Inflation, Risk Premia, and the Business Cycle*

J. R. Scott†

MIT Sloan

December 2023

Click Here for Latest Version

Abstract

This paper documents and explores the implications of the *risk price puzzle*—the empirical disconnect between inflation and risk premium shocks. I show that existing New Keynesian models struggle to rationalize the risk price puzzle with an upward-sloping Phillips curve. To resolve the puzzle, I develop a novel macro-finance model that integrates a two-sector real business cycle framework with the government debt valuation equation, which determines the price level without nominal rigidities. In the model, risk premium shocks generate the business-cycle comovement of macroeconomic quantities without implying counterfactual inflationary dynamics. Empirically, the response of inflation to risk premium shocks switches from positive to negative around 1998, mirroring the change in the stock-bond correlation. The model attributes this phenomenon to the changing covariance between shocks to the risk premium and real risk-free rate, which is consistent with both the heightened responsiveness of monetary policy to the stock market and the increasing prominence of the flight-to-safety phenomenon.

*I am deeply grateful to Emil Verner, Jonathan Parker, Larry Schmidt, and Adrien Verdelhan for their invaluable guidance and advice. This paper benefited greatly from conversations with Pat Adams, Tomás Caravello, Tim de Silva, Martin Eichenbaum, Dan Greenwald, Debbie Lucas, Rodney Ramcharan, and Eric Rasmusen, along with helpful comments from seminar audiences at MIT Sloan.

†100 Main Street, E62-388, Cambridge, MA 02142. Email: justinsc@mit.edu

1 Introduction

While an extensive literature has examined the effects of monetary policy shocks on macroeconomic outcomes, most of the variation in discount rates arises from innovations in risk premia, not risk-free rates (Campbell & Ammer (1993)). Motivated by this finding, a macro-finance research agenda has sought to jointly match asset pricing and macroeconomic dynamics by incorporating time-varying risk premia into New Keynesian models.[1] In these models, shocks to risk premia function as aggregate demand shocks that generate procyclical movements in macroeconomic quantities and inflation.

In this paper, I document a robust empirical pattern: risk premium shocks have statistically and economically significant effects on macroeconomic quantities, but no effect on inflation. Figure 1 shows the empirical impulse responses of output, consumption, investment, unemployment, and inflation to a one-standard deviation risk premium shock.[2] Consistent with the demand shock logic, an increase in the risk premium raises the unemployment rate and reduces output, consumption, and investment. Inflation, however, appears to be largely disconnected from fluctuations in risk premia, and, if anything, increases in risk premia lead to *higher* inflation. I refer to this empirical disconnect between risk premia and inflation as the *risk price puzzle*, akin to the longstanding "price puzzle" in which an unexpected tightening of monetary policy predicts higher inflation (Sims (1992)).

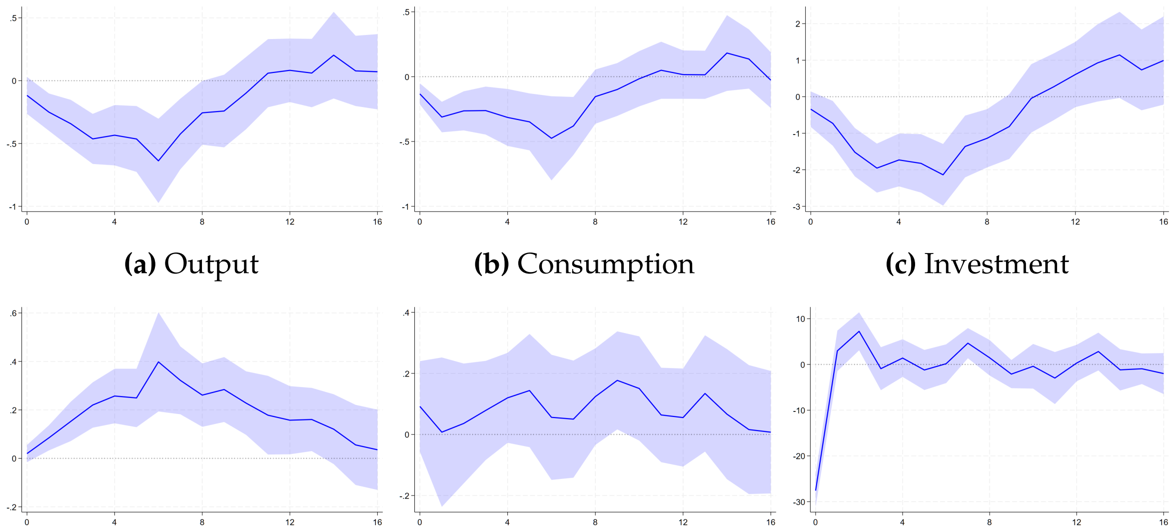
The existing New Keynesian literature has rationalized the weak relationship between inflation and aggregate demand via flat structural Phillips curves. Notwithstanding the considerable uncertainty about the slope of the Phillips curve, I show that a benchmark dynamic stochastic general equilibrium (DSGE) model exhibits the risk price puzzle even if the Phillips curve is nearly flat. In particular, when the magnitude of the risk premium shock is calibrated to generate an empirically realistic stock return response, inflation falls substantially after a positive risk premium shock. This suggests looking beyond the New Keynesian paradigm for a model in which the relationship between inflation and aggregate demand is not governed by the Phillips curve.

To resolve the risk price puzzle, I develop a novel macro-finance model that integrates a two-sector real business cycle (RBC) framework with the *government debt valuation equation*, which determines the price level without nominal rigidities. The

[1] See, for example, Dew-Becker (2014), Swanson (2016), Campbell et al. (2020), Pflueger & Rinaldi (2022), and Pflueger (2023).

[2] Risk premium shocks are the first principal component of shocks measured using three different approaches: (i) the Volatility Financial Conditions Index of Adrian et al. (2023); (ii) the news decomposition of Meeuwis et al. (2023); and (iii) the structural VAR of Cieslak & Pang (2021). Appendix B shows that similar impulse responses are obtained when each measure is used individually.

Figure 1: Empirical Impulse Responses to Risk Premium Shock



(a) Output **(b)** Consumption **(c)** Investment

(d) Unemployment **(e)** Inflation **(f)** Real Stock Return

Note: The figures above plot the empirical local-projection impulse responses of output, consumption, investment, unemployment, inflation, and the real stock return to a one-standard deviation risk premium shock. Risk premium shocks are measured as the first principal component of shocks to the Volatility Financial Conditions Index of Adrian et al. (2023), discount rate news from Meeuwis et al. (2023), and structural risk premium shocks from Cieslak & Pang (2021). The y-axis measures the response in percentage points, while the x-axis represents the number of quarters after the shock. Shaded areas are 90% confidence intervals. The sample is from 1968Q4 to 2023Q2.

government debt valuation equation implies that the price level adjusts so that the real market value of government debt is equal to the present discounted value of future primary (i.e., net of interest) surpluses. To gain intuition, assume that the expected return on government debt and the surplus growth rate are constant. Then, the following analogue of the Gordon growth model for government debt holds:

$$\frac{B_{t-1}}{P_t} = \frac{\mathbb{E}_t\left[S_{t+1}\right]}{R^b - G},$$

where B_{t-1} is the amount of nominal bonds outstanding, P_t is the price level, and S_t is the surplus with growth rate G. Since tax revenues are procyclical and government spending is countercyclical, surpluses covary with output. Thus, a positive risk premium shock that reduces output as in Figure 1 will lead to lower current or expected future surpluses. Since B_{t-1} is predetermined at time t, a lower value of $\mathbb{E}_t\left[S_{t+1}\right]$ implies that the price level must increase. I refer to the impact of risk premium shocks on surpluses as the *cash flow channel* of inflation. Risk premium shocks can also affect inflation by changing the discount rate on government debt, R^b. Since surpluses covary negatively with the stochastic discount factor, the government debt portfolio earns a risk premium (Jiang et al. (2019)). Thus, a positive risk premium shock will raise R^b and lead to an increase in the price level. I refer to this as the *discount rate channel* of inflation.

The real side of the model builds upon the production-based asset pricing literature that incorporates time-varying risk premia into RBC models. To introduce heterogeneity, I deploy a two-agent framework in which a representative shareholder owns all financial wealth in the economy, while a representative worker supplies labor. Asset prices are determined by the stochastic discount factor (SDF) of the representative shareholder with a time-varying price of risk and a time-varying time discount factor. Since an SDF reflects both preferences and beliefs, an increase in the price of risk can be interpreted as either an increase in risk aversion or greater pessimism about future consumption. Employment dynamics follow from a standard Diamond–Mortensen–Pissarides (DMP) search model, and frictions in forming or severing employment relationships imply that labor, like capital, is a risky investment good. Thus, a higher price of risk raises the discount rates for capital and workers, reducing investment and labor demand. The inclusion of both capital and labor in the production function is crucial to break the separation result of Tallarini (2000), which suggests that risk premia have a negligible effect on macroeconomic quantities. On the contrary, investment and unemployment in the model are highly sensitive to the magnitude and volatility of the price of risk.

As explained by Barro & King (1984), RBC models have difficulty rationalizing the simultaneous reduction in consumption, investment, and employment in response to a non-technology shock. Intuitively, consumption smoothing causes investment and employment to respond more negatively than output to a positive risk premium shock. Hence, dividends—which are approximately equal to output minus investment—increase. Since the representative shareholder consumes the dividends of the firm, a positive risk premium shock will lead to an *increase* in shareholder consumption, all else equal. In order to generate realistic comovement between consumption and other macroeconomic quantities, I model wealth effects by solving for the representative shareholder's optimal consumption rule.

In order to simulate time-varying risk premia, I solve the model using a third-order perturbation around the non-stochastic steady state. I estimate the model's parameters by matching the empirical impulse responses to a risk premium shock in Figure 1. The model successfully matches the level and volatility of the equity premium with an average risk aversion of 20, which is low relative to existing macro-finance models. Importantly, the model avoids generating the risk price puzzle, as a positive risk premium shock leads to a small yet positive increase in inflation. Finally, even though the estimation procedure does not directly target macroeconomic moments, the model-implied volatilities of output, consumption, investment, and inflation are similar to those in the data, although the model generates the Shimer (2005) puzzle of too little volatility in unemployment.

The unconditional evidence in Figure 1 masks the instability of the risk price puzzle over time; in particular, risk premium shocks are associated with higher inflation before 1998 and lower inflation thereafter. This changing relationship coincides with a similar shift in the sign of the stock-bond correlation, in which government bonds transformed from risky (positive beta) to hedging (negative beta) assets. In the model, the time-varying correlation between shocks to the risk premium and real risk-free rate jointly explains the changing dynamics of the risk price puzzle and the stock-bond correlation. When risk premium and risk-free rate shocks are positively correlated, the discount rate on government debt increases, and the government debt valuation equation predicts higher inflation. Alternatively, when risk premium shocks are negatively correlated with risk-free rate shocks, the higher risk premium on government debt is offset by a lower risk-free rate, which can lead to disinflation. The time-varying correlation between the shocks is consistent with both the heightened responsiveness of monetary policy to the stock market since the mid-1990s (Cieslak & Vissing-Jorgensen (2021)) and the increasing prominence of the flight-to-safety phenomenon due to the shortage of safe assets (Caballero et al. (2017)).

The rest of the paper proceeds as follows. Section 2 documents the risk price puzzle using three state-of-the art methods for inferring risk premium shocks. Section 3 explores how existing New Keynesian models have difficulty rationalizing the risk price puzzle with an upward-sloping Phillips curve. Section 4 integrates a two-sector RBC framework with the government debt valuation equation. Section 5 details the estimation procedure and discusses the model-implied impulse responses to a risk premium shock. Finally, Section 6 concludes with suggestions for further research.

Related Literature

While calls to integrate finance and macroeconomics date back to at least Fischer & Merton (1984), only recently have researchers made strides in incorporating time-varying risk premia into macroeconomic models. Campbell et al. (2020) integrate the external habit preferences of Campbell & Cochrane (1999) into a standard three-equation New Keynesian model, which Pflueger & Rinaldi (2022) use to analyze the role of countercyclical risk aversion in generating the response of stock returns to monetary policy shocks. Swanson (2016) incorporates time-varying risk premia into the New Keynesian framework using Epstein-Zin preferences. In contrast to the aforementioned papers, I use the government debt valuation equation to determine the price level, which obviates the need to introduce nominal rigidities in the manner of Rotemberg (1982) or Calvo (1983).

The real side of my model builds upon an older literature that incorporates risk premia into RBC models with capital adjustment costs. Nearly all of these models—

such as Jermann (1998), Kaltenbrunner & Lochstoer (2010), or Chen (2017)—exclude labor as an input to the production function and feature a single shock to productivity. More recently, Bai & Zhang (2022) embed a DMP model into the RBC framework and use Epstein-Zin preferences to match the equity premium with a low level of risk aversion. While I follow Bai & Zhang (2022) by including both capital and labor in the production function, I deviate from their model by deploying the two-agent setup of Greenwald et al. (2019), in which the SDF features a time-varying price of risk and a time-varying time discount factor.

Cochrane (2023) summarizes decades of research on the Fiscal Theory of the Price Level (FTPL) and traces the impact of adding the government debt valuation equation to an otherwise standard New Keynesian model. Caramp & Feilich (2022) show that wealth effects weaken the impact of monetary policy shocks when fiscal policy is non-Ricardian, and Corhay et al. (2023) show that time-varying risk premia allow changes in the maturity structure of government debt to affect expected inflation. To my knowledge, this paper is the first to integrate the FTPL into a quantitatively realistic RBC model. As in Jiang et al. (2019), the cyclical dynamics of tax revenues and government spending in the model generate a risk premium on government debt.

Unlike the voluminous literature on monetary policy shocks, very little research has been devoted to the identification and estimation of risk premium shocks. Adrian et al. (2023) develop an index of financial conditions that closely tracks the market price of risk. The authors show that a tightening of the index induces a persistent contraction in output with no response of inflation. Meeuwis et al. (2023) examine the response of unemployment to discount rate news, defined as revisions in rational expectations of future stock returns. Instead of following the VAR-based approach of Campbell (1991), the authors measure discount rate news directly using forecasting regressions. To infer structural shocks to risk premia, Cieslak & Pang (2021) estimate a VAR with sign and monotonicity restrictions using high-frequency data on asset prices. In Section 2, I use all three approaches to estimate empirical impulse responses of macroeconomic variables to risk premium shocks.

Basu et al. (2021) and Li & Merkel (2020) are the most similar papers to this one, so it's worth highlighting the differences among them. Empirically, Basu et al. (2021) identify a single shock that accounts for the bulk of fluctuations in five-year-ahead expected excess stock returns. The authors show that the shock generates the business-cycle comovement of macroeconomic quantities and inflation. The authors develop an RBC model with Epstein-Zin preferences to rationalize this finding, although their model does not attempt to explain the response of inflation. Li & Merkel (2020) build a New Keynesian model with idiosyncratic risk and incomplete markets to study the transmission of uncertainty shocks. As in this paper, government bonds are in positive net supply, and fiscal policy can be used to ameliorate aggregate demand recessions induced by precautionary savings. However, uncertainty shocks produce substantial disinflation in their model, and variation in idiosyncratic risk does not necessarily generate variation in aggregate risk premia, the object of this paper.

2 Empirical Evidence

This section documents the risk price puzzle, in which risk premium shocks have negative impacts on output, consumption, investment, and employment, and a negligible and/or positive impact on inflation. As there is no widely accepted measure of risk premium shocks, I infer shocks using three state-of-the-art methods: (i) the Volatility Financial Conditions Index of Adrian et al. (2023); (ii) the news decomposition of Meeuwis et al. (2023); and (iii) the structural VAR of Cieslak & Pang (2021). Each of these methods produce a long time series of risk premium shocks dating back to the Great Inflation of the 1970s.

2.1 Data Sources

Quarterly macroeconomic data are obtained from the Federal Reserve Economic Database (FRED). Output, consumption, and investment are reported in real, per capita terms. Consumption includes nondurable goods and services, while investment includes durable goods and gross private domestic investment. Monthly unemployment rates are averaged to a quarterly frequency. Inflation is measured as the annualized log difference in the GDP deflator. The primary surplus is the Bureau of Economic Analysis (BEA)'s measure of net federal government saving plus interest payments. Utilization-adjusted total factor productivity is obtained from the Federal Reserve Bank of San Francisco. All variables are seasonally adjusted. When calculating impulse responses, I use the cyclical components of output, consumption, and investment extracted using the filtering procedure of Hamilton (2018); I take logs of these variables before applying the filter in order to make the resulting series scale invariant.

Stock returns are inferred from the cum-dividend CRSP value-weighted index. The price-dividend ratio is calculated as the end-of-quarter market value of the index divided by the sum of dividends paid over the previous four quarters. The nominal risk-free rate is the end-of-quarter 3-month Treasury yield from the CRSP Treasury Risk-Free Rate file. A real risk-free rate is constructed by subtracting expected inflation from the nominal rate, where expected inflation is measured using the median one-quarter-ahead forecast of the GDP deflator from the Survey of Professional Forecasters (SPF). End-of-quarter yields on longer-term Treasuries and Moody's corporate bond indices are obtained from FRED. Returns on the Fama & French (2015) factors are obtained from Ken French's website. Stock returns and interest rates are annualized and continuously compounded.

2.2 Volatility Financial Conditions Index

For a representative agent with time separable preferences, Breeden (1979) shows that the market price of risk—i.e., the volatility of the pricing kernel—can be measured as the volatility of aggregate consumption.[3] Motivated by this theoretical result, Adrian et al. (2023) develop the Volatility Financial Conditions Index (VFCI), defined as the conditional volatility of output spanned by financial factors. The VFCI has two main advantages over more commonly used equity-implied volatility indices like the VIX. First, since the VFCI does not rely on options data, it is available over a longer sample beginning in the mid-1960s, whereas the VIX is only available beginning in 1990. Second, given that the VIX is derived from the option prices of publicly traded stocks, it is an imperfect measure of the true market price of risk, as only a fraction of the overall capital stock is publicly traded.

The construction of the VFCI proceeds in two steps. First, a small number of financial variables are designated as "base assets" that span a sizable fraction of variation in the market price of risk. I follow Adrian et al. (2023) and use the six financial variables in Table 1 as base assets. In order to construct an orthogonal basis for the space spanned by the assets, I compute the first four principal components, which explain approximately 90% of the variation.

Table 1: Financial Variables

Name	Description
ret	Annualized Return on CRSP Value-Weighted Stock Index
vol	Standard Deviation of Daily CRSP Stock Returns
term	10-Year Yield Minus 3-Month Yield on Treasuries
liq	3-Month Treasury Yield Minus Effective Federal Funds Rate
cred	Moody's AAA Corporate Bond Yield Minus 10-Year Treasury Yield
def	Moody's AAA Minus BAA Corporate Bond Yield

Second, I estimate a multiplicative heteroskedastic linear regression of real GDP growth on the principal components using maximum likelihood:

$$\Delta y_{t+1} = \boldsymbol{\theta PC}_t + \varepsilon_t, \tag{1}$$

$$\text{Var}\left(\varepsilon_t^2\right) = \sigma_t^2 = \exp\left(\boldsymbol{\delta PC}_t\right), \tag{2}$$

where Δy_{t+1} is log real GDP growth and $\boldsymbol{PC}_t \equiv [1, PC_1, PC_2, PC_3, PC_4]$ is the vector of principal components. Equation (1) states that expected GDP growth is a linear function of the principal components, while Equation (2) guarantees that volatility

[3]When preferences are not time separable, the volatility of aggregate consumption is a forward-looking measure of current and future expected prices of risk.

is non-negative. The VFCI is constructed as the fitted values of the logarithm of the conditional volatility of GDP growth:

$$VFCI \equiv \log \sqrt{\hat{\sigma}_t^2} = \hat{\boldsymbol{\delta}} \boldsymbol{PC}_t. \tag{3}$$

The correlation between the VFCI and the VIX is 0.78 in the post-1990 sample for which both series are available. I measure shocks to the VFCI as the residuals from an estimated AR(1) process, where the optimal number of lags was selected using the Akaike and Bayesian information criteria.

2.3 News Decomposition

Campbell (1991) proposes a method to infer fluctuations in risk premia by decomposing innovations in real stock returns into revisions in rational expectations of future dividends ("cash flow news") and real returns ("discount rate news"). Mathematically, this decomposition can be expressed as:

$$r_t^s - \mathbb{E}_{t-1} r_t^s = \underbrace{(\mathbb{E}_t - \mathbb{E}_{t-1}) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+j}}_{\text{Cash Flow News}} - \underbrace{(\mathbb{E}_t - \mathbb{E}_{t-1}) \sum_{j=1}^{\infty} \rho^j r_{t+j}^s}_{\text{Discount Rate News}}, \tag{4}$$

where r_t^s is the log real stock return, d_t is the log real dividend, and ρ is a linearization parameter.[4] The traditional approach measures the news terms by estimating a first-order VAR and inferring long-horizon expectations from the estimated coefficients. To avoid the restrictions imposed by assuming a time-invariant VAR structure, I follow Meeuwis et al. (2023) and measure the news terms directly using forecasting regressions. In particular, let $\Gamma_t \equiv \sum_{j=1}^{J} \rho^j r_{t+j}^s$ be the discounted sum of future excess stock returns up to some fixed horizon J. One can estimate the loadings, w_k, on the returns of traded factors that best predict Γ_t:

$$\Gamma_t = a + \boldsymbol{b} \boldsymbol{\Omega}_{t-1} + \sum_{k=1}^{N} w_k f_{k,t} + \eta_t, \tag{5}$$

where $\boldsymbol{\Omega}_{t-1}$ is a vector of lagged state variables that includes: (i) the log real stock return, r_{t-1}^s; (ii) the nominal 3-month Treasury yield, y_{t-1}^{3mo}; (iii) the spread between the nominal 10-year and 3-month Treasury yields, s_{t-1}; and (iv) the log price-dividend ratio, pd_{t-1}. Under the assumption that $\boldsymbol{\Omega}_{t-1}$ captures existing prior beliefs about Γ_t, $w_k f_{k,t}$ can be interpreted as discount rate news.

Consistent with Meeuwis et al. (2023), I define the set of traded factors to be the

[4] See Chapter 5.3 of Campbell (2017) for derivation. The linearization parameter, ρ, is equal to $1/\left(1 + \overline{DP}\right)$, where $\overline{DP}$ is the steady-state dividend-price ratio.

five-factor model of Fama & French (2015) plus momentum. I estimate the factor loadings at a quarterly frequency, assuming that stock returns after 40 quarters are not predictable. Armed with the estimated loadings $\widehat{w}_k$, discount rate and cash flow news can be inferred as:

$$N_{DR,t} \equiv \sum_{k=1}^{N} \widehat{w}_k \widetilde{f}_{k,t} \tag{6}$$

$$N_{CF,t} \equiv N_{MKT,t} + N_{DR,t}, \tag{7}$$

where $\widetilde{f}_{k,t}$ and $N_{MKT,t}$ are the traded factors and real stock returns, respectively, orthogonalized with respect to $\mathbf{\Omega}_{t-1}$.

2.4 Structural VAR

One drawback of the news decomposition approach is that correlations between the news terms introduce ambiguity in the interpretation of impulse responses. Ideally, one would have a measure of *structural* risk premium shocks that represent unanticipated movements in the price of risk that are orthogonal to other shocks. Along these lines, Cieslak & Pang (2021) provide a method to infer risk premium shocks using a structural VAR with sign and monotonicity restrictions. The procedure assumes that there are three types of shocks: (i) growth shocks, ω^g; (ii) risk-free rate shocks, ω^m; and (iii) risk premium shocks, ω^p. Growth shocks represent shocks to investor expectations of future cash flows—a positive growth shock raises both stock returns and bond yields, with an impact that weakens with bond maturity. A positive risk-free rate shock depresses stock returns and raises bond yields, with an impact that declines with maturity. Finally, a positive risk premium shock lowers stock returns and raises bond yields, but with an impact that is stronger at longer maturities. Since the orthogonality assumption only holds contemporaneously, these restrictions do not preclude situations where, for example, a negative shock to growth expectations leads to a higher risk premium in subsequent periods. Table 2 summarizes the sign and monotonicity restrictions.

Table 2: Sign & Monotonicity Restrictions

	Growth $\boldsymbol{\omega^g} \uparrow$	**Risk-Free Rate** $\boldsymbol{\omega^m} \uparrow$	**Risk Premium** $\boldsymbol{\omega^p} \uparrow$
Yield Changes	$(+)$	$(+)$	$(+)$
Stock Returns	$(+)$	$(-)$	$(-)$
Impact Maturity	Short	Short	Long

In the terminology of Cieslak & Pang (2021), ω^p is a "common premium" shock that simultaneously increases the risk premium on bonds and stocks. Alternatively, one could assume that a risk premium shock lowers bond yields. However, as shown by Lettau & Wachter (2011), in order for the yield curve to be unconditionally upward-sloping, a positive risk premium shock must raise the yields of both nominal and real bonds, with a greater impact at longer maturities. Given the strong evidence that the average slope of the yield curve is positive,[5] I prefer the sign restrictions in Table 2.

In order to extract the structural shocks, I first estimate a VAR(1) using quarterly innovations in the log price-dividend ratio and the nominal yields of 3-year and 10-year Treasuries.[6] Starting from the Cholesky decomposition of the variance-covariance matrix of the reduced-form residuals, I generate rotation matrices using the QR factorization approach of Rubio-Ramirez et al. (2010) and impose the sign and monotonicity restrictions. Since this procedure only set-identifies the impulse responses, I follow Fry & Pagan (2011) and select the shocks that produce impulse responses that are as close to the median as possible.

2.5 Impulse Responses

In order to quantify the impact of risk premium shocks on macroeconomic variables, I estimate local projection impulse responses in the manner of Jordà (2005). In particular, I run regressions of the form:

$$y_{t+h} = \alpha + \beta^h v_t^{rp} + \mathbf{\Gamma}'\mathbf{Z}_t + \varepsilon_{t+h}, \tag{8}$$

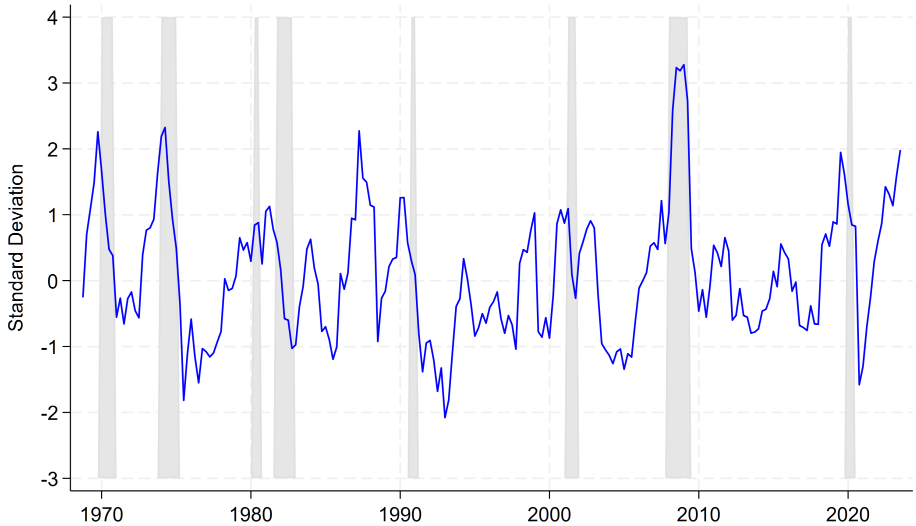
where v_t^{rp} is the risk premium shock in quarter t and y_{t+h} is the variable of interest in quarter $t+h$. The vector of controls $\boldsymbol{Z}_t$ includes four lags of output, consumption, investment, unemployment, inflation, and the real stock return. Impulse responses are computed for horizons $h = \{0, \ldots, 20\}$ quarters, and standard errors are Newey-West with five lags. As a baseline, I define v_t^{rp} as the first principal component of the VFCI shock, discount rate news, and structural risk premium shock. Figure 2 plots the time series of v_t^{rp}. Increases in risk premia occur either immediately before or at the start of recessions, with the largest shocks occurring during the 1973 oil embargo, the "Black Monday" crash of 1987, and the Global Financial Crisis.

As noted in the Introduction, Figure 1 plots the empirical local-projection impulse

[5]Using the Gürkaynak et al. (2007) estimates of zero-coupon yields maintained by the Federal Reserve, I calculate the average spread between 10-year and 2-year yields to be 1.22% using nominal Treasuries and 1.18% using Treasury inflation-protected securities (TIPS). Nominal yields are available beginning in 1971, whereas TIPs yields are available since 1999.

[6]Since I estimate the VAR at the quarterly frequency, I use changes in the log price-dividend ratio instead of stock returns. Using a Campbell-Shiller linearization, innovations in stock returns can be expressed as: $\Delta r_{t+1}^s - \mathbb{E}_t \Delta r_{t+1}^s \approx \kappa_1 \left(pd_{t+1} - \mathbb{E}_t pd_{t+1}\right) + \left(d_{t+1} - \mathbb{E}_t d_{t+1}\right)$. The first term, $pd_{t+1} - \mathbb{E}_t pd_{t+1}$, captures shocks to the relevant state variables, while the second term, $d_{t+1} - \mathbb{E}_t d_{t+1}$, captures shocks to the current realization of log dividends. The noise arising from the second term can be substantial at the quarterly frequency.

Figure 2: Time Series of Risk Premium Shocks



Note: The figure above plots the time series of risk premium shocks, defined as the first principal component of the VFCI shock, discount rate news, and structural risk premium shock. Shaded bars represent NBER recessions.

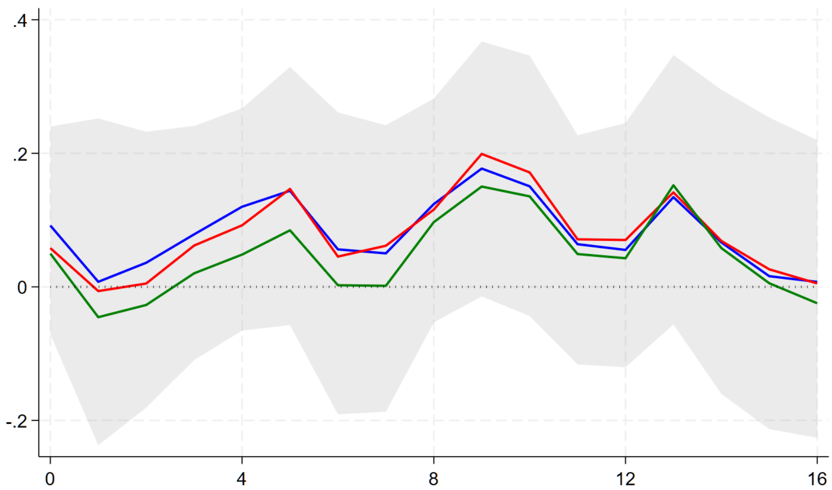
responses to a one-standard deviation increase in v_t^{rp}. The maximum impact of a risk premium shock on macroeconomic quantities is obtained at the six-quarter horizon: a positive risk premium shock is associated with a 0.63 percentage-point decrease in output, a 0.47 percentage-point decrease in consumption, a 2.13 percentage-point decrease in investment, and a 0.39 percentage-point increase in the unemployment rate. All of these effects are statistically significant at the 5% level. Inflation increases by approximately 0.1 percentage points in the quarter after the shock, but the response is statistically insignificant at all horizons. The annualized real stock return falls by about 25 percentage points on impact, followed by three quarters of higher than average returns, which reflect elevated expected excess returns. For robustness, Appendix B reports the empirical impulse responses to VFCI shocks, discount rate news, and structural risk premium shocks separately. All three measures of risk premium shocks produce a similar pattern of impulse responses for macroeconomic quantities: output, consumption, and investment fall, and the unemployment rate increases. Consistent with the risk price puzzle, no measure generates a decline in inflation, and a positive structural risk premium shock actually leads to a statistically significant increase in inflation in the two quarters after the shock.

Research dating back to Sims (1992) seeks to resolve the price puzzle by controlling for measures of commodity prices when computing impulse responses to monetary policy shocks: since commodity prices predict future inflation, they plausibly belong in the information set of the central bank when setting interest rates.[7] Does the risk

[7]Christiano et al. (1994) also point out that including commodity prices helps control for the confounding effects of oil price supply shocks.

price puzzle similarly disappear when controlling for measures of expected inflation? Figure 3 compares the empirical impulse response of inflation in the baseline specification to those obtained when adding commodity prices or SPF inflation forecasts as controls.[8] The impulse responses are essentially unchanged.

Figure 3: Empirical Impulse Response of Inflation to Risk Premium Shock



Note: The figure above plots the empirical local-projection impulse responses of inflation to a one-standard deviation risk premium shock in the baseline specification (blue), controlling for SPF inflation forecasts (red), and controlling for commodity prices (green). The y-axis measures the response in percentage points, while the x-axis represents the number of quarters after the shock. Shaded areas are the maximum 90% confidence intervals across the three specifications.

Finally, Figure 4 plots the empirical impulse responses of two auxiliary variables—the log surplus-output ratio and total factor productivity (TFP). Since government surpluses are frequently negative, I calculate the log surplus-output ratio as $sy_t = \log(1 + S_t/Y_t)$.[9] A positive risk premium shock has a small yet statistically significant impact on the surplus-output ratio in the short-to-medium term; combined with the fact that output falls in response to a positive risk premium shock, there is strong evidence that risk premium shocks affect surpluses. If anything, TFP appears to increase in response to a positive risk premium shock, indicating that the risk price puzzle is not an artifact of the confounding effects of a negative demand shock and a negative supply shock, insofar as supply shocks affect TFP.

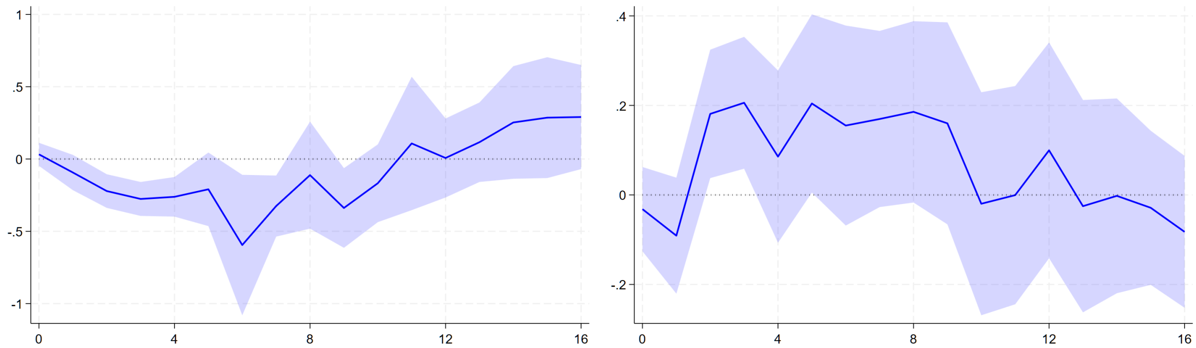
2.6 Stock-Bond Correlation

The changing correlation between the returns on stocks and government bonds has been extensively documented, with a consensus that the correlation switched from

[8] In order to maximize the length of the sample, I use the Commodity Research Bureau (CRB)'s raw industrials index, which is available beginning in 1947Q1. Hanson (2004) shows that the raw industrials index has superior forecasting power for future inflation than the more commonly used spot index.

[9] Given that the minimum and maximum surplus-output ratios in my sample are -0.27 and 0.055, respectively, the $\log(1 + x)$ approximation is fairly accurate.

Figure 4: Empirical Impulse Responses of Auxiliary Variables to Risk Premium Shock



(a) Surplus-Output Ratio

(b) TFP

Note: The figures above plot the empirical local-projection impulse responses of the log surplus-output ratio and TFP to a one-standard deviation risk premium shock. The y-axis measures the response in percentage points, while the x-axis represents the number of quarters after the shock. Shaded areas are 90% confidence intervals.

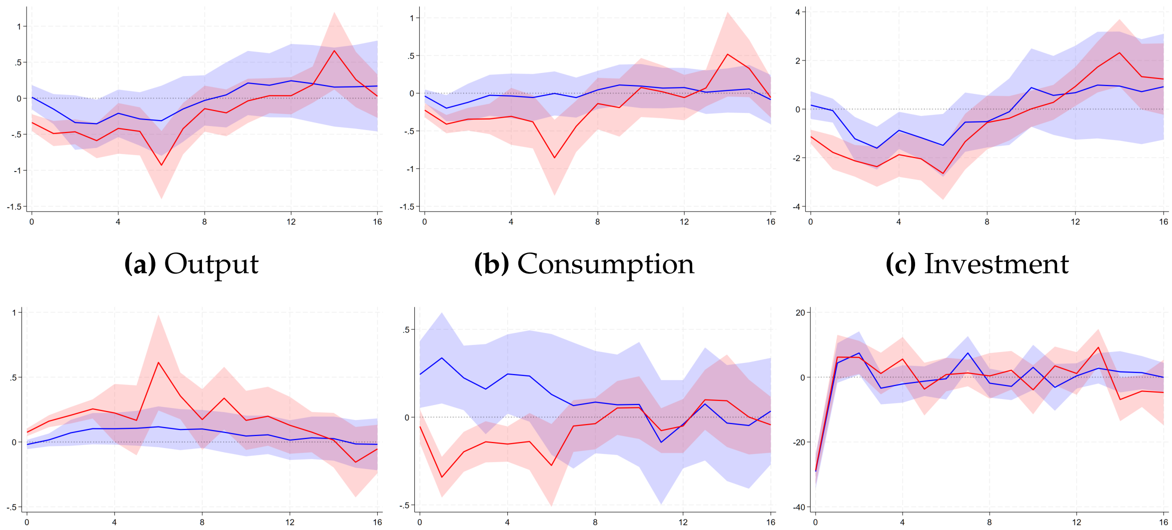
positive to negative around 1998 in the United States.[10] Is there a similar change in the relationship between risk premium shocks and inflation at this time? Figure 5 compares the empirical local projection impulse responses in the pre-1998 and post-1998 subsamples. Positive risk premium shocks generate an increase in inflation in the pre-1998 subsample, with attenuated effects on macroeconomic quantities. However, in the post-1998 subsample, positive risk premium shocks lower inflation, with stronger effects on macroeconomic quantities. In Section 5, I rationalize this divergence in impulse responses via the changing correlation between shocks to the risk premium and real risk-free rate. In particular, in the pre-1998 subsample, risk premium shocks were positively correlated with risk-free rate shocks, inducing a positive correlation between stock and bond returns. Since both a higher risk premium and a higher risk-free rate raise the discount rate on government debt, inflation will increase from the government debt valuation equation. On the other hand, in the post-1998 subsample, risk premium shocks were negatively correlated with risk-free rate shocks, inducing a negative correlation between stock and bond returns. In this case, the effect of a higher risk premium is offset by a lower risk-free rate, which can lead to disinflation.

3 New Keynesian

In this section, I examine whether existing New Keynesian models can rationalize the risk price puzzle. In nearly all of these models, macroeconomic dynamics can be

[10]See Campbell et al. (2020) and Chernov et al. (2021). Laarits (2020) shows that the stock-bond correlation calculated using real bond prices is similar to that calculated using nominal bond prices.

Figure 5: Empirical Impulse Responses to Risk Premium Shock: Pre- and Post-1998



(a) Output **(b)** Consumption **(c)** Investment

(d) Unemployment **(e)** Inflation **(f)** Real Stock Return

Note: The figures above plot the empirical local-projection impulse responses of output, consumption, investment, unemployment, inflation, and the real stock return to a one-standard deviation risk premium shock in pre-1998 (blue) and post-1998 (red) periods. The y-axis measures the response in percentage points, while the x-axis represents the number of quarters after the shock. Shaded areas are 90% confidence intervals. The sample is from 1968Q4 to 2023Q2.

summarized in the following three equations:

$$\textbf{Euler Equation: } x_t = \mathbb{E}_t x_{t+1} - \psi\left(i_t - \mathbb{E}_t \pi_{t+1}\right) + v_{xt} \tag{9}$$

$$\textbf{Phillips Curve: } \pi_t = \kappa x_t + \rho \mathbb{E}_t \pi_{t+1} + v_{\pi t} \tag{10}$$

$$\textbf{Monetary Policy Rule: } i_t = \gamma^x x_t + \gamma^\pi \pi_t + v_{it}, \tag{11}$$

where x_t is the output gap, i_t is the nominal risk-free rate, and π_t is inflation. The demand shock, v_{xt}, in Equation (9) represents anything that affects the output gap at a given risk-free rate. As such, v_{xt} subsumes shocks to the risk premium arising from a time-varying price of risk. For instance, with external habit preferences as in Campbell et al. (2020), demand shocks can arise from shocks to the surplus consumption ratio. The slope of the Phillips curve, κ, captures the sensitivity of inflation to the output gap. For $\kappa > 0$, a negative demand shock will lead to lower inflation, all else equal.

To illustrate the effects of a risk premium shock in the New Keynesian model, I compute the model-implied impulse responses of output, consumption, investment, unemployment, and inflation to a one-standard deviation risk premium shock in the benchmark DSGE model of Galí et al. (2012).[11] The authors estimate their model on U.S. data from 1966Q1 to 2007Q4 using Bayesian full-system estimation techniques. While the authors refer to their demand shock as a "risk premium shock", they do not

[11] Galí et al. (2012) reformulates the medium-scale DSGE model of Smets & Wouters (2007) to allow for involuntary unemployment.

provide a structural interpretation of the shock in terms of preferences or attempt to match asset pricing moments. Nevertheless, the basic mechanism—a positive risk premium shock lowers consumption and the output gap from the Euler equation, which in turn lowers inflation by moving along an upward-sloping the Phillips curve—carries over into finance-integrated New Keynesian models like Swanson (2016) or Pflueger & Rinaldi (2022).

Figure 6 shows that a positive risk premium shock in Galí et al. (2012) generates sizable reductions in output, consumption, employment, and investment, although it fails to capture the lags in these responses. Interestingly, the magnitude of the investment response is smaller than that of output or consumption, which is counterfactual given the empirical evidence in Section 2. Panel (e) shows that inflation declines by 0.05 percentage points on impact, with an effect that lasts about 10 quarters. The weak response of inflation reflects the model's flat structural Phillips curve. Since estimates of the slope of the Phillips curve vary greatly—Schorfheide (2008) gives a range between 0 and 4—the response of inflation to a risk premium shock in Galí et al. (2012) can be considered an effective lower bound.[12] Furthermore, the model-implied impulse response of the real stock return in Panel (f) is 15 times smaller than the corresponding empirical impulse response. If one calibrates the volatility of the risk premium shock to match the response of the real stock return, inflation declines by 0.2 percent instead, as indicated by the dashed green line in Panel (e).

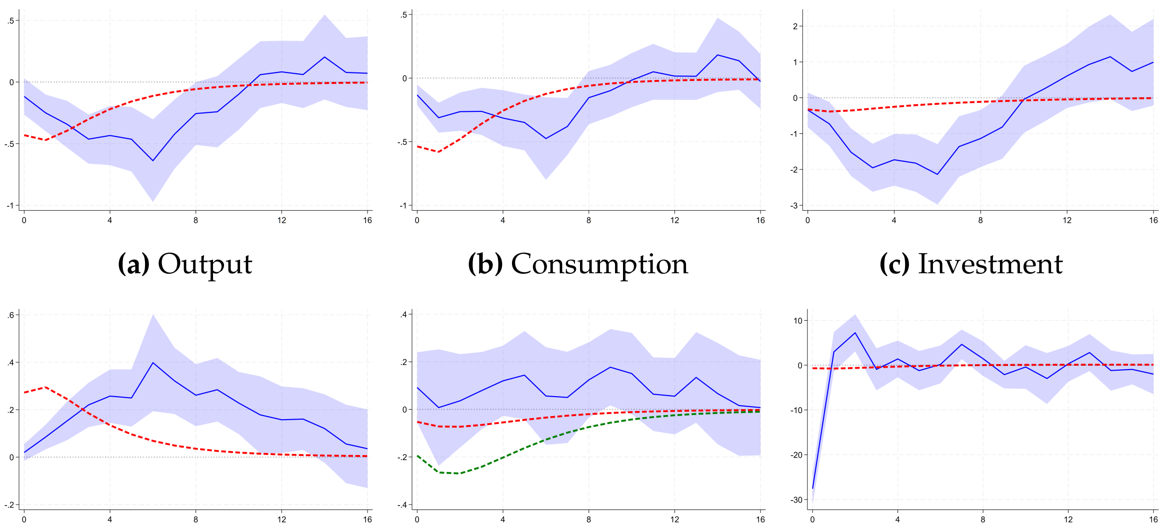
What about monetary policy? McLeay & Tenreyro (2020) argue that the empirical disconnect between inflation and aggregate demand can be rationalized by highly effective monetary policy, even if the structural Phillips curve is steep. Intuitively, if monetary policy seeks to minimize welfare losses—conventionally measured as deviations of inflation from its target and output from its potential—then the central bank will seek to increase inflation when the output gap is negative. Doing so will induce a negative correlation between inflation and the output gap in the data. However, this argument does not explain the empirical impulse responses in Section 2, which require either a flat structural Phillips curve or an alternative model in which risk premium shocks generate the procyclical movement of macroeconomic quantities even when monetary policy replicates flexible-price allocations.

4 Model

This section develops a novel macro-finance model that rationalizes the risk price puzzle. The real side is a two-sector RBC model with capital adjustment costs and equilib-

[12] As an illustration of just how fickle the empirical Phillips curve has proven to be, St. Louis Fed President James Bullard delivered remarks in 2018 entitled "The Case of the Disappearing Phillips Curve". A few years later, researchers at the Chicago Fed published a report with new estimates of the slope of the Phillips curve (Hobijn et al. (2023)). The title? "Recent Steepening of Phillips Curves".

Figure 6: Model-Implied Impulse Responses to Risk Premium Shock in a Benchmark New Keynesian Model



(a) Output **(b)** Consumption **(c)** Investment

(d) Unemployment **(e)** Inflation **(f)** Real Stock Return

Note: The figures above plot the empirical (solid blue) and model-implied (dashed red) impulse responses of output, consumption, investment, unemployment, inflation, and the real stock return to a positive one-standard deviation risk premium shock in Galí et al. (2012). The dashed green line in Panel (e) shows the model-implied response of inflation when the volatility of the risk premium shock is calibrated to match the empirical response of the real stock return. The y-axis measures the response in percentage points, while the x-axis represents the number of quarters after the shock. Shaded areas are 90% confidence intervals.

rium unemployment. The two-agent "shareholder-worker" setup decouples the labor supply from consumption-savings decisions. Risk premium shocks affect inflation through the government debt valuation equation, which links the price level to the market value of government debt.

4.1 Preferences

The economy is populated by two types of agents. The first type are shareholders who own equity in a representative firm and debt issued by the government. The second type are hand-to-mouth workers who finance consumption out of labor income and government transfers. A representative shareholder has time-separable power utility:

$$U\left(C_t^s\right) = \frac{C_t^{s\,1-x_{t-1}} - 1}{1 - x_{t-1}}, \tag{12}$$

where C_t^s is real shareholder consumption in period t. Marginal utility is $U'\left(C_t^s\right) = C_t^{s\,-x_{t-1}}$. From Equation (12), assets are priced by a stochastic discount factor of the form:

$$M_{t+1} = \hat{\beta}_t \left(\frac{C_{t+1}^s}{C_t^s}\right)^{-x_t}. \tag{13}$$

(See Appendix A.1 for derivation.) Equation (13) contains three innovations relative to the standard SDF obtained with constant relative risk aversion: (i) the price of risk, x_t, is time-varying; (ii) the (scaled) time discount factor, $\hat{\beta}_t$, is time-varying; and (iii) shareholder consumption appears in the SDF, as workers do not own any financial assets. I assume that x_t follows an exogenous AR(1) process:

$$x_t = \overline{x} + \phi_x x_{t-1} + \varepsilon_{xt}, \quad \varepsilon_{xt} \sim N\left(0, \sigma_x^2\right). \tag{14}$$

Since an SDF reflects both preferences and beliefs, an increase in x_t can be interpreted as either an increase in effective risk aversion or greater pessimism about future shareholder consumption. As in Greenwald et al. (2014), I specify that the time discount factor takes the form:

$$\hat{\beta}_t = \frac{\exp(-\delta_t)}{\mathbb{E}_t\left[\left(\frac{C^s_{t+1}}{C^s_t}\right)^{-x_t}\right]}. \tag{15}$$

Equation (15) implies that the log real risk-free rate exactly follows the exogenous stochastic process δ_t, which is also an AR(1):

$$\delta_t = \overline{\delta} + \phi_\delta \delta_{t-1} + \varepsilon_{\delta t}, \quad \varepsilon_{\delta t} \sim N\left(0, \sigma_\delta^2\right). \tag{16}$$

One can interpret the SDF in Equation (13) as a generalization of the external habit model of Campbell & Cochrane (1999). Instead of specifying an AR(1) process for the log surplus consumption ratio, Equation (14) directly specifies an AR(1) process for the price of risk. Additionally, the functional form of $\hat{\beta}_t$ in Equation (15) is similar in spirit to the habit sensitivity function in Campbell & Cochrane (1999) that is used to reverse engineer a constant real risk-free rate.

4.2 Production

A representative firm uses capital, K_t, and labor, N_t, to produce output, Y_t, according to a Cobb-Douglas production function:

$$Y_t = A_t K_t^\alpha N_t^{1-\alpha}, \tag{17}$$

where α is the capital share and A_t is TFP. Log TFP is assumed to follow an AR(1) process:

$$a_{t+1} = \bar{a} + \phi_a a_{t-1} + \varepsilon_{a,t+1}, \quad \varepsilon_{a,t+1} \sim N\left(0, \sigma_a^2\right). \tag{18}$$

The firm incurs adjustment costs when investing. The cost of investment, Φ_t, is

$$\Phi_t \equiv \Phi\left(I_t, K_t\right) = \left[a_1 + \frac{a_2}{1-1/\nu}\left(\frac{I_t}{K_t}\right)^{1-1/\nu}\right] K_t, \tag{19}$$

where I_t is investment and $\nu > 0$ is the supply elasticity of capital. I follow Jermann (1998) and set $a_1 = \xi/(1-\nu)$ and $a_2 = \xi^{1/\nu}$ to ensure that there are no adjustment costs in the steady state. Capital accumulates according to

$$K_{t+1} = (1-\xi) K_t + \Phi_t, \tag{20}$$

where ξ is the constant depreciation rate.[13]

The representative firm posts a number of job vacancies, V_t, in order to attract unemployed workers, U_t. Since population is normalized to unity, N_t and U_t are the employment and unemployment rates, respectively. As in Hagedorn & Manovskii (2008), vacancies are filled via a constant returns to scale matching function:

$$G(U_t, V_t) = \frac{U_t V_t}{\left(U_t^{\iota} + V_t^{\iota}\right)^{\frac{1}{\iota}}}, \tag{21}$$

where $\iota > 0$ is a curvature parameter that governs the degree of matching frictions. Accordingly, employment evolves as

$$N_{t+1} = (1-s) N_t + G(U_t, V_t), \tag{22}$$

where s is the constant separation rate.[14]

Define labor market tightness as $\theta_t = \frac{V_t}{U_t}$. The probability that a vacancy is filled (i.e., the "job filling rate") is

$$q(\theta_t) = \frac{G(U_t, V_t)}{V_t} = \left(1 + \theta_t^{\iota}\right)^{-1/\iota}. \tag{23}$$

Intuitively, when the labor market is tight—due to a high vacancy rate and/or low unemployment rate—it takes longer for firms to fill vacancies. The representative firm incurs a fixed cost κ when posting vacancies. As a result, aggregate investment in hiring is κV_t.

The firm pays no taxes and doesn't issue debt. Thus, the representative firm's dividends are

$$D_t = Y_t - W_t N_t - \kappa V_t - I_t, \tag{24}$$

where W_t is the real wage. Taking W_t, the stochastic discount factor, M_{t+1}, and the job filling rate, $q(\theta_t)$, as given, the firm chooses investment and vacancies in order to

[13] As in Kaltenbrunner & Lochstoer (2010), I model adjustment costs as a reduction to the accumulation of capital, rather than as a subtraction from profits. In other words, I_t can be more accurately described as investment expenditure gross of adjustment costs.

[14] For $V_t > 0$, s is equal to the maximum drop in the employment level. I follow Kilic & Wachter (2018) and operate under this assumption rather than introducing it as another constraint in the firm's problem.

maximize the cum-dividend market value of the firm:

$$E_t \equiv \max_{\{V_{t+\tau}, N_{t+\tau+1}, I_{t+\tau}, K_{t+\tau+1}\}_{\tau=0}^{\infty}} \mathbb{E}_t \left[\sum_{\tau=0}^{\infty} M_{t+\tau} D_{t+\tau} \right], \tag{25}$$

subject to the laws of motion for capital and employment in Equations (20) and (22). From the first-order conditions with respect to I_t and K_{t+1}, one can derive the *investment Euler equation*:

$$1 \equiv \mathbb{E}_t \left[M_{t+1} R^k_{t+1} \right] = \mathbb{E}_t \left[M_{t+1} \left(\frac{\alpha \frac{Y_{t+1}}{K_{t+1}} + \frac{1}{a_2} \left(\frac{I_{t+1}}{K_{t+1}} \right)^{1/\nu} (1 - \xi + a_1) + \frac{1}{\nu - 1} \frac{I_{t+1}}{K_{t+1}}}{\frac{1}{a_2} \left(\frac{I_t}{K_t} \right)^{1/\nu}} \right) \right], \tag{26}$$

where R^k_{t+1} is the gross investment return. (See Appendix A.2 for derivation.) Equation (26) quantifies the tradeoff between the marginal benefit of investment at $t+1$ and the marginal cost of investment at t. The marginal cost of investment is $1/\left(\partial \Phi_t / \partial I_t\right) = \frac{1}{a_2} \left(\frac{I_t}{K_t} \right)^{1/\nu}$, and the marginal benefit of investment includes the marginal product of capital, $\partial Y_{t+1} / \partial K_{t+1} = \alpha Y_{t+1} / K_{t+1}$, and the marginal continuation value net of depreciation, $(1 - \xi) / \left(\partial \Phi_{t+1} / \partial I_{t+1}\right)$. The remaining terms in the numerator capture the marginal impact of an additional unit of capital on the installation technology.

Similarly, using the first-order conditions with respect to V_t and N_{t+1}, one can derive the *hiring Euler equation*:

$$1 \equiv \mathbb{E}_t \left[M_{t+1} R^n_{t+1} \right] = \mathbb{E}_t \left[M_{t+1} \left(\frac{(1 - \alpha) \frac{Y_{t+1}}{N_{t+1}} - W_{t+1} + (1 - s) \frac{\kappa}{q(\theta_{t+1})}}{\frac{\kappa}{q(\theta_t)}} \right) \right], \tag{27}$$

where R^n_{t+1} is the gross hiring return. Equation (27) quantifies the tradeoff between the marginal benefit of hiring at $t+1$ and the marginal cost of hiring at t. The marginal benefit of hiring includes the marginal product of labor, $\partial Y_{t+1} / \partial N_{t+1} = (1 - \alpha) Y_{t+1} / N_{t+1}$, net of the wage, plus the marginal cost of hiring, $\kappa / q\left(\theta_{t+1}\right)$, net of separation.

Since the Cobb-Douglas production function exhibits constant returns to scale, the gross stock return, R^s_{t+1}, is a weighted average of the investment and hiring returns:

$$R^s_{t+1} = w^k_t R^k_{t+1} + \left(1 - w^k_t\right) R^n_{t+1}, \tag{28}$$

where $w^k_t \equiv \mu^k_t K_{t+1} / \left(\mu^k_t K_{t+1} + \mu^n_t N_{t+1} \right)$ is the value weight of the investment return in the stock return. The shadow value of capital, μ^k_t, is equal to the marginal cost of investment, $(1/a_2) \left(I_t / K_t \right)^{(1/\nu)}$, and the shadow value of labor, μ^n_t, is equal to the marginal cost of hiring, $\kappa / q\left(\theta_t\right)$.

4.3 Wages and Consumption

The equilibrium real wage, W_t, is determined endogenously by applying a sharing rule per the outcome of a Nash bargaining process between employed workers and the firm. Define $\eta \in (0,1)$ as the workers' relative bargaining weight, b as the workers' flow value of unemployment, and τ as the income tax rate. Since workers own no financial assets, the canonical Nash-bargained wage equation does not hold in this setting. Instead, Appendix A.4 shows that W_t is given by:

$$W_t = \frac{1}{1-\tau+\tau\eta}\left[\eta\left((1-\alpha)\frac{Y_t}{N_t}+(1-s)\frac{\kappa}{q(\theta_t)}\right)+(1-\eta)\,b\right]. \tag{29}$$

The real wage is increasing in the marginal product of labor, $(1-\alpha)\,Y_t/N_t$, and the marginal cost of hiring net of separation, $(1-s)\kappa/q(\theta_t)$. Intuitively, the more productive workers are and the more costly it is to fill a vacancy, the higher the wage. The higher the bargaining power of the worker, the smaller the impact of the income tax rate on the wage: in the limiting case in which the worker has full bargaining power and $\eta = 1$, the income tax rate does not affect the wage, as the worker receives all of the surplus from the match. Since workers consume their net-of-tax labor income plus unemployment benefits, worker consumption is equal to:

$$C_t^w = N_t W_t - T_t + bU_t = (1-\tau)\,N_t W_t + bU_t. \tag{30}$$

As explained by Barro & King (1984), RBC models typically have counterfactual implications for the comovement of consumption, investment, and employment in response to a non-technology shock. To see why, consider the expression for dividends in Equation (24). Since consumption smoothing makes investment more procyclical than output, absent an increase in the wage, dividends will increase in response to a risk premium shock. Since the representative shareholder consumes the dividends of the firm, shareholder consumption will increase as well. Thus, in order to generate a realistic response of shareholder consumption to a risk premium shock, I model wealth effects by solving for the representative shareholder's optimal consumption rule as a fraction of financial wealth every period.[15] As shown by Carroll (2004), for a constant price of risk, x, time discount factor, β, and return on wealth, R^w, the optimal fraction of wealth consumed is $\psi = \frac{R^w-(R^w\beta)^{1/x}}{R^w}$.[16] For the SDF in Equation (13) with a

[15]Barro & King (1984) entertain the possibility that wealth effects can overcome the negative comovement between investment and consumption. However, the authors conclude that for a representative agent with time-separable preferences and variable labor supply, the wealth effect will never be strong enough to generate a decline in consumption when investment increases. Thus, the two-agent setup is necessary to overcome the Barro & King (1984) curse.

[16]For log utility, $x = 1$ and ψ simplifies to $1-\beta$. This is the case considered by Caballero & Simsek (2022).

time-varying price of risk, time discount factor, and return on wealth, there is no such closed form solution, but the optimal consumption rule can be found using dynamic programming.

4.4 Government

Government spending is divided into discretionary and non-discretionary components. Non-discretionary spending consists of unemployment benefits, bU_t, while discretionary spending is equal to a constant fraction χ of output. The government's real primary surplus, S_t, is

$$S_t = T_t - G_t = \tau W_t N_t - bU_t - \chi Y_t, \tag{31}$$

where T_t is tax revenue and G_t is total government spending. For simplicity, I assume that the government only issues zero-coupon, one-period debt, rolled over every period. In nominal terms, the government's intertemporal budget constraint is

$$B_{t-1} = P_t S_t + Q_t B_t, \tag{32}$$

where B_t is the amount of nominal bonds outstanding at the end of period t, P_t is the price level, and Q_t is the nominal bond price. The government begins with a initial number of bonds, B_{-1}, and issues new bonds every period to finance budget deficits. The Euler equation for the nominal bond implies that

$$Q_t = \mathbb{E}_t \left[M_{t+1} \frac{P_t}{P_{t+1}} \right]. \tag{33}$$

Divide both sides of Equation (32) by P_t and substitute in Equation (33) to get that

$$\frac{B_{t-1}}{P_t} = S_t + \mathbb{E}_t \left[M_{t+1} \frac{P_t}{P_{t+1}} \right] \frac{B_t}{P_t}.$$

Iterating forward and applying the transversality condition, $\lim_{T\to\infty} \mathbb{E}_t \left[M_{t+1} \frac{B_{T-1}}{P_T} \right] = 0$, one obtains the *government debt valuation equation*[17]:

$$\frac{B_{t-1}}{P_t} = \mathbb{E}_t \left[\sum_{j=0}^{\infty} M_{t+j} S_{t+j} \right]. \tag{34}$$

Debt B_{t-1} is predetermined at time t. Since the right-hand side of Equation (34) does not depend on the price level, the price level in the denominator on the left-hand side

[17] Since there is no growth in the steady state and the discount rate on government debt is positive, the transversality condition will hold.

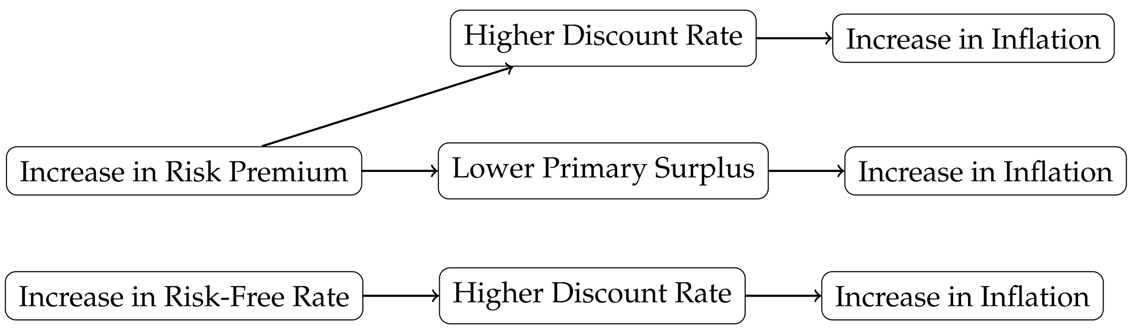
must adjust so that Equation (34) holds. In other words, the price level adjusts so that the real market value of government debt is equal to the present discounted value of future surpluses.[18] Intuitively, if government bonds are worth more than the present value of future surpluses, shareholders have an incentive to sell them and buy more goods and services. This raises aggregate demand and bids up prices.

As pointed out by Jiang et al. (2019), the procyclicality of tax revenues and the countercyclicality of government spending makes the government debt portfolio risky. In other words, the relevant discount rate for government debt contains a risk premium that reflects shareholder compensation for bearing business-cycle risk. To see how risk premium shocks affect inflation through the government debt valuation equation, rewrite Equation (34) using the ex-post return on debt, R^b_{t+1}, as a discount factor:

$$\frac{B_{t-1}}{P_t} = \mathbb{E}_t \left[\sum_{j=0}^{\infty} \left(\prod_{k=1}^{j} \frac{1}{R^b_{t+k}} \right) S_{t+j} \right]. \tag{35}$$

Since the surplus claim is risky, a positive risk premium shock will increase the discount rate on government debt. All else equal, the present discounted value of future surpluses will decline, and the price level will increase. I refer to this as the *discount rate channel* of inflation. The procyclicality of the surplus process implies that there is also a *cash flow channel* of inflation, in which a positive risk premium shock reduces surpluses and pushes up the price level. Finally, shocks to the real risk-free rate can independently affect the discount rate on government debt: an increase in the real risk-free rate raises the discount rate and increases the price level. Figure 7 summarizes the channels in which risk premium and risk-free rate shocks affect inflation through the government debt valuation equation.

Figure 7: Impact of Risk Premium and Risk-Free Rate Shocks on Inflation



[18]The government debt valuation holds even when allowing for sovereign default—government debt is still backed by future surpluses, but bond prices adjust to reflect the possibility of default. See the proof in Appendix A of Jiang et al. (2019).

5 Solution & Mechanism

To solve the model, I compute perturbations around the non-stochastic steady state. A first-order approximation would reduce all risk premia to zero, while a second-order approximation yields nonzero but constant risk premia. Thus, in order to generate time-varying risk premia, the model must be solved to at least the third order. Since approximations higher than the first-order move the ergodic distributions of the endogenous variables away from their non-stochastic steady states, I follow Basu & Bundick (2017) and compute model-implied impulse responses to risk premium shocks around the *stochastic* steady state. Essentially, this compares the path of the economy over an extended period in which all shocks are identically zero to a counterfactual path in which a single one-standard deviation shock to the price of risk, x_t, is realized.

5.1 Estimation Procedure

I calibrate, rather than estimate, the parameters in Table 3. I set the mean, persistence, and volatility of the real risk-free rate to match its empirical dynamics in the sample period from 1968Q4 to 2023Q2. As is standard in the literature, I set the labor share equal to 1/3, while the quarterly depreciation and separation rates of 3.8% and 8.73% are consistent with the values in Bai & Zhang (2022), adjusted to a quarterly frequency. Finally, I set the discretionary spending-to-output ratio equal to its historical average of 0.16 as calculated by Elenev et al. (2022).

Table 3: Calibrated Parameters

Name	Description	Value
α	Labor Share	1/3
ξ	Depreciation Rate	3.8%
s	Separation Rate	8.73%
$\overline{\delta}$	Risk-Free Rate Mean	0.3%
σ_δ	Risk-Free Rate Volatility	0.2%
ρ_δ	Risk-Free Rate Persistence	0.88
χ	Discretionary Spending-to-Output Ratio	0.16

Let $\hat{\mathbf{\Psi}}$ be a column vector that stacks the point estimates of the empirical impulse responses to a risk premium shock across all horizons, along with the unconditional average equity premium. If $\mathbf{\Psi}(\Pi)$ denotes the corresponding model-implied impulse responses and unconditional equity premium, then the impulse response matching procedure minimizes the following objective function:

$$\mathcal{L}(\Pi) \equiv (\hat{\mathbf{\Psi}} - \mathbf{\Psi}(\Pi))' \boldsymbol{W} (\hat{\mathbf{\Psi}} - \mathbf{\Psi}(\Pi)), \tag{36}$$

where W is a diagonal weighting matrix consisting of the inverse of the variances of the empirical impulse responses in $\hat{\mathbf{\Upsilon}}$, along with a large weight for the unconditional equity premium.

Table 4 lists the estimated parameters obtained from the impulse response matching procedure. A few estimates are worth highlighting. First, the average risk price of 20 is low relative to the existing macro-finance literature.[19] Second, the low supply elasticity of capital implies sizable capital adjustment costs that increase the volatility of shareholder consumption. Finally, the flow value of unemployment and worker bargaining weight combine to generate a real wage that is far less inertial than in Bai & Zhang (2022).

Table 4: Estimated Parameters

Name	Description	Value
ρ_a	TFP Persistence	0.99
σ_a	TFP Volatility	0.005
$\overline{x}$	Risk Price Mean	20
ρ_x	Risk Price Persistence	0.85
σ_x	Risk Price Volatility	4.5
ν	Supply Elasticity of Capital	0.145
η	Worker Bargaining Weight	0.79
ι	Curvature of Matching Function	0.65
κ	Vacancy Cost Parameter	0.106
b	Flow Value of Unemployment	0.05
τ	Income Tax Rate	0.33

In order to compute unconditional macroeconomic and asset pricing moments, I simulate the model for 5,000 periods. Of these moments, only the average equity premium was directly targeted in the estimation procedure, so the moments provide an assessment of the external validity of the estimated model. Table 5 compares the model-implied and empirical moments. The average annualized equity premium of 4.87% and Sharpe ratio of 0.43 compare favorably with the data. The risk premium on government debt is 9.54%, which is similar to the 3-month risk premium estimated by Jiang et al. (2019) in their dynamic asset pricing model.[20] Consistent with the "unemployment volatility puzzle" documented by Shimer (2005), the model generates too little volatility in the unemployment rate. However, the model-implied volatilities of output, consumption, and investment growth are closer to their empirical counterparts. Since the model features a volatile price of risk in order to match the equity

[19] For instance, Basu et al. (2021) estimate a steady-state risk aversion of 42 with Epstein-Zin preferences.

[20] Jiang et al. (2019) estimate the risk premium on a short-maturity tax revenue strip to be 9.11%. This is a lower bound on the debt risk premium, though, as the authors show that the risk premium on debt exceeds that on a tax revenue claim.

premium, the model-implied volatility of inflation is higher than in the data, but not excessively so.

Table 5: Unconditional Macroeconomic and Asset Pricing Moments

Moment	Model	Empirical
Equity Risk Premium	4.87%	4.97%
Sharpe Ratio	0.43	0.35
Debt Risk Premium	9.54%	—
Inflation Volatility	4.2%	2.5%
Consumption Growth Volatility	2.68%	2.0%
Unemployment Volatility	0.5%	1.65%
Output Growth Volatility	2.7%	2.18%
Investment Growth Volatility	10.4%	11.1%
Transfer Spending-Output Ratio	2.36%	3.45%
Income Tax Revenue-Output Ratio	19.7%	16.4%

Note: Unconditional model moments are based on a single simulation of 5,000 periods. Empirical moments for the ratios of transfer spending and income taxes to output are from Elenev et al. (2022). All moments are reported in annualized terms.

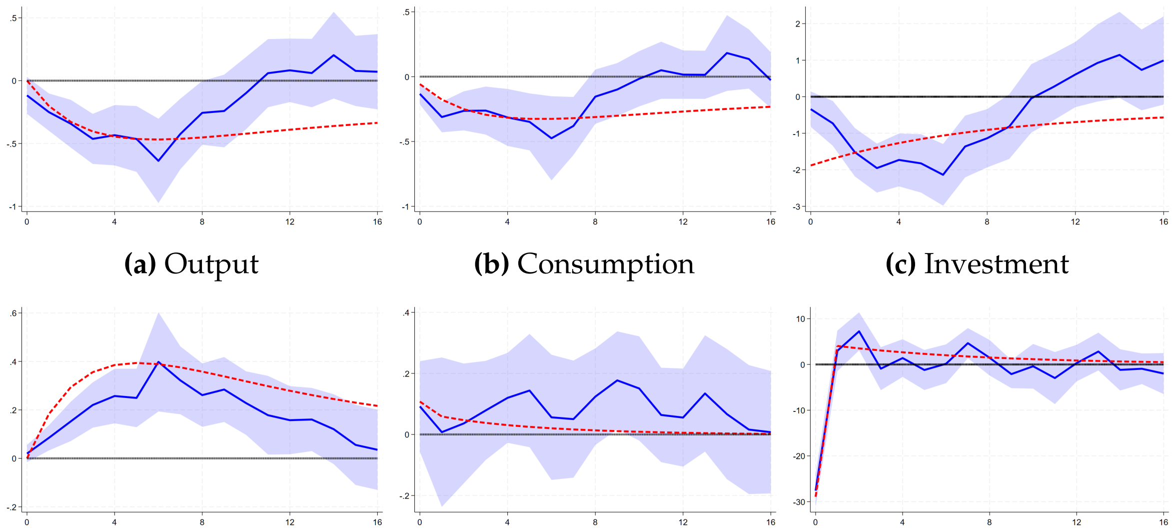
5.2 Model-Implied Impulse Responses

Figure 8 compares the model-implied impulse responses to a one-standard deviation risk premium shock to their empirical counterparts. In general, the model matches the qualitative comovement of output, consumption, investment, and unemployment quite well, although it fails to capture the lagged response of investment.[21] Importantly, the model generates a small yet positive response of inflation in the period after the shock, with an effect that dies out quickly. The transitory response of inflation is a product of the maturity of the debt portfolio; with one-period debt, a risk premium shock generates a one-time price level jump. Unlike Galí et al. (2012), the response of the real stock return is nearly identical in the model and in the data, which reflects the model's ability to match the equity premium.

Recall that the impact of a positive risk premium shock on inflation switched from positive to negative around 1998, coincident with a shift in the correlation between stock and bond returns. Since the price of risk and real risk-free rate follow exogenous stochastic processes in the model, risk premium shocks have no effect on the real risk-free rate. Thus, to capture the changing dynamics of the risk price puzzle, I compute impulse responses to two combinations of shocks: (i) a shock that increases the risk premium and real risk-free rate, generating a positive stock-bond correlation; and (ii) a shock that increases the risk premium and lowers the real risk-free rate, generating

[21] In order to match the delayed response of investment to monetary policy shocks, Christiano et al. (2005) specify that capital adjustment costs are proportional to investment growth, rather than the investment-to-capital ratio as in Equation (19).

Figure 8: Model-Implied Impulse Responses to Risk Premium Shock



(a) Output **(b)** Consumption **(c)** Investment

(d) Unemployment **(e)** Inflation **(f)** Real Stock Return

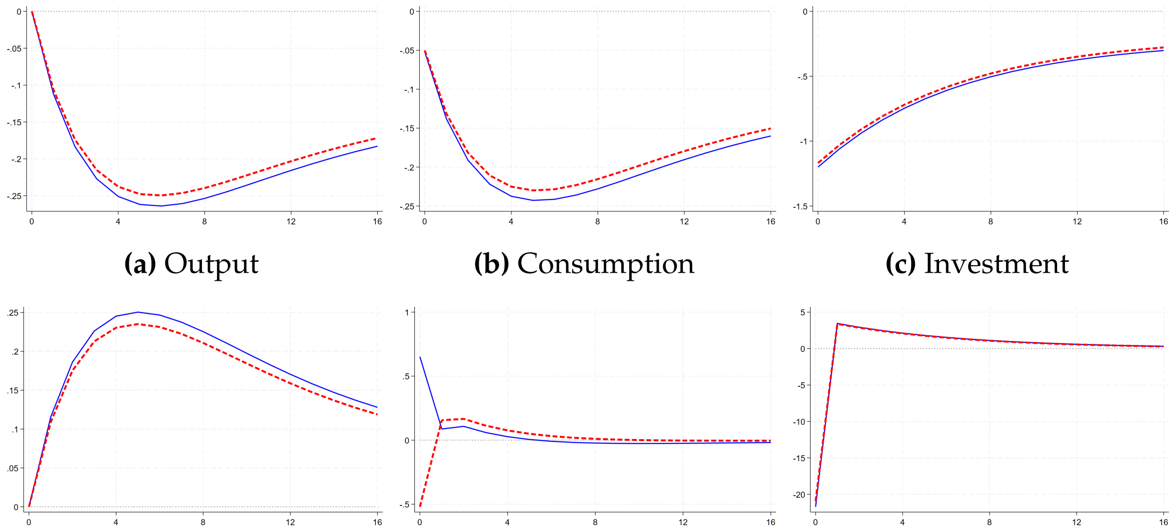
Note: The figures above plot the empirical (solid blue) and model-implied (dashed red) impulse responses of output, consumption, investment, unemployment, inflation, and the real stock return to a positive one-standard deviation risk premium shock. The y-axis measures the response in percentage points, while the x-axis represents the number of quarters after the shock. Shaded areas are 90% confidence intervals.

a negative stock-bond correlation. According to the government debt valuation equation, a simultaneous increase in the risk premium and real risk-free rate magnifies the discount rate channel and generates a larger inflationary response than a risk premium shock alone. On the other hand, a reduction in the real risk-free rate will produce a smaller response of inflation, or even disinflation, depending on the magnitude of the shock.

One interpretation of the simultaneous risk premium and real risk-free rate shocks is the conduct of monetary policy: as documented by Cieslak & Vissing-Jorgensen (2021), in the positive stock-bond correlation era, monetary policy did not respond as aggressively in lowering short-term interest rates in response to declines in the stock market. Alternatively, one could interpret the combination of shocks as capturing the strength of the flight-to-quality phenomenon, in which investors reallocate their portfolios away from risky assets like stocks and toward safe assets like short-term government debt.

Figure 9 plots the model-implied impulse responses for the two different combinations of risk premium and real risk-free rate shocks. The impulse responses to a shock that simultaneously increases the risk premium and real risk-free rate are plotted as solid blue lines. While the responses of macroeconomic quantities are virtually identical to those obtained for a pure risk premium shock in Figure 8, the response of inflation is twice as large, consistent with the empirical impulse responses in the pre-1998 subsample in Figure 5. On the other hand, the model-implied impulse response

Figure 9: Model-Implied Impulse Responses to Risk Premium and Risk-Free Rate Shocks



(a) Output **(b)** Consumption **(c)** Investment

(d) Unemployment **(e)** Inflation **(f)** Real Stock Return

Note: The figures above plot the model-implied impulse responses of output, consumption, investment, unemployment, inflation, and the real stock return to shocks that simultaneously increase the risk premium and real risk-free rate (blue solid) and lower the real risk-free rate (red dashed). The y-axis measures the response in percentage points, while the x-axis represents the number of quarters after the shock.

of inflation to a shock that simultaneously increases the risk premium and decreases the real risk-free rate is negative, consistent with the disinflationary empirical impulse response of inflation in the post-1998 subsample.

6 Conclusion

Although the risk price puzzle is difficult to reconcile in existing New Keynesian models, incorporating nominal rigidities into the RBC model developed in this paper is likely to enhance the realism of the model. In particular, nominal rigidities may help deliver a smoother impulse response of inflation to a risk premium shock even under the assumption of one-period debt. Long-term debt would bring even more realism, given that the average duration of U.S. government debt is approximately six years. Furthermore, the availability of long-run macroeconomic and financial data for advanced countries—e.g., the Jordà-Schularick-Taylor Macrohistory Database—suggests that documenting the risk price puzzle internationally is a natural next step.

While the effects of monetary policy shocks in a heterogeneous-agent economy have been the object of recent study (Kekre & Lenel (2022)), the extent to which risk premium shocks affect the wealth distribution is unknown. Relatedly, while this paper suggests that risk premium shocks have a meaningful impact on macroeconomic aggregates, it is unclear whether similar effects would be observed at the micro level. For instance, once could adapt the framework of Ottonello & Winberry (2020) to investigate the heterogeneous effects of risk premium shocks on investment and employment at the firm level.

Finally, the model in this paper abstracts completely from monetary policy. Given that a simultaneous increase in the risk premium and reduction in the real interest rate can induce disinflation, monetary policymakers may be able to calibrate the magnitude of their response to a risk premium shock in order to eliminate the distortionary effects of inflation. In doing so, one can envision Taylor-type rules that explicitly include measures of risk premia.

References

Adrian, T., Duarte, F., & Iyer, T. (2023). The Market Price of Risk and Macro-Financial Dynamics. *IMF Working Paper*.

Bai, H., & Zhang, L. (2022). Searching for the Equity Premium. *Journal of Financial Economics*, *143*(2), 897–926.

Barro, R. J., & King, R. G. (1984). Time-Separable Preferences and Intertemporal-Substitution Models of Business Cycles. *The Quarterly Journal of Economics*, *99*(4), 817–839.

Basu, S., & Bundick, B. (2017). Uncertainty Shocks in a Model of Effective Demand. *Econometrica*, *85*(3), 937–958.

Basu, S., Candian, G., Chahrour, R., & Valchev, R. (2021). *Risky Business Cycles* (Tech. Rep.). National Bureau of Economic Research.

Breeden, D. T. (1979). An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities. *Journal of Financial Economics*, *7*(3), 265–296.

Caballero, R. J., Farhi, E., & Gourinchas, P.-O. (2017). The Safe Assets Shortage Conundrum. *Journal of Economic Perspectives*, *31*(3), 29–46.

Caballero, R. J., & Simsek, A. (2022). *A Monetary Policy Asset Pricing Model* (Tech. Rep.). National Bureau of Economic Research.

Calvo, G. A. (1983). Staggered Prices in a Utility-Maximizing Framework. *Journal of Monetary Economics*, *12*(3), 383–398.

Campbell, J. Y. (1991). A Variance Decomposition for Stock Returns. *The Economic Journal*, *101*(405), 157–179.

Campbell, J. Y. (2017). *Financial Decisions and Markets: A Course in Asset Pricing*. Princeton University Press.

Campbell, J. Y., & Ammer, J. (1993). What Moves the Stock and Bond Markets? A Variance Decomposition for Long-Term Asset Returns. *The Journal of Finance*, *48*(1), 3–37.

Campbell, J. Y., & Cochrane, J. H. (1999). By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy*, *107*(2), 205–251.

Campbell, J. Y., Pflueger, C., & Viceira, L. M. (2020). Macroeconomic Drivers of Bond and Equity Risks. *Journal of Political Economy*, *128*(8), 3148–3185.

Caramp, N., & Feilich, E. (2022). Monetary Policy and Government Debt. *Unpublished Working Paper*.

Carroll, C. D. (2004). *Theoretical Foundations of Buffer Stock Saving*. National Bureau of Economic Research Cambridge, Mass., USA.

Chen, A. Y. (2017). External Habit in a Production Economy: A Model of Asset Prices and Consumption Volatility Risk. *The Review of Financial Studies*, *30*(8), 2890–2932.

Chernov, M., Lochstoer, L. A., & Song, D. (2021). *The Real Channel for Nominal Bond-Stock Puzzles* (Tech. Rep.). National Bureau of Economic Research.

Christiano, L., Eichenbaum, M., & Evans, C. (1994). *Identification and the Effects of Monetary Policy Shocks* (Tech. Rep.). Federal Reserve Bank of Chicago.

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

Cieslak, A., & Pang, H. (2021). Common Shocks in Stocks and Bonds. *Journal of Financial Economics*, *142*(2), 880–904.

Cieslak, A., & Vissing-Jorgensen, A. (2021). The Economics of the Fed Put. *The Review of Financial Studies*, *34*(9), 4045–4089.

Cochrane, J. H. (2023). *The Fiscal Theory of the Price Level*. Princeton University Press.

Corhay, A., Kind, T., Kung, H., & Morales, G. (2023). Discount Rates, Debt Maturity, and the Fiscal Theory. *The Journal of Finance*, *78*(6), 3561-3620.

Dew-Becker, I. (2014). Bond Pricing with a Time-Varying Price of Risk in an Estimated Medium-Scale Bayesian DSGE Model. *Journal of Money, Credit and Banking*, *46*(5), 837–888.

Elenev, V., Landvoigt, T., Shultz, P., & Van Nieuwerburgh, S. (2022). Can Monetary Policy Create Fiscal Capacity? *Unpublished Working Paper*.

Fama, E. F., & French, K. R. (2015). A Five-Factor Asset Pricing Model. *Journal of Financial Economics*, *116*(1), 1–22.

Fischer, S., & Merton, R. C. (1984). *Macroeconomics and Finance: The Role of the Stock Market* (Tech. Rep.). National Bureau of Economic Research.

Fry, R., & Pagan, A. (2011). Sign Restrictions in Structural Vector Autoregressions: A Critical Review. *Journal of Economic Literature*, *49*(4), 938–960.

Galí, J., Smets, F., & Wouters, R. (2012). Unemployment in an Estimated New Keynesian Model. *NBER Macroeconomics Annual*, *26*(1), 329–360.

Greenwald, D. L., Lettau, M., & Ludvigson, S. C. (2014). *Origins of Stock Market Fluctuations* (Tech. Rep.). National Bureau of Economic Research.

Greenwald, D. L., Lettau, M., & Ludvigson, S. C. (2019). *How the Wealth Was Won: Factors Shares as Market Fundamentals* (Tech. Rep.). National Bureau of Economic Research.

Gürkaynak, R. S., Sack, B., & Wright, J. H. (2007). The US Treasury Yield Curve: 1961 to the Present. *Journal of Monetary Economics*, *54*(8), 2291–2304.

Hagedorn, M., & Manovskii, I. (2008). The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited. *American Economic Review*, *98*(4), 1692–1706.

Hamilton, J. D. (2018). Why You Should Never Use the Hodrick-Prescott Filter. *Review*

of Economics and Statistics, *100*(5), 831–843.
Hanson, M. S. (2004). The Price Puzzle Reconsidered. *Journal of Monetary Economics*, *51*(7), 1385–1413.
Hobijn, B., Miles, R., Royal, J., & Zhang, J. (2023). The Recent Steepening of Phillips Curves. *Chicago Fed Letter*(475).
Jermann, U. J. (1998). Asset Pricing in Production Economies. *Journal of Monetary Economics*, *41*(2), 257–275.
Jiang, Z., Lustig, H., Van Nieuwerburgh, S., & Xiaolan, M. Z. (2019). *The US Public Debt Valuation Puzzle* (Tech. Rep.). National Bureau of Economic Research.
Jordà, Ò. (2005). Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review*, *95*(1), 161–182.
Kaltenbrunner, G., & Lochstoer, L. A. (2010). Long-Run Risk Through Consumption Smoothing. *The Review of Financial Studies*, *23*(8), 3190–3224.
Kekre, R., & Lenel, M. (2022). Monetary Policy, Redistribution, and Risk Premia. *Econometrica*, *90*(5), 2249–2282.
Kilic, M., & Wachter, J. A. (2018). Risk, Unemployment, and the Stock Market: A Rare-Event-Based Explanation of Labor Market Volatility. *The Review of Financial Studies*, *31*(12), 4762–4814.
Laarits, T. (2020). Precautionary Savings and the Stock-Bond Covariance. *NYU Stern School of Business*.
Lettau, M., & Wachter, J. A. (2011). The Term Structures of Equity and Interest Rates. *Journal of Financial Economics*, *101*(1), 90–113.
Li, Z., & Merkel, S. (2020). Flight-to-Safety in a New Keynesian Model. *Unpublished Working Paper*.
McLeay, M., & Tenreyro, S. (2020). Optimal Inflation and the Identification of the Phillips Curve. *NBER Macroeconomics Annual*, *34*(1), 199–255.
Meeuwis, M., Papanikolaou, D., Rothbaum, J., & Schmidt, L. (2023). Discount Rates, Labor Market Dynamics, and Income Risk. *Unpublished Working Paper*.
Ottonello, P., & Winberry, T. (2020). Financial Heterogeneity and the Investment Channel of Monetary Policy. *Econometrica*, *88*(6), 2473–2502.
Pflueger, C. (2023). *Back to the 1980s or Not? The Drivers of Inflation and Real Risks in Treasury Bonds* (Tech. Rep.). National Bureau of Economic Research.
Pflueger, C., & Rinaldi, G. (2022). Why Does the Fed Move Markets So Much? A Model of Monetary Policy and Time-Varying Risk Aversion. *Journal of Financial Economics*, *146*(1), 71–89.
Rotemberg, J. J. (1982). Sticky Prices in the United States. *Journal of Political Economy*, *90*(6), 1187–1211.
Rubio-Ramirez, J. F., Waggoner, D. F., & Zha, T. (2010). Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference. *The Review of*

Economic Studies, *77*(2), 665–696.

Schorfheide, F. (2008). DSGE Model-Based Estimation of the New Keynesian Phillips Curve. *FRB Richmond Economic Quarterly*, *94*(4), 397–433.

Shimer, R. (2005). The Cyclical Behavior of Equilibrium Unemployment and Vacancies. *American Economic Review*, *95*(1), 25–49.

Sims, C. A. (1992). Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy. *European Economic Review*, *36*(5), 975–1000.

Smets, F., & Wouters, R. (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review*, *97*(3), 586–606.

Swanson, E. T. (2016). A Macroeconomic Model of Equities and Real, Nominal, and Defaultable Debt. *Unpublished Manuscript*.

Tallarini, T. D. (2000). Risk-Sensitive Real Business Cycles. *Journal of Monetary Economics*, *45*(3), 507–532.

Appendix A

A.1 Derivation of Equation (13)

As in Greenwald et al. (2014), define the time discount factor, β_t, as

$$\beta_t \equiv \frac{\exp(-\delta_t)}{\mathbb{E}_t\left[\frac{{C^s_{t+1}}^{-x_t}}{C_t^{s-x_{t-1}}}\right]}. \tag{A.1}$$

For the preference specification in Equation (12), the intertemporal marginal rate of substitution is

$$M_{t+1} = \beta_t \left(\frac{{C^s_{t+1}}^{-x_t}}{C_t^{s-x_{t-1}}}\right). \tag{A.2}$$

Substitute Equation (A.1) into (A.2) and simplify to get that

$$M_{t+1} = \exp(-\delta_t)\left(\frac{{C^s_{t+1}}^{-x_t}}{\mathbb{E}_t\left[{C^s_{t+1}}^{-x_t}\right]}\right). \tag{A.3}$$

Multiply the right-hand side of Equation (A.3) by $\frac{\frac{1}{C_t^{s-x_t}}}{\frac{1}{C_t^{s-x_t}}}$ and reorganize:

$$M_{t+1} = \frac{\exp(-\delta_t)}{\mathbb{E}_t\left[\left(\frac{C^s_{t+1}}{C^s_t}\right)^{-x_t}\right]}\left(\frac{C^s_{t+1}}{C^s_t}\right)^{-x_t}. \tag{A.4}$$

Equation (13) follows from the definition of the scaled time discount factor, $\hat{\beta}_t$, in Equation (15).

A.2 Derivations of Equations (26) & (27)

The Lagrangian for the firm's problem in Equation (25) is:

$$\begin{aligned}\mathcal{L} = &\cdots + Y_t - W_t N_t - \kappa V_t - I_t - \mu^n_t\left[N_{t+1} - (1-s)N_t - q(\theta_t)V_t\right] - \mu^k_t\left[K_{t+1} - (1-\xi)K_t - \Phi(I_t, K_t)\right] \\ &+ \mathbb{E}_t\Big[M_{t+1}\big(Y_{t+1} - W_{t+1}N_{t+1} - \kappa V_{t+1} - I_{t+1} - \mu^n_{t+1}\left[N_{t+2} - (1-s)N_{t+1} - q(\theta_{t+1})V_{t+1}\right] \\ &-\mu^k_{t+1}\left[K_{t+2} - (1-\xi)K_{t+1} - \Phi(I_{t+1}, K_{t+1})\right] + \cdots\big)\Big].\end{aligned}$$

The first-order conditions with respect to I_t and K_{t+1} are, respectively,

$$\mu_t^k = \frac{1}{\partial\Phi_t/\partial I_t} \tag{A.5}$$

$$\mu_t^k = \mathbb{E}_t\left[M_{t+1}\left[\frac{\partial Y_{t+1}}{\partial K_{t+1}} + \left(1-\xi+\frac{\partial\Phi_{t+1}}{\partial K_{t+1}}\right)\frac{1}{\partial\Phi_{t+1}/\partial I_{t+1}}\right]\right]. \tag{A.6}$$

From Equation (19),

$$\frac{\partial\Phi_t}{\partial I_t} = a_2\left(\frac{I_t}{K_t}\right)^{-\frac{1}{\nu}} \tag{A.7}$$

$$\frac{\partial\Phi_t}{\partial K_t} = a_1 + \frac{a_2}{\nu-1}\left(\frac{I_t}{K_t}\right)^{1-\frac{1}{\nu}}. \tag{A.8}$$

For the Cobb-Douglas production function in Equation (17), the marginal product of capital, $\frac{\partial Y_{t+1}}{\partial K_{t+1}}$, is equal to $\alpha\frac{Y_{t+1}}{K_{t+1}}$. Substitute Equations (A.7) and (A.8) into Equation (A.6) and divide both sides by μ_t^k to obtain Equation (26).

The first-order conditions with respect to V_t and N_{t+1} are, respectively,

$$\mu_t^n = \frac{\kappa}{q(\theta_t)} \tag{A.9}$$

$$\mu_t^n = \mathbb{E}_t\left[M_{t+1}\left[\frac{\partial Y_{t+1}}{\partial N_{t+1}} - W_{t+1} + (1-s)\mu_{t+1}^n\right]\right]. \tag{A.10}$$

The marginal product of labor, $\frac{\partial Y_{t+1}}{\partial N_{t+1}}$, is equal to $(1-\alpha)\frac{Y_{t+1}}{N_{t+1}}$. Substitute $\mu_{t+1}^n = \frac{\kappa}{q(\theta_{t+1})}$ into Equation (A.10) and divide both sides by μ_t^n to obtain Equation (27).

A.3 Derivation of Equation (28)

In order to prove Equation (28), I use a guess-and-verify approach. First, assume that $S_{t+1} = \mu_{t+1}^k K_{t+2} + \mu_{t+1}^n N_{t+2}$. The recursive formulation of Equation (25) is

$$S_t = \mathbb{E}_t\left[M_{t+1}\left(S_{t+1}+D_{t+1}\right)\right]. \tag{A.11}$$

Use $S_{t+1} = \mu_{t+1}^k K_{t+2} + \mu_{t+1}^n N_{t+2}$ to rewrite the right-hand side of Equation (A.11):

$$\begin{aligned}
S_t =& \mathbb{E}_t \left[M_{t+1} \left(\mu_{t+1}^k K_{t+2} + \mu_{t+1}^n N_{t+2} + D_{t+1} \right) \right] \\
=& \mathbb{E}_t \left[M_{t+1} \left(\mu_{t+1}^k \left((1-\xi) K_{t+1} + \Phi_{t+1} \right) + \mu_{t+1}^n \left((1-s) N_{t+1} + q(\theta_{t+1}) V_{t+1} \right) \right. \right. \\
& \left. \left. + Y_{t+1} - W_{t+1} N_{t+1} - \kappa V_{t+1} - I_{t+1} \right) \right] \\
=& \mathbb{E}_t \left[M_{t+1} \left(\mu_{t+1}^k \left((1-\xi) K_{t+1} + \frac{\partial \Phi_{t+1}}{\partial I_{t+1}} I_{t+1} + \frac{\partial \Phi_{t+1}}{\partial K_{t+1}} K_{t+1} \right) + \mu_{t+1}^n \left((1-s) N_{t+1} + q(\theta_{t+1}) V_{t+1} \right) \right. \right. \\
& \left. \left. + \frac{\partial Y_{t+1}}{\partial K_{t+1}} K_{t+1} + \frac{\partial Y_{t+1}}{\partial N_{t+1}} N_{t+1} - W_{t+1} N_{t+1} - \kappa V_{t+1} - I_{t+1} \right) \right] \\
=& K_{t+1} \mathbb{E}_t \left[M_{t+1} \left(\frac{\partial Y_{t+1}}{\partial K_{t+1}} + \left(1 - \xi + \frac{\partial \Phi_{t+1}}{\partial K_{t+1}} \right) \mu_{t+1}^k \right) \right] + \mu_{t+1}^k \frac{\partial \Phi_{t+1}}{\partial I_{t+1}} I_{t+1} \\
& + N_{t+1} \mathbb{E}_t \left[M_{t+1} \left(\frac{\partial Y_{t+1}}{\partial N_{t+1}} - W_{t+1} + (1-s) \mu_{t+1}^n \right) \right] + \mu_{t+1}^n q(\theta_{t+1}) V_{t+1} - \kappa V_{t+1} - I_{t+1} \\
=& \mu_t^k K_{t+1} + \mu_t^n N_{t+1},
\end{aligned} \tag{A.12}$$

in which the third equality follows from constant returns to scale for Y_{t+1} and Φ_{t+1}, and the last equality follows from the first-order conditions in Equations (A.5)-(A.10). Substitute Equation (A.12) into the expression for the stock return:

$$R_{t+1}^s = \frac{S_{t+1} + D_{t+1}}{S_t} = \frac{\mu_{t+1}^k K_{t+2} + \mu_{t+1}^n N_{t+2} + D_{t+1}}{\mu_t^k K_{t+1} + \mu_t^n N_{t+1}}. \tag{A.13}$$

Use the laws of motion for capital and employment to substitute out K_{t+2} and N_{t+2}, respectively, and substitute out D_{t+1} using Equation (24):

$$R_{t+1}^s = \frac{\mu_{t+1}^k \left((1-\xi) K_{t+1} + \Phi_{t+1} \right) + \mu_{t+1}^n \left((1-s) N_{t+1} + q(\theta_{t+1}) V_{t+1} \right) + Y_{t+1} - W_{t+1} N_{t+1} - \kappa V_{t+1} - I_{t+1}}{\mu_t^k K_{t+1} + \mu_t^n N_{t+1}}. \tag{A.14}$$

Since Y_{t+1} and Φ_{t+1} have constant returns to scale, we can reorganize Equation (A.14) and simplify to get that:

$$R_{t+1}^s = \frac{\left(\frac{\partial Y_{t+1}}{\partial K_{t+1}} + \left(1 - \xi + \frac{\partial \Phi_{t+1}}{\partial K_{t+1}} \right) \mu_{t+1}^k \right) K_{t+1}}{\mu_t^k K_{t+1} + \mu_t^n N_{t+1}} + \frac{\left(\frac{\partial Y_{t+1}}{\partial N_{t+1}} - W_{t+1} + (1-s) \mu_{t+1}^n \right) N_{t+1}}{\mu_t^k K_{t+1} + \mu_t^n N_{t+1}}. \tag{A.15}$$

Define the investment return as $R_{t+1}^k \equiv \frac{\frac{\partial Y_{t+1}}{\partial K_{t+1}} + \left(1 - \xi + \frac{\partial \Phi_{t+1}}{\partial K_{t+1}}\right) \mu_{t+1}^k}{\mu_t^k}$ and the hiring return as $R_{t+1}^n \equiv \frac{\frac{\partial Y_{t+1}}{\partial N_{t+1}} - W_{t+1} + (1-s) \mu_{t+1}^n}{\mu_t^n}$. Then,

$$\begin{aligned}
R_{t+1}^s &= \frac{\mu_t^k K_{t+1}}{\mu_t^k K_{t+1} + \mu_t^n N_{t+1}} R_{t+1}^k + \frac{\mu_t^n N_{t+1}}{\mu_t^k K_{t+1} + \mu_t^n N_{t+1}} R_{t+1}^n \\
&= w_t^k R_{t+1}^k + \left(1 - w_t^k \right) R_{t+1}^n,
\end{aligned} \tag{A.16}$$

where $w_t^k = \frac{\mu_t^k K_{t+1}}{\mu_t^k K_{t+1} + \mu_t^n N_{t+1}}$ is the value weight of the investment return in the stock return, which is equivalent to the capital share in the value of the firm.

A.4 Derivation of Equation (29)

Let $\partial J_t / \partial N_t$ be the marginal value of an employed worker to the representative household of workers, $\partial J_t / \partial U_t$ be the marginal value of an unemployed worker, $\partial E_t / \partial N_t$ be the marginal value of an employed worker to the representative firm, and $\partial E_t / \partial V_t$ be the marginal value of an unfilled vacancy to the firm. For simplicity, I assume that workers have log utility and are "myopic" in the sense that they only value the current payoff of employment or unemployment.

A worker-firm match turns an unemployed worker into an employed worker as well as an unfilled vacancy into an employed worker. The total surplus from the Nash bargain is:

$$H_t \equiv \left(\frac{\partial J_t}{\partial N_t} - \frac{\partial J_t}{\partial U_t} \right) / \phi_t + \frac{\partial E_t}{\partial N_t} - \frac{\partial E_t}{\partial V_t},$$

where ϕ_t is marginal utility. The equilibrium wage that arises from the bargain is

$$\max_{\{W_t\}} \left[\left(\frac{\partial J_t}{\partial N_t} - \frac{\partial J_t}{\partial U_t} \right) / \phi_t \right]^\eta \left(\frac{\partial E_t}{\partial N_t} - \frac{\partial E_t}{\partial V_t} \right)^{1-\eta},$$

where $0 < \eta < 1$ is the worker's bargaining weight. The outcome is the surplus-sharing rule:

$$\left(\frac{\partial J_t}{\partial N_t} - \frac{\partial J_t}{\partial U_t} \right) / \phi_t = \eta H_t = \eta \left[\left(\frac{\partial J_t}{\partial N_t} - \frac{\partial J_t}{\partial U_t} \right) / \phi_t + \frac{\partial E_t}{\partial N_t} - \frac{\partial E_t}{\partial V_t} \right],$$

in which the worker receives a fraction η of the total surplus. Since the representative worker is hand-to-mouth, his budget constraint is

$$C_t^w = (1 - \tau) W_t N_t + b U_t.$$

From the first-order condition of the firm's problem in Equation (25) with respect to V_t, we have that

$$\begin{aligned} \frac{\partial E_t}{\partial V_t} &= -\kappa + \mathbb{E}_t \left[M_{t+1} \frac{\partial E_{t+1}}{\partial N_{t+1}} q(\theta_t) \right] = 0 \\ \Longrightarrow \quad \frac{\kappa}{q(\theta_t)} &= \mathbb{E}_t \left[M_{t+1} \frac{\partial E_{t+1}}{\partial N_{t+1}} \right]. \end{aligned} \tag{A.17}$$

Similarly, from the first-order condition with respect to N_t, we have that

$$\frac{\partial E_t}{\partial N_t} = \frac{\partial Y_t}{\partial N_t} - W_t + (1-s)\mathbb{E}_t\left[M_{t+1}\frac{\partial E_{t+1}}{\partial N_{t+1}}\right]. \tag{A.18}$$

Given the assumptions above, $\partial J_t/\partial N_t = (1-\tau)W_t/C_t^w$ and $\partial J_t/\partial U_t = b/C_t^w$. Thus, the total surplus of the worker-firm relationship is

$$\begin{aligned} H_t &= (1-\tau)W_t - b + \frac{\partial Y_t}{\partial N_t} - W_t + (1-s)\mathbb{E}_t\left[M_{t+1}\frac{\partial E_{t+1}}{\partial N_{t+1}}\right] \\ &= (1-\alpha)\frac{Y_t}{N_t} - \tau W_t - b + (1-\eta)(1-s)\mathbb{E}_t\left[M_{t+1}H_{t+1}\right], \end{aligned} \tag{A.19}$$

as $\partial E_t/\partial N_t = (1-\eta)H_t$ from the surplus-sharing rule. Rewrite Equation (A.18) as

$$(1-\eta)H_t = (1-\alpha)\frac{Y_t}{N_t} - W_t + (1-\eta)(1-s)\mathbb{E}_t\left[M_{t+1}H_{t+1}\right]. \tag{A.20}$$

Combining Equations (A.19) and (A.20) yields:

$$\begin{aligned} (1-\alpha)\frac{Y_t}{N_t} - W_t + (1-\eta)(1-s)\mathbb{E}_t\left[M_{t+1}H_{t+1}\right] &= (1-\eta)\left((1-\alpha)\frac{Y_t}{N_t} - \tau W_t - b\right) \\ &\quad + (1-\eta)^2(1-s)\mathbb{E}_t\left[M_{t+1}H_{t+1}\right]. \end{aligned} \tag{A.21}$$

Simplifying,

$$\begin{aligned} W_t\left(1-(1-\eta)\tau\right) &= \eta(1-\alpha)\frac{Y_t}{N_t} + b(1-\eta) + \eta(1-\eta)(1-s)\mathbb{E}_t\left[M_{t+1}H_{t+1}\right] \\ &= \eta(1-\alpha)\frac{Y_t}{N_t} + b(1-\eta) + \eta(1-s)\mathbb{E}_t\left[M_{t+1}\frac{\partial E_{t+1}}{\partial N_{t+1}}\right] \\ &= \eta(1-\alpha)\frac{Y_t}{N_t} + b(1-\eta) + \eta(1-s)\frac{\kappa}{q\left(\theta_t\right)}, \end{aligned} \tag{A.22}$$

where the last line follows from Equation (A.17). Solve for W_t and reorganize to get that

$$W_t = \frac{1}{1-\tau+\tau\eta}\left[\eta\left((1-\alpha)\frac{Y_t}{N_t} + (1-s)\frac{\kappa}{q\left(\theta_t\right)}\right) + (1-\eta)\,b\right]. \tag{A.23}$$

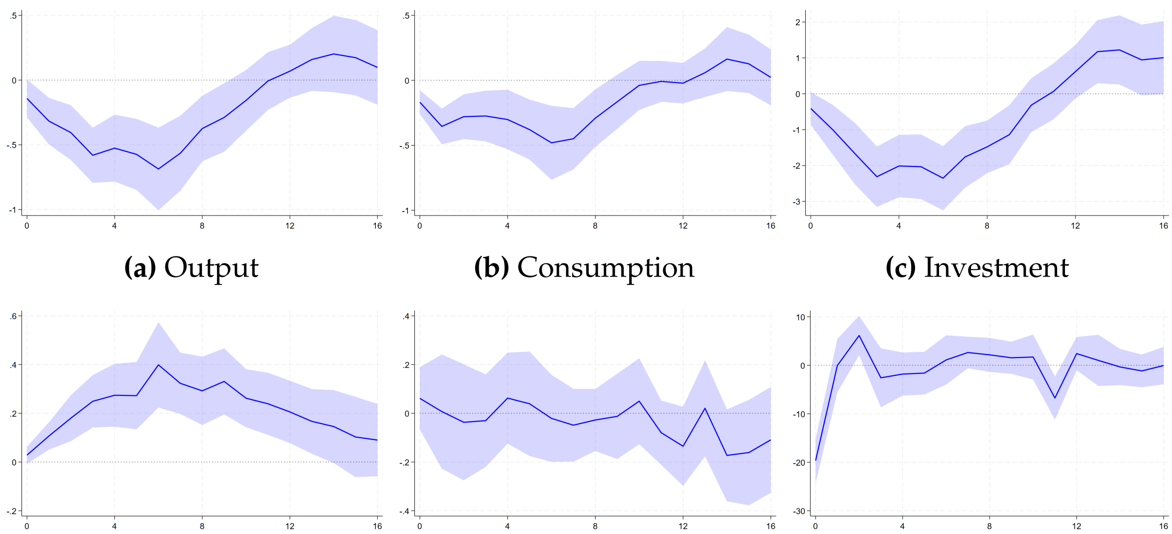
Appendix B

B.1 Empirical Impulse Responses to Risk Premium Shocks

B.1.1 Volatility Financial Conditions Index

Figure 10 plots the empirical local-projection impulse responses to a one-standard deviation VFCI shock. The maximum impact on macroeconomic quantities is obtained at the six-quarter horizon: a positive risk premium shock is associated with a 0.68 percentage-point decrease in output, a 0.48 percentage-point decrease in consumption, a 2.35 percentage-point decrease in investment, and a 0.39 percentage-point increase in the unemployment rate. All of these effects are statistically significant at the 5% level. Inflation increases by approximately 0.06 percentage points in the quarter after the shock, but the response is statistically insignificant at all horizons.

Figure 10: Empirical Impulse Responses to VFCI Shock



(a) Output **(b)** Consumption **(c)** Investment

(d) Unemployment **(e)** Inflation **(f)** Real Stock Return

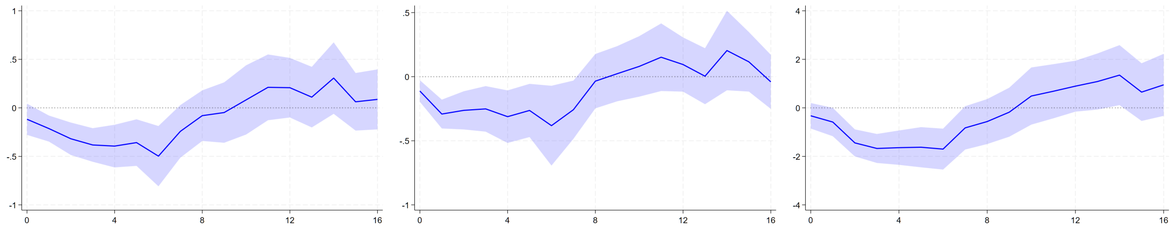
Note: The figures above plot the empirical local-projection impulse responses of output, consumption, investment, unemployment, inflation, and the real stock return to a one-standard deviation VFCI shock. The y-axis measures the response in percentage points, while the x-axis represents the number of quarters after the shock. Shaded areas are 90% confidence intervals. The sample is from 1968Q4 to 2023Q2.

B.1.2 Discount Rate News

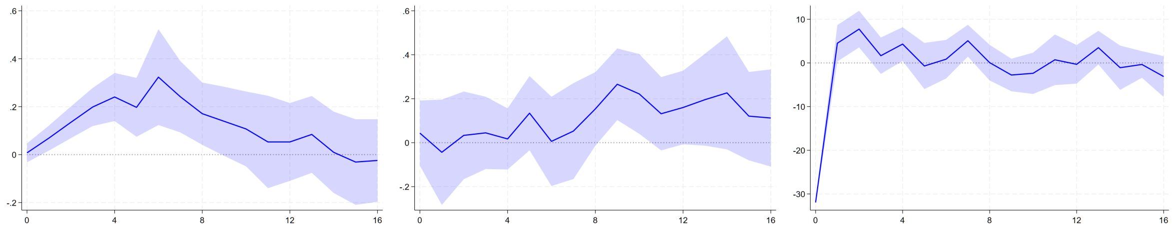
Figure 11 plots the empirical local-projection impulse responses to a one-standard deviation increase in discount rate news. The maximum impact on macroeconomic quantities is obtained at the six-quarter horizon: a positive risk premium shock is associated with a 0.49 percentage-point decrease in output, a 0.38 percentage-point decrease in consumption, a 1.7 percentage-point decrease in investment, and a 0.32

percentage-point increase in the unemployment rate. All of these effects are statistically significant at the 5% level. The response of inflation is insignificant at all horizons except 10 and 11 quarters after the shock, in which inflation increases by approximately 0.25 percentage points.

Figure 11: Empirical Impulse Responses to Discount Rate News



(a) Output **(b)** Consumption **(c)** Investment



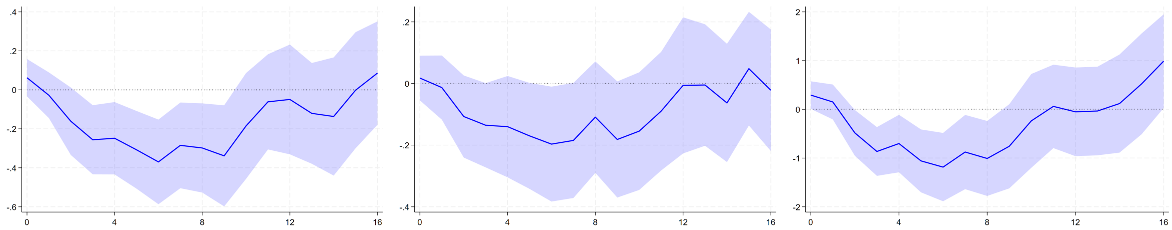
(d) Unemployment **(e)** Inflation **(f)** Real Stock Return

Note: The figures above plot the empirical local-projection impulse responses of output, consumption, investment, unemployment, inflation, and the real stock return to a one-standard deviation increase in discount rate news. The y-axis measures the response in percentage points, while the x-axis represents the number of quarters after the shock. Shaded areas are 90% confidence intervals. The sample is from 1968Q4 to 2023Q2.

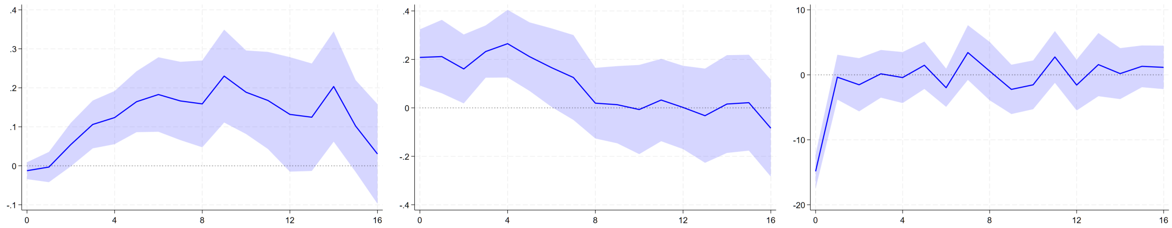
B.1.3 Structural Risk Premium Shock

Figure 12 plots the empirical local-projection impulse responses to a one-standard deviation structural risk premium shock. The maximum impact on macroeconomic quantities is obtained at the six-quarter horizon: a positive risk premium shock is associated with a 0.37 percentage-point decrease in output, a 0.19 percentage-point decrease in consumption, a 1.18 percentage-point decrease in investment, and a 0.18 percentage-point increase in the unemployment rate. All of these effects are statistically significant at the 5% level, although the response of consumption is insignificant at all other horizons. Inflation increases by approximately 0.2 percentage points, and the response is statistically significant for the six quarters after the shock.

Figure 12: Empirical Impulse Responses to Structural Risk Premium Shock



(a) Output **(b)** Consumption **(c)** Investment



(d) Unemployment **(e)** Inflation **(f)** Real Stock Return

Note: The figures above plot the empirical local-projection impulse responses of output, consumption, investment, unemployment, inflation, and the real stock return to a one-standard deviation structural risk premium shock. The y-axis measures the response in percentage points, while the x-axis represents the number of quarters after the shock. Shaded areas are 90% confidence intervals. The sample is from 1968Q4 to 2023Q2.

Appendix C

C.1 Affine Term Structure Model

To motivate the monotonicity restriction for the structural risk premium shock in Section 2.3, I consider an affine yield curve model with two state variables: the real risk-free rate, δ_t, and the price of risk, x_t. As in Section 4.1, I assume that the state variables follow AR(1) processes:

$$\delta_t = \bar{\delta} + \phi_\delta \delta_{t-1} + \varepsilon_{\delta t}, \quad \varepsilon_{\delta t} \sim N\left(0, \sigma_\delta^2\right) \tag{C.1}$$

$$x_t = \bar{x} + \phi_x x_{t-1} + \varepsilon_{xt}, \quad \varepsilon_{xt} \sim N\left(0, \sigma_x^2\right). \tag{C.2}$$

Let $p_t^{(n)}$ denote the log price of an n-period real bond with yield $y_t^{(n)} = -\frac{1}{n} p_t^{(n)}$. Conjecture that both bond prices and yields are affine in the vector of state variables, $z_t = [\delta_t, x_t]'$:

$$p_t^{(n)} = A_n + B_n' z_t \tag{C.3}$$

$$y_t^{(n)} = -\frac{1}{n} p_t^{(n)} = -\frac{1}{n} A_n - \frac{1}{n} B_n' z_t.$$

The real log SDF has the form:

$$\log M_{t+1} = -\delta_t - \frac{1}{2}\Lambda_t' \Sigma \Lambda_t - \Lambda_t' \varepsilon_{t+1},$$

where $\Lambda_t = \Lambda_0 + \Lambda_1 z_t$ determines the time variation in risk premia and $\varepsilon_{t+1} = [\varepsilon_{\delta t}, \varepsilon_{xt}]'$. The nominal log SDF is $m_{t+1}^{\$} = m_{t+1} - \pi_{t+1}$.

The conjecture in Equation (C.3) is trivially verified for $n = 0$, with $p_t^{(0)} = 0$ implying the initialization $A_0 = 0$ and $B_0' = 0$. To prove by induction, assume the conjecture holds for n. Then,

$$\begin{aligned}
p_t^{(n+1)} &= \log \mathbb{E}_t \exp\left\{-\delta_t - \frac{1}{2}\Lambda_t' \Sigma \Lambda_t - \Lambda_t' \varepsilon_{t+1} + A_n + B_n' \Phi_z z_t + B_n' \varepsilon_{t+1}\right\} \\
&= \log \mathbb{E}_t \exp\left\{-\delta_t - \frac{1}{2}\Lambda_t' \Sigma \Lambda_t + \left(B_n' - \Lambda_t'\right) \varepsilon_{t+1} + A_n + B_n' \Phi_z z_t\right\} \\
&= \log \mathbb{E}_t \exp\left\{-\delta_0 - \delta_1' z_t - \frac{1}{2}\Lambda_t' \Sigma \Lambda_t + \left(B_n' - \Lambda_t'\right) \varepsilon_{t+1} + A_n + B_n' \Phi_z z_t\right\} \\
&= -\delta_0 - \delta_1' z_t - \frac{1}{2}\Lambda_t' \Sigma \Lambda_t + \frac{1}{2} B_n' \Sigma B_n - B_n' \Sigma \Lambda_t + \frac{1}{2}\Lambda_t' \Sigma \Lambda_t + A_n + B_n' \Phi_z z_t \\
&= -\delta_0 - \delta_1' z_t + \frac{1}{2} B_n' \Sigma B_n - B_n' \Sigma \Lambda_0 - B_n' \Sigma \Lambda_1 z_t + A_n + B_n' \Phi_z z_t \\
&= \left(-\delta_0 + \frac{1}{2} B_n' \Sigma B_n - B_n' \Sigma \Lambda_0 + A_n\right) + \left(-\delta_1' - B_n' \Sigma \Lambda_1 + B_n' \Phi_z\right) z_t,
\end{aligned}$$

which implies that

$$A_{n+1} = -\delta_0 + \frac{1}{2} B_n' \Sigma B_n - B_n' \Sigma \Lambda_0 + A_n \tag{C.4}$$

$$B_{n+1}' = -\delta_1' - B_n' \Sigma \Lambda_1 + B_n' \Phi_z. \tag{C.5}$$