{t+j}\otimes\tilde{\mathbf{x}}^f_{t+j}-\mathbf{x}^f_{t+j}\otimes\mathbf{x}^f_{t+j}\right]$$

So

$$E_t\left[\tilde{\mathbf{x}}^s_{t+1}-\mathbf{x}^s_{t+1}\right]=0$$

$$E_t\left[\tilde{\mathbf{x}}^s_{t+2}-\mathbf{x}^s_{t+2}\right]=\sum_{j=1}^{1}\mathbf{h}^{1-j}_{\mathbf{x}}\mathbf{H}_{\mathbf{xx}}E_t\left[\tilde{\mathbf{x}}^f_{t+j}\otimes\tilde{\mathbf{x}}^f_{t+j}-\mathbf{x}^f_{t+j}\otimes\mathbf{x}^f_{t+j}\right]$$
$$=\mathbf{H}_{\mathbf{xx}}E_t\left[\tilde{\mathbf{x}}^f_{t+1}\otimes\tilde{\mathbf{x}}^f_{t+1}-\mathbf{x}^f_{t+1}\otimes\mathbf{x}^f_{t+1}\right]$$

$$E_t\left[\tilde{\mathbf{x}}^s_{t+3}-\mathbf{x}^s_{t+3}\right]=\sum_{j=1}^{2}\mathbf{h}^{2-j}_{\mathbf{x}}\mathbf{H}_{\mathbf{xx}}E_t\left[\tilde{\mathbf{x}}^f_{t+j}\otimes\tilde{\mathbf{x}}^f_{t+j}-\mathbf{x}^f_{t+j}\otimes\mathbf{x}^f_{t+j}\right]$$
$$=\mathbf{h}_{\mathbf{x}}\mathbf{H}_{\mathbf{xx}}E_t\left[\tilde{\mathbf{x}}^f_{t+1}\otimes\tilde{\mathbf{x}}^f_{t+1}-\mathbf{x}^f_{t+1}\otimes\mathbf{x}^f_{t+1}\right]$$
$$+\mathbf{H}_{\mathbf{xx}}E_t\left[\tilde{\mathbf{x}}^f_{t+2}\otimes\tilde{\mathbf{x}}^f_{t+2}-\mathbf{x}^f_{t+2}\otimes\mathbf{x}^f_{t+2}\right]$$
$$=\mathbf{h}_{\mathbf{x}}E_t\left[\tilde{\mathbf{x}}^s_{t+2}-\mathbf{x}^s_{t+2}\right]+\mathbf{H}_{\mathbf{xx}}E_t\left[\tilde{\mathbf{x}}^f_{t+2}\otimes\tilde{\mathbf{x}}^f_{t+2}-\mathbf{x}^f_{t+2}\otimes\mathbf{x}^f_{t+2}\right]$$

So in general

$$E_t\left[\tilde{\mathbf{x}}^s_{t+k}-\mathbf{x}^s_{t+k}\right]=\mathbf{h}_{\mathbf{x}}E_t\left[\tilde{\mathbf{x}}^s_{t+k-1}-\mathbf{x}^s_{t+k-1}\right]+\mathbf{H}_{\mathbf{xx}}E_t\left[\tilde{\mathbf{x}}^f_{t+k-1}\otimes\tilde{\mathbf{x}}^f_{t+k-1}-\mathbf{x}^f_{t+k-1}\otimes\mathbf{x}^f_{t+k-1}\right]$$

For the total state variable:

$$E_t\left[\tilde{\mathbf{x}}_{t+l}-\mathbf{x}_{t+l}\right]=E_t\left[\tilde{\mathbf{x}}^f_{t+l}-\mathbf{x}^f_{t+l}\right]+E_t\left[\tilde{\mathbf{x}}^s_{t+l}-\mathbf{x}^s_{t+l}\right]$$

For the control variables:

$$\mathbf{y}^s_{t+l}=\mathbf{g}_0+\mathbf{g}_{\mathbf{x}}\left(\mathbf{x}^f_{t+l}+\mathbf{x}^s_{t+l}\right)+\mathbf{G}_{\mathbf{xx}}\left(\mathbf{x}^f_{t+l}\otimes\mathbf{x}^f_{t+l}\right)$$
$$\tilde{\mathbf{y}}^s_{t+l}=\mathbf{g}_0+\mathbf{g}_{\mathbf{x}}\left(\tilde{\mathbf{x}}^f_{t+l}+\tilde{\mathbf{x}}^s_{t+l}\right)+\mathbf{G}_{\mathbf{xx}}\left(\tilde{\mathbf{x}}^f_{t+l}\otimes\tilde{\mathbf{x}}^f_{t+l}\right)$$

$$E_t\left[\tilde{\mathbf{y}}^s_{t+l}-\mathbf{y}^s_{t+l}\right]=\mathbf{g}_{\mathbf{x}}\left(E_t\left[\tilde{\mathbf{x}}^f_{t+l}-\mathbf{x}^f_{t+l}\right]+E_t\left[\tilde{\mathbf{x}}^s_{t+l}-\mathbf{x}^s_{t+l}\right]\right)+\mathbf{G}_{\mathbf{xx}}E_t\left[\tilde{\mathbf{x}}^f_{t+l}\otimes\tilde{\mathbf{x}}^f_{t+l}-\mathbf{x}^f_{t+l}\otimes\mathbf{x}^f_{t+l}\right]$$

6 Additional Model Output

This section reports additional output from the model.

6.1 Correlation matrix for the estimated state innovations

The table below shows the correlation matrix for the estimated states in the baseline model, i.e., $\mathcal{M}^{M,CS}$.

	$\epsilon_{\mu_z,t}$	$\epsilon_{d,t}$	$\epsilon_{n,t}$	$\epsilon_{\pi^*,t}$	$\epsilon_{a,t}$
$\epsilon_{\mu_z,t}$	1	-0.12	0.12	0.10	0.09
$\epsilon_{d,t}$		1	-0.21	-0.11	0.08
$\epsilon_{n,t}$			1	0.03	0.26
$\epsilon_{\pi^*,t}$				1	0.07
$\epsilon_{a,t}$					1

6.2 Correlation matrix for the estimated measurement errors

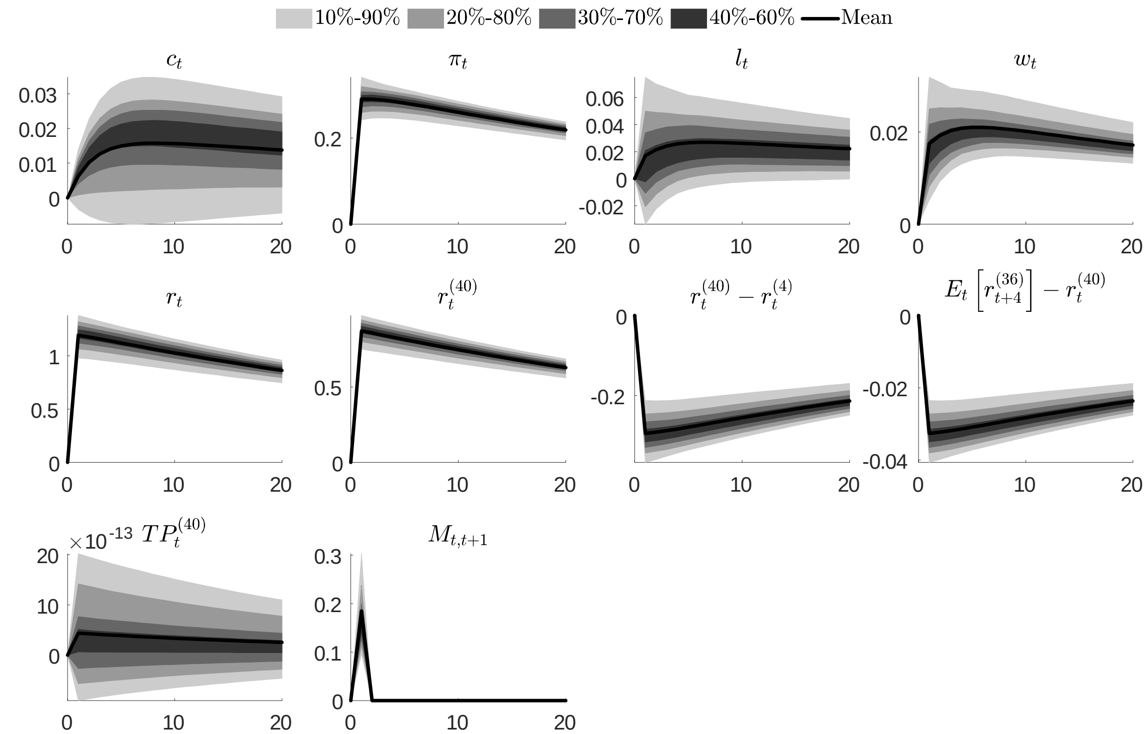
The table below shows the correlation matrix for the estimated measurement errors in by the baseline model, i.e., $\mathcal{M}^{M,CS}$.

	$\hat{l}_t$	$\hat{w}_t$	Δc_t	π_t	r_t	$r_t^{(4)}$	$r_t^{(12)}$	$r_t^{(20)}$	$r_t^{(28)}$	$r_t^{(40)}$
$\hat{l}_t$	1.00	0.23	0.44	0.25	-0.16	-0.26	-0.09	0.07	0.08	-0.11
$\hat{w}_t$	-	1.00	-0.11	-0.44	0.03	0.29	0.25	0.06	-0.21	-0.42
Δc_t	-	-	1.00	0.23	-0.23	-0.50	-0.55	-0.46	-0.35	-0.17
π_t	-	-	-	1.00	-0.09	-0.11	0.03	0.20	0.34	0.32
r_t	-	-	-	-	1.00	0.09	0.05	0.05	0.07	0.15
$r_t^{(4)}$	-	-	-	-	-	1.00	0.72	0.26	-0.08	-0.20
$r_t^{(12)}$	-	-	-	-	-	-	1.00	0.80	0.41	-0.13
$r_t^{(20)}$	-	-	-	-	-	-	-	1.00	0.83	0.12
$r_t^{(28)}$	-	-	-	-	-	-	-	-	1.00	0.58
$r_t^{(40)}$	-	-	-	-	-	-	-	-	-	1.00

6.3 Impulse Reponse Functions to Demand Shocks for a perfect foreslight model solution

Figure 1: A Demand Shock under Perfect Foresight (no Risk)

This figure shows the generalized impulse response functions (IRFs) of a positive one-standard deviation shock to d_t, where the various shadings cover the indicated fraction of the distribution obtained using 1,000 randomly generated initial states. These IRFs are computed in closed-form for the second order projection solution using the approach in Andreasen et al. (2018), except for $M_{t,t+1}$ that is computed using Monte Carlo integration with 1,000 draws. Except for the term premium $TP_t^{(40)}$ and the nominal stochastic discount factor $M_{t,t+1}$, all impulse response functions are expressed in percentage deviations from the steady state (i.e., scaled by 100), with consumption expressed in deviations from the balanced growth path and all bond yields and inflation measured in annualized terms. The term premium $TP_t^{(40)}$ is expressed in annualized basis points.



References

Andreasen, M. M. (2013), 'Non-Linear DSGE Models and the Central Difference Kalman Filter', *Journal of Applied Econometrics* **28**, 929–955.

Andreasen, M. M., Fernandez-Villaverde, J. & Rubio-Ramirez, J. F. (2018), 'The Pruned State Space System for Non-Linear DSGE Models: Theory and Empirical Applications to Estimation', *Review of Economic Studies* **85**(1), 1–49.

Andreasen, M. M. & Jørgensen, K. (2020), 'The Importance of Timing Attitudes in Consumption-Based Asset Pricing Models', *Journal of Monetary Economics* **111**, 95–117.

Blanchard, O. & Gali, J. (2005), 'Real Wage Rigidities and the New Keynesian Model', *Journal of Money, Credit and Banking* **39**(1), 35–65.

Chernozhukov, V. & Hong, H. (2003), 'An MCMC Approach to Classical Estimation', *Journal of Econometrics* **115**(2), 293–346.

Christiano, L. J., Trabandt, M. & Walentin, K. (2011), 'Introducing Financial Frictions and Unemployment into a Small Open Economy Model', *Journal of Economic Dynamic and Control* **35**(12), 1999–2041.

Hamilton, J. D. (2018), 'Why You Should Never Use the Hodrick-Prescott Filter', *The Review of Economics and Statistics* **100**(5), 831–843.

Hansen, L. P. (1982), 'Large Sample Properties of Generalized Method of Moments Estimators', *Econometrica* **50**(4), 1029–1054.

Hayashi, F. (2000), *Econometrics*, Princeton University Press, New Jersey.

Newey, W. K. & West, K. D. (1987), 'A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix', *Econometrica* **55**(3), 703–708.

Norgaard, M., Poulsen, N. K. & Ravn, O. (2000), 'Advances in Derivative-Free State Estimation for Nonlinear Systems', *Automatica* **36**(11), 1627–1638.

Rudebusch, G. D. & Swanson, E. T. (2008), 'Examining the Bond Premium Puzzle with a DSGE Model', *Journal of Monetary Economics* **55**, 111–126.

Rudebusch, G. D. & Swanson, E. T. (2012), 'The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks', *American Economic Journal: Macroeconomics* **4**(1), 1–43.

Swanson, E. (2012), 'Risk aversion and the labor margin in dynamic equilibrium models', *Working Paper version* .