Macro Risks and the Term Structure of Interest Rates*

Geert Bekaert, Columbia University and the National Bureau of Economic Research,

Eric Engstrom, Board of Governors of the Federal Reserve System

Andrey Ermolov, Gabelli School of Business, Fordham University

September 9, 2019

Abstract

We use non-Gaussian features in U.S. macroeconomic data to identify aggregate supply and demand shocks while imposing minimal economic assumptions. Macro risks represent the variables that govern the time-varying variance, skewness and higher-order moments of these two shocks, with "good" ("bad") variance associated with positive (negative) skewness. We document that macro risks significantly contribute to the variation of yields and risk premiums for nominal bonds. While overall bond risk premiums are counter-cyclical, an increase in aggregate demand variance significantly lowers risk premiums. Macro risks also significantly predict future realized bond return variances.

Keywords: bond return predictability, term premium, macroeconomic volatility, business cycles, macro risk factors

JEL codes: E31, E32, E43, E44, G12, G13

---

*Contact information: Geert Bekaert - `gb241@gsb.columbia.edu`, Eric Engstrom - `eric.c.engstrom@frb.gov`, Andrey Ermolov - `aermolov1@fordham.edu`. Authors thank Michael Bauer, Mikhail Chernov, Isaiah Hull, Philipp Illeditsch, Aaron Pancost, Andrea Vedolin, seminar participants at Baruch College, Bilkent University, University of British Columbia, Bundesbank, City University of London, Duke, Fordham, University of Illinois at Urbana-Champaign, Imperial College, Oxford, Riksbank, Sabanci Business School, Tulane, University of North Carolina at Chapel-Hill, and Stevens Institute of Technology and conference participants at the 2015 Federal Reserve Bank of San Francisco and Bank of Canada Conference on Fixed Income Markets, 2016 NBER Summer Institute, 2016 Society of Financial Econometrics Meeting, 2017 European Finance Association Meeting, 2017 NBER-NSF Time Series Conference, Spring 2018 Midwest Macroeconomics Meetings, 2018 SFS Cavalcade North America, 2018 Örebro University and Kommuninvest Workshop "Predicting Asset Returns", 2018 Paris December Finance Meeting, and 2019 International Finance and Banking Society Medellin Conference for useful comments. All errors are the sole responsibility of the authors. The views expressed in this document do not necessarily reflect those of the Board of Governors of the Federal Reserve System, or its staff.

1 Introduction

A growing literature has established empirical links between various features of the macroeconomic environment and the term structure of interest rates. Ang and Piazessi (2003) and Bikbov and Chernov (2010) find that macroeconomic variables account for a large part of the variation in bond yields. Ludvigson and Ng (2009) show that macro factors predict holding period returns on bonds, even when also conditioning on forward rates and yield spreads. While Ludvigson and Ng employ many macro factors, similar results are present in Joslin, Priebsch, and Singleton (2014) who use a few key measures of economic activity and inflation, Cooper and Priestley (2009) who use the output gap, and Cieslak and Pavola (2015), who use a measure of long-run inflation expectations.

This evidence is important because it rejects the "spanning hypothesis", the idea that the yield curve spans all information relevant for forecasting future yields and returns, and that no variables other than those embodied by the current yield curve are needed for such forecasting (see Gürkaynak and Wright, 2012; and Duffee, 2013, for instance). The evidence that macro factors help forecast yields also has important economic ramifications. Ludvigson and Ng (2009) show that using macro factors to help measure risk premiums on bonds produces estimates of risk premiums that are counter-cyclical[1], in line with the predictions of some equilibrium models (see, e.g., Wachter, 2006).

More recently, however, this evidence has come under increased scrutiny for its statistical significance. Bauer and Hamilton (2017) show that conventional methods of inference are unreliable in the context of the predictive regressions that are commonly used to measure risk premiums, and they propose an alternative bootstrap methodology with adequate size and power properties. They find the evidence against the spanning hypothesis to be significantly weaker than is suggested by some published results and show that those results may be mostly spurious.

In this article, we introduce a new set of macro factors to re-examine the spanning hypothesis, and we link the new factors to the term structure of interest rates. Our "macro risks" represent the variables that govern the time-varying variance, skewness and higher-order moments of two distinct types of macroeconomic shocks, which we categorize as aggregate supply (AS) or aggregate demand (AD) shocks. Specifically, we

[1]Piazzesi and Swanson (2008) find counter-cyclical risk premiums in excess returns on federal funds futures, using employment growth as a predictor.

define aggregate supply shocks as shocks that move inflation and real activity in the opposite direction, whereas demand shocks are defined as innovations that move inflation and real activity in the same direction. This is the textbook Keynesian definition for which Blanchard (1989) finds empirical evidence examining the joint behavior of output, unemployment, prices, wages, and nominal money in the United States.

Our focus on these new risk factors is useful for two main reasons. First, a fundamental implication of many asset-pricing paradigms (e.g., the habit model of Buraschi and Jiltsov, 2007, or the long-run risk model of Bansal and Shaliastovich, 2013) is that bond risk premiums should be a function of expected *second*- and higher-order moments of macroeconomic fundamentals and explicitly not a function of first moments such as the expected rate of inflation. However, puzzlingly, the literature referenced above has mostly focused on explaining expected bond returns and risk premiums with the expectations of the *levels* or growth rates of macroeconomic variables or, even more simply, actual realized macroeconomic data (see, e.g., Ludvigson and Ng, 2009). Notable exceptions are Wright (2011) and Bansal and Shaliastovich (2013). Wright (2011) links term premiums to inflation uncertainty, whereas Bansal and Shaliastovich (2013) link bond risk premiums to consumption and inflation volatility. Bekaert, Engstrom, and Xing (2009) link the term structure to consumption growth volatility but do not explore risk premiums.

Second, we document that because the relative variances of supply versus demand shocks varies over time, the covariance between inflation and real activity potentially changes through time as well. Theoretically, the sign and magnitude of this covariance are important determinants of the risk premium for nominal bonds. When supply (demand) shocks dominate, real activity and inflation are negatively (positively) correlated, and bonds are a poor (good) hedge against macroeconomic fluctuations, presumably leading to relatively higher (lower) nominal term and risk premiums. Variations of this economic intuition have surfaced before. Fama (1981) shows how a negative correlation between inflation and real activity (stagflation) can induce a negative correlation between stock returns and inflation, a well-known puzzle in asset pricing (see Fama and Schwert, 1977, for a classic paper). Fama (1981) refers to the quantity theory of money to generate the negative correlation between real activity and inflation. In contrast, Kaul (1987) shows that if monetary authorities follow pro-cyclical monetary policies, the relationship between real activity and inflation may well be positive, using the pre-World War II period as an example. Neither Kaul (1987) nor Fama (1981) explore the relationship between these economic correlations and the term structure of interest rates. More recently, Campbell, Sunderam, and Viceira (2017) focus on the changes in sign in the covariation between the real stochastic discount factor and inflation, and what they imply for bond risk premiums, although they do not link their real stochastic discount factor to macroeconomic variables. When this covariation is negative, they argue, bonds are good hedges for the economic environment and vice versa.

Methodologically, we first extract aggregate supply and demand shocks for the U.S. economy from data on inflation, real GDP growth, core inflation, and the unemployment gap. Defining supply and demand shocks using only sign restrictions presents an identification problem. We resolve this issue with minimal further economic assumptions using a novel approach that exploits unconditional higher-order moments in the data, which we show to be highly statistically significant. This identification strategy is explored in more detail in a companion macroeconomics paper (Bekaert, Engstrom, and Ermolov, 2018). Despite this economically agnostic approach, we show that the "structural" supply and demand shocks that we identify exhibit dynamic properties consistent with some classic definitions of demand and supply shocks in the macroeconomic literature.[2]

Our second methodological step is to model the time variation in these risk factors that govern the conditional distributions of supply and demand shocks. For this purpose, we use the Bad Environment-Good Environment model ("BEGE", see Bekaert and Engstrom, 2017). In the BEGE model, separate macro risk factors drive "good" (positively skewed) and "bad" (negatively skewed) variances for each shock. As the "good" variance increases, the distribution for the shock becomes more positively skewed. Increases in "bad" variance may pull skewness into negative territory. Because we identify good and bad variance factors for both demand and supply shocks, we identify up to four risk factors in total, although we use standard statistical criteria to determine how many risk factors are necessary to fit the data parsimoniously.

After estimating the time series for the macro risks, we proceed to explore their relationship with the term structure of interest rates. First, we quantify the importance of our macro risk factors for explaining the variation in the shape of the yield curve, as quantified by the standard level, slope, and curvature factors. We find, perhaps not

[2]In particular, our demand shocks affect output temporarily, whereas our supply disturbances have a permanent effect on output, with neither having a long-run effect on the unemployment rate, just as in the classic Blanchard and Quah (1989) paper.

surprisingly, that our risk factors are not as important as more standard macroeconomic variables, such as expected inflation and real activity, for explaining variation in the yield curve over time.

A more novel finding is that we strongly reject the spanning hypothesis, which states, roughly, that yield curve factors - level, slope, and curvature - span all relevant information about risk premiums for government bonds. Specifically, we show that macro risk factors are economically and statistically significant predictors of excess bond returns even when also conditioning on yield curve factors. We establish our predictability results using the stringent framework of Bauer and Hamilton (2017), therefore resurrecting the evidence against the spanning hypothesis.

We also extend the analysis of Wright (2011) and Bansal and Shaliastovich (2013) by showing the importance of decomposing macroeconomic variation into components due to the variance of supply and demand shocks, and into the good and bad types of variance. We find that the time-variation in the macro risk factors for supply and demand implies that the covariance between inflation and real activity changes through time and sometimes switches sign. Our analysis links this time-variation to bond risk premiums by showing that demand (supply) variance negatively (positively) predicts bond excess returns. We also show that while overall the expected excess bond returns are countercyclical, in line with other findings in the literature, an increase in demand (supply) variance is associated with lower (higher) expected returns.

Finally, our novel macro risk factors prove to be statistically significant predictors of future realized bond return variances and are relatively more important predictors than are level macro factors and factors extracted from the term structure. There is a well-established literature linking equity return variances to macro factors (e.g., Engle, Ghysels, and Sohn, 2013)[3] but less work on bond return variances.[4] That yields alone cannot capture time variation in volatility for fixed income returns has been well-established by the literature on "unspanned volatility" (see, for instance, Collin-Dufresne and Goldstein, 2002; or Joslin, 2018). Our results suggest that a promising route for extending term structure models so that they can span volatility may be to add macroeconomic risk

[3]There is also work on the effect of inflation disagreement embedded in surveys on the term structure (see, e.g., Hong, Sraer, and Yu, 2017; or Ehling et al., 2018), and heteroskedasticity and disagreement are likely positively correlated.

[4]Baele, Bekaert and Inghelbrecht (2010) is an exception, but their focus is on comovements between bond and stock returns.

factors to the set of state variables. A similar strategy was explored by Joslin and Konchitchki (2018), who link uncertainty about corporate earnings performance to interest rate volatility.

The remainder of the paper is organized as follows. In section 2, we describe how we identify and model aggregate supply and demand shocks and define the macro risk factors. Section 3 describes the econometric methodology that we use to extract the structural shocks and the macro risk factors and provides empirical estimates for the U.S. economy. In section 4, we link the macro risk factors to bond market variables. The final section summarizes our key results and sets out an agenda for future research.

2 Modeling Macro Risks

2.1 An informal investigation of AS/AD Macro Risks

The main idea of the article is that macro uncertainty has different implications when it is associated with supply shocks, then when it is associated with demand shocks. Demand (supply) shocks move inflation and GDP growth in the same (opposite direction). As we discuss below, identifying and extracting such risks from macro data is no easy task. In this section, we construct informal proxies to AS and AD macro risks using data on US real GDP growth and inflation. Imagine we have identified unexpected shocks (residuals) to GDP growth and inflation (we do so formally using a standard vector autoregression, VAR, in Section 2.2) and define them as ϵ_t^i for $i = \pi$ (inflation), g (GDP growth). Our proxy for the time t demand variance variable is then:

$\sum_{i=0}^{n} \frac{\alpha^i}{\sum_{j=0}^{n} \alpha^j} \epsilon_{t-i}^{g} \epsilon_{t-i}^{\pi} \mathbb{1}_{\epsilon_{t-i}^{g} \epsilon_{t-i}^{\pi} > 0}$,

where ϵ_t^g is the GDP growth shock, ϵ_t^π is the aggregate inflation shock, α is a constant $\in (0; 1]$ and $\mathbb{1}$ is an indicator function. That is, the proxy is a weighted average of past cross-products of inflation and GDP growth shocks with terms being non-zero and positive only when the shocks have the same sign. It therefore proxies for demand shock variation and is high when demand shocks are large. Analogously, the time t supply variance proxy is defined as:

$-\sum_{i=0}^{n} \frac{\alpha^i}{\sum_{j=0}^{n} \alpha^j} \epsilon_{t-i}^{g} \epsilon_{t-i}^{\pi} \mathbb{1}_{\epsilon_{t-i}^{g} \epsilon_{t-i}^{\pi} < 0}$.

Below we use the parameters $\alpha = 1$ and $n = 9$ (that is, the equally weighted average of the past ten quarters), but our results are robust to varying α between 0.8 and 1 and n

between 5 and 15. For ease of interpretation, we rescale the resulting demand and supply variances to have unit variance. Figure 1 plots the time series. The supply variance is high in the seventies, but also around the 1990 and 2001 recessions. The demand variance is high in the early 1980s and the Great Recession, but the supply variance is elevated in the Great Recession as well, suggesting it was not a pure demand driven recession. These results conform with the usual macroeconomic intuition regarding the identification of various recessions in the US. It turns out that these variance series show a high correlation with the conditional variance series we formally derive in Section 2.3 (0.58 for the demand and 0.65 for the supply variance).

If financial markets recognize these macro risks as we surmise, supply risks should increase bond risk premiums, demand risks should decrease them. We run a regression, which we discuss more formally in Section 4, of excess bond returns on 4 level macro variables (including expected inflation and expected GDP growth as suggested by Cieslak and Povala, 2015, and Bauer and Rudebusch, 2017) and these demand and supply variances. We use excess returns on bonds of maturities ranging from 1 year, to 2, 5 and 10 years. Table 1 reports partial results for this exercise, focusing on the results for the macro risks. The coefficients can be interpreted as the percentage effects on the bond risk premium for a one standard deviation increase in the supply/demand variances. The signs are as expected, and the economic effects increase with horizon. While the supply variance coefficients are not statistically significantly different from zero, higher demand variances lower bond risk premiums in a statistically significant fashion. For 5-year maturity bonds, a one standard deviation increase in the demand variance is associated with a 1.5% lower bond risk premium (annualized).

The remainder of this article verifies whether this simple intuition holds when AS/AD risks are more formally extracted from the data. In section 2.2, we construct the macro shocks, and in section 2.3 we estimate the actual macro risk factors, representing the second and higher order moments of the AS and AD shocks.

2.2 Macro Shocks

We consider a four variable macro model, with the standard real GDP growth and inflation variables, but also core inflation and the unemployment gap. Core inflation, which strips out components of overall inflation that are particularly volatile such as energy and food prices, is, of course, a variable that is closely followed by monetary policy makers and it has been shown to be useful in forecasting future inflation. Ajello, Benzoni and Chyhruk (2012) in fact claim that adding core inflation to a macro system results in inflation forecasts that are as accurate as forecasts based on survey data (see Ang, Bekaert and Wei, 2007, for more on the accuracy of survey based inflation forecasts). This is relevant, because we use quarterly data starting in 1962 and thus cannot easily use survey forecasts (for instance, the quarterly Survey of Professional Forecasters started in 1969). Analogously, for many practitioners, the unemployment rate gap is preferred to GDP growth as an indicator of economic activity. Moreover, as Bauer and Rudebusch (2016) demonstrate, the unemployment rate shows little correlation with GDP growth and therefore contains useful alternative information about real economic activity.

Because we want to identify shocks to these four variables, it is important that we estimate their conditional means carefully. Following a long tradition in macroeconomics, we work within the vector autoregression moving average (VARMA) class of models so that our primary instruments for forecasting conditional means of the macro variables are lagged realizations and lagged shocks for our endogenous variables. In addition, because bond yields have well-established predictive power for economic variables (see Harvey, 1988, and many others, for the predictive ability of the term spread for real activity, for example), we add yields to our set of conditioning variables. Specifically, let the vector X_t consist of the 4 macro variables, and the one quarter and 10-year Treasury yields. We use a VARMA model to extract macroeconomic shocks from X_t:

$$X_t = B(L)X_{t-1} + C(L)u_t. \tag{1}$$

Next, we model the shocks to the macro variables, u_t, as functions of two structural shocks, u_t^s and u_t^d. The first fundamental economic shock, u_t^s, is an aggregate supply shock, defined so that it moves GDP growth and inflation in opposite directions, as happens, for instance, in episodes of stagflation. The second fundamental shock, u_t^d , is an aggregate demand shock, defined so that it moves GDP growth and inflation in the same direction as would be the case in a typical economic boom or recession. While we do not impose further sign restrictions, we expect core inflation to load on structural shocks with the same signs as aggregate inflation and we expect the unemployment gap to load negatively on both demand and supply shocks.

More concretely, we assume that the residuals in (1) depend on structural shocks and measurement errors:

$$u_t = \Sigma u_t^m + \Omega e_t \tag{2}$$

where $u_t^m = [u_t^s, u_t^d]'$ are structural AS/AD shocks and Σ is a 6x2 matrix containing the exposures of macroeconomic and yield shocks to structural shocks. In particular, we also impose the sign restrictions discussed earlier on Σ where the upper 4x2 block is:

$$\begin{bmatrix} -\sigma_{\pi s} & \sigma_{\pi d} \\ \sigma_{gs} & \sigma_{gd} \\ -\sigma_{\pi^c s} & \sigma_{\pi^c d} \\ -\sigma_{u^e s} & -\sigma_{u^e d} \end{bmatrix},$$

where only the $\sigma_{i,j}$ parameters in the top two rows are constrained to be positive; π^c represents core inflation and u^e the unemployment gap. Supply and demand shocks are assumed to be uncorrelated and normalized to have unit variance. Specifically, $Cov(u_t^d, u_t^s) = 0$ and $Var(u_t^d) = Var(u_t^s) = 1$.

The vector e_t in (2) represent "measurement error" shocks uncorrelated with u_t, with mean zero, unit variance and zero skewness and excess kurtosis and Ω is diagonal except for the interest rate block.[5] It is necessary to add these uncorrelated innovations to the macro series to avoid having a singularity in their covariance matrix as it would be impossible to invert two structural shocks from four macroeconomic shocks. Note that the orthogonal shocks may not just represent measurement error (as, e.g., in Wilcox, 1992). They may also represent important variation, not modeled in our framework, such as that arising from monetary policy shocks, stressed, e.g., in Campbell, Pflueger and Viceira (2015). Importantly, our model implies that any time-variation in higher order macro moments and all covariance dynamics are generated by the structural shocks u_t, not by e_t.

Finally, because macroeconomic data exhibit substantial non-Gaussian features (see, e.g., Evans and Wachtel (1993) for inflation, and Hamilton (1989) for GDP growth), we assume that demand and supply shocks are potentially non-Gaussian in that they may

[5]This interest rate block will only be relevant in the impulse response analysis described below.

have non-zero univariate skewness and excess kurtosis.[6] While there are several third and fourth-order cross-moments between supply and demand shocks, we also assume that these are zero. In particular, $E[(u_t^s)^i(u_t^d)^j] = E[(u_t^s)^i]E[(u_t^d)^j]$ for $i, j = 1, 2, 3$ and $i + j \leq 4$. We make these assumptions mostly for parsimony, but find that relaxing these additional restrictions does not affect the fit of the model.

The system in equation (2) that focuses on the first 4 macro variables is clearly under-identified if we employ only second order moments of the macro data: we have 8 "structural" exposures to the AS/AD shocks; there are 4 elements of Ω, and the structural shocks are assumed non-Gaussian, so we must also estimate their skewness and kurtosis (another 4 parameters). However, the covariance matrix of residuals would deliver only 10 moments. To achieve identification, we use unconditional higher order moments of the macro variables. For example, there are 8 available unconditional skewness and co-skewness moments. These moments, in conjunction with 10 available second moments, could in principle be used to identify the parameters of the system. Higher-order cross moments offer yet more moment restrictions to help identify the parameters.

While econometrically it is clear that non-Gaussianity achieves identification (see Lanne, Meitz, and Saikkonen, 2017, for a theoretical paper on obtaining identification through higher-order moments in a VAR), it is useful to clarify the economic sources of identification.[7] Co-skewness and co-kurtosis moments, for example, reveal information about the relative sensitivity of inflation and GDP growth to the structural shocks, depending on whether the latter are skewed or not. A particularly intuitive case would be one where the supply shocks are relatively Gaussian (zero skewness) and the demand shock relatively non-Gaussian (and negatively skewed). Suppose for ease of exposition that the skewness of supply shocks is literally zero (which, as we will see, is not far from the truth). Then, given the value of demand skewness, two co-skewness moments would aid identification of $\sigma_{\pi d}$ and σ_{gd}. If $E[u_t^g(u_t^\pi)^2]$, the "inflation squared" moment, is much more negative than $E[(u_t^g)^2 u_t^\pi]$, the "GDP growth squared" moment, inflation must be more sensitive to demand shocks than are GDP growth shocks and vice versa. Impor-

[6] We assume that orthogonal "measurement error" shocks e_t in (2) have zero skewness and excess kurtosis mostly for convenience, but this assumption also aids in the identification of the supply and demand shocks. That is, all the excess skewness and kurtosis among the macro variables must solely arise from the structural shocks.

[7] In a companion macro-oriented paper (Bekaert, Engstrom, and Ermolov, 2018), we expand on the economics behind the identification scheme.

tantly, our identification relies on unconditional non-Gaussianities and is valid whether the conditional moments vary through time or not.

2.3 Macro Risks

Finally, we formulate a dynamic model for the demand and supply shocks. To deliver the economic intuition described before, the variance of demand and supply shocks should vary through time. In this case, the model also implies that the conditional variance between inflation and GDP growth shocks is time-varying:

$$Cov_{t-1}[u_t^g, u_t^\pi] = -\sigma_{\pi s}\sigma_{gs} Var_{t-1} u_t^s + \sigma_{\pi d}\sigma_{gd} Var_{t-1} u_t^d, \tag{3}$$

where the subscripts on the Cov and Var operators denote that they may vary over time. Thus, when demand shocks dominate the covariance is relatively high but when supply shocks dominate it is low.

We define macro risks as the time-varying determinants of the second and higher-order moments of supply and demand shocks. We parameterize the distribution of supply and demand shocks using a model that accommodates conditionally non-Gaussian distributions, the Bad Environment-Good Environment (BEGE) model (Bekaert and Engstrom, 2017).

Following a BEGE structure, demand and supply shocks are component models of two independent distributions:

$$\begin{aligned} u_t^s &= \sigma_p^s \omega_{p,t}^s - \sigma_n^s \omega_{n,t}^s, \\ u_t^d &= \sigma_p^d \omega_{p,t}^d - \sigma_n^d \omega_{n,t}^d, \end{aligned} \tag{4}$$

where t is a time index, and σ_p^s, σ_n^s, σ_p^d, and σ_n^d are positive constants. We use the notation:

$$\begin{aligned} \omega_{p,t+1}^d &\sim \tilde{\Gamma}(p_t^d, 1), \\ \omega_{n,t+1}^d &\sim \tilde{\Gamma}(n_t^d, 1), \\ \omega_{p,t+1}^s &\sim \tilde{\Gamma}(p_t^s, 1), \\ \omega_{n,t+1}^s &\sim \tilde{\Gamma}(n_t^s, 1), \end{aligned} \tag{5}$$

to denote that $\omega_{p,t}^d$ follows a centered gamma distribution with shape parameter p_t^d and a unit scale parameter. The corresponding probability density function, $\phi(\omega^d_{p,t})$, is given by:

$$\phi(\omega^d_{p,t+1}) = \frac{1}{\Gamma(p^d_t)}(\omega^d_{p,t+1} + p^d_t)^{p^d_t - 1} exp(-\omega^d_{p,t+1} - p^d_t),$$

for $\omega^d_{p,t+1} > -p^d_t$; with $\Gamma(\cdot)$ representing the gamma function. Similar definitions apply to $\omega^d_{n,t+1}$, $\omega^s_{p,t+1}$, and $\omega^s_{n,t+1}$. Unlike the standard gamma distribution, the centered gamma distribution has mean zero. For such a distribution, the shape parameter equals the variance of the random variable.

The top of panel A in Figure 2 illustrates that the probability density function of $\sigma^d_p \omega^d_{p,t}$ (the "good" component of the demand shock) is bounded from the left and has a right tail. Similarly, the middle of panel A in Figure 2 shows that the probability density function of $-\sigma^d_n \omega^d_{n,t}$ (the "bad" component) is bounded from the right and has a left tail. Finally, the bottom of panel A in Figure 2 plots the component model of these two components which has both tails. The components of u^s_t have the same distributional properties. Hence, we define a "good" ("bad") shape parameter as one associated with a ω_p (ω_n)-shock.

The good (p^d_t, p^s_t) and bad (n^d_t, n^s_t) shape parameters of our macro shocks are assumed to vary through time in an autoregressive fashion as in Gourieroux and Jasiak (2006):

$$\begin{aligned} p^d_t &= \bar{p}^d(1 - \phi^d_p) + \phi^d_p p^d_{t-1} + \sigma^d_p \omega^d_{p,t}, \\ p^s_t &= \bar{p}^d(1 - \phi^s_p) + \phi^s_p p^s_{t-1} + \sigma^s_p \omega^s_{p,t}, \\ n^d_t &= \bar{n}^d(1 - \phi^d_n) + \phi^d_n n^d_{t-1} + \sigma^d_n \omega^d_{n,t}, \\ n^s_t &= \bar{n}^s(1 - \phi^s_n) + \phi^s_n p^s_{t-1} + \sigma^s_n \omega^s_{n,t}, \end{aligned} \tag{6}$$

where σ^j_i, $i = p/n$, $j = d/s$ are assumed to be positive. Note that positive $\omega^d_{p,t}$ shocks drive up GDP growth, as do the $\omega^s_{p,t}$ shocks, and those shocks are associated with an increase in both p^d_t and p^s_t. We call this "good volatility" because it induces more positive skewness in GDP growth. Conversely, positive realizations of $\omega^d_{n,t}$ and $\omega^s_{n,t}$ shocks drive down GDP growth and they are associated with an increase in "bad" volatility and more negative skewness. This explains the "BEGE" moniker.

Using the demand shock as an example, Panel B of Figure 2 illustrates possible conditional distributions of demand shocks which could arise as a result of the time variation in shape parameters in equation (6). In particular, the probability density function in the top of Panel B characterizes the situation where good volatility is relatively large and the component distribution has a pronounced right tail, while the probability density function in the bottom of Panel B corresponds to the case where bad volatility is relatively large and the component distribution exhibits a pronounced left tail.

At this point, we have set out an economy with four structural shocks ($\omega^d_{p,t}$, $\omega^d_{n,t}$, $\omega^s_{p,t}$, and $\omega^s_{n,t}$) and four state variables that govern higher-order moments, which we collect in $X^{mr}_t = [p^s_t, n^s_t, p^d_t, n^d_t]'$. These four state variables summarize the macroeconomic risks in the economy. Using the properties of the centered gamma distribution, we have, for example:

$$\begin{aligned}
&E_{t-1}[u^s_t] = 0, \\
&E_{t-1}[(u^s_t)^2] = (\sigma^s_p)^2 p^s_t + (\sigma^s_n)^2 n^s_t, \\
&E_{t-1}[(u^s_t)^3] = 2(\sigma^s_p)^3 p^s_t - 2(\sigma^s_n)^3 n^s_t, \\
&E_{t-1}[(u^s_t)^4] - 3(E_{t-1}[(u^s_t)^2])^2 = 6(\sigma^s_p)^4 p^s_t + 6(\sigma^s_n)^4 n^s_t,
\end{aligned} \tag{7}$$

and analogously for u^d_t.

Thus, the BEGE structure implies that the conditional variances of the macro variables vary through time, with the time-variation potentially coming from either demand or supply shocks, and either bad or good volatility. In addition, the distribution of the macro shocks is conditionally non-Gaussian, with time variation in the higher order moments driven by variation in X^{mr}_t.

3 Identifying Macro Risks in the US economy

While there are multiple ways to estimate the system in equations (1)-(2) and (4)-(6), the presence of the gamma distributed shocks makes the exercise nontrivial. We therefore split the problem into three manageable steps. First, we use standard techniques to estimate the VAR model and determine its order. Second, we filter the demand and supply shocks from the system in equation (2) by estimating a GMM system that includes higher-order unconditional moments of the macroeconomic variables. Third, once the demand and supply shocks are filtered, we can estimate the BEGE model on supply and demand shocks separately (exploiting our identifying assumptions) using approximate maximum likelihood as in Bates (2006). Importantly, the three steps are internally consistent.

A disadvantage of using a multi-step estimation process is that statistical inference is complicated by the fact that all steps after the first one use pre-estimated coefficients or filtered variables that are subject to sampling error. To account for these errors, we also execute the entire multi-step estimation process using data bootstrapped under the estimated parameters. The bootstrap procedure is described in Appendix A. Moreover, we conduct additional Monte Carlo analysis (see Section 3.4.4) to assess the finite sample performance of the estimators in steps 1 and 2. Theoretically, our model could be estimated in one step using Bayesian methods. However, given the high dimensionality of the parameter space, Bayesian estimation is difficult without tight priors. We begin by describing the data we use.

3.1 Data

The data are quarterly from 1962:Q2 to 2016:Q4 (219 quarters). Potentially, we could have included data back to 1947:Q1 (the starting date for GDP data). The later start date is chosen to exclude a period when there was higher measurement error in the GDP data (Bureau of Economic Analysis, 1993). Moreover, US long-term rates were pegged by the Federal Reserve prior to the Treasury Accord of 1951. For inflation (core inflation) we use 100 times log changes in the headline CPI index (CPI excluding food and energy) measured for the last month of each quarter, from the Bureau of Labor Statistics (BLS). Real GDP growth is 100 times the log difference in real GDP (in chained 2009 dollars) from the Bureau of Economic Analysis. The unemployment rate gap is the difference between the unemployment rate (in percent) from the last month of each quarter from the BLS, and the estimated level of the natural rate of unemployment published by the Congressional Budget Office.

Interest rate data consists of yields, prices and returns for nominal U.S. Treasury securities. For maturities of length 1 quarter and 1, 2, 3, 4 and 5 years, estimated yields for zero-coupon securities are taken from the Fama-Bliss (1987) data set (part of the CRSP database). For yields of maturity 10 years, data from 1962:Q2 through 1971:Q1 are from the McCullouch-Kwon (1993) data set. From 1971:Q1-2016:Q4, data for 10-year yields are from Gürkaynak, Sack, and Wright (2010). Yields at maturities other than those discussed above are estimated by linear interpolation. We use continuously compounded yields, expressed as annualized percentages.

3.2 Estimating VAR(p) and VARMA (p, q) models

To estimate the time series model for X_t, including inflation, real GDP growth, core inflation, the unemployment rate gap and short- and long-term interest rates, we first de-mean the variables. We then choose from a set of time series models, in particular, VARMA(i,j) for $i = 1, 2, 3$ and $j = 0, 1, 2, 3$, using standard information criteria. We only consider diagonal ("own lag") specifications for the MA components. As emphasized, for instance, by Dufour and Pelletier (2014), any identified VARMA model can be represented by using full (unrestricted) VAR specifications together with a sufficient number of diagonal MA terms.

Because some of these models are heavily parameterized (the highest-order ones have over 100 parameters), our estimation relies on a two-step projection-based procedure that was proposed by Hannan and Rissanen (1982) rather than maximum likelihood. Specifically, we first estimate by OLS a vector-autoregression with a large number of lags. We use 6 lags, but that choice does not appear material for the results. We then recover the estimated residuals from this step, $\hat{u}_t$. These residuals serve as a "plug-in" estimator of lagged shocks for the VARMA model, and then we estimate the VARMA model by OLS. We again recover the residuals from this step, providing new estimates of $\hat{u}_t$. This procedure is repeated until all of the estimated parameters of the VARMA and all of the estimated residuals converge, which we define as changing by less than 1e-6.

Model selection criteria are reported in Table 2. We use the standard Bayesian information criterion (BIC), but the Akaike information criterion (AIC) is modified to correct for small sample biases (Sugiura, 1978; Burnham and Anderson, 2004). The AIC model identifies the VAR(2) model as optimal. The BIC criterion identifies the VAR(1) model as optimal, but the VAR(2) comes in second place. We proceed by using the VAR(2) specification to identify shocks to the macro variables.[8]

[8] In a previous version of the paper, we used a VARMA(1,1) model to identify shocks to the macro variables and found similar results.

3.3 Identifying supply and demand shocks

3.3.1 Methodology

The VAR(2) model delivers time series observations on u_t, with their distributional properties driven by 4 unobserved state variables (the X_t^{mr} vector) which have non-Gaussian innovations. However, note that identification of the coefficients in Σ in equation (2), enables us to filter the supply and demand shocks from the original macro shocks u_t. With these structural shocks in hand, univariate BEGE systems on each of the demand and supply shocks can be estimated separately.

We use information in 2^{nd}, 3^{rd} and 4^{th} order unconditional moments of the reduced-form macroeconomic shocks to identify their loadings onto supply and demand shocks in a classical minimum distance (CMD) estimation framework (see, e.g., Wooldridge, 2002, pp. 445-446). Specifically, we calculate 48 statistics using the four macroeconomic shocks. These are the unconditional standard deviations (4), correlations (6), univariate (scaled) skewness and excess kurtosis (8), selected co-skewness (12), and selected co-excess kurtosis measures (18). In particular, we exclude third and fourth order moments that involve more than two different shocks such as $E(x_1 \times x_2 \times x_3)$.

With 48 moments to match and many fewer parameters in the structural model of equation (2), our system is substantially overidentified, thus requiring a weighting matrix. To generate a weighting matrix, we estimate the covariance matrix of the statistics, using a block bootstrapping routine. Specifically, we sample, with replacement, blocks of length 20 quarters of the 4 variable - vector of macroeconomic shocks, to build up a synthetic sample of length equal to that of our data. We calculate the same set of 2^{nd}, 3^{rd}, and 4^{th} order statistics for each of 10,000 synthetic samples. We then calculate the covariance matrix of these statistics across bootstrap samples. In principle, the inverse of this covariance matrix should be a good candidate as a weighting matrix for our CMD system. However, inspecting the bootstrapped covariance matrix, we found that the sampling errors for some statistics are highly correlated, leading to ill-conditioning of the covariance matrix. We therefore use a diagonal weighting matrix with the inverses of the bootstrapped variances of the statistics on the diagonal and zero elsewhere.[9]

[9]This weighting matrix is not asymptotically efficient and it also does not reflect sampling error associated with the VAR(2) parameters that were used to identify the macroeconomic shocks, but it ensures that all moments receive an easily interpretable positive weight in the objective function.

Table 3 reports the higher-order moments we use in the estimation. Not surprisingly, all volatility statistics are statistically significantly different from zero, but so are the coefficients of excess kurtosis. However, among the skewness coefficients, only the positive skewness of shocks to the unemployment gap is statistically significant while 4 of 12 co-skewness coefficients are significant. Over half of the co-kurtosis measures are statistically significant. The p-value for the joint significance of all the 3^{rd} and 4^{th} order moments is < 0.0001, which we interpret as a strong rejection of the hypothesis that the data are distributed unconditionally according to a multivariate Gaussian distribution.

We next use the information in these higher order moments to identify the loadings on our supply and demand shocks. We estimate a total of 13 parameters using our 48 estimated statistics. These can be grouped into three sets:

- The loadings of four macro shocks onto supply and demand shocks (8 parameters) in the matrix Σ in (2), imposing the sign restrictions described above.
- The share of variation of the macro shocks that comes from idiosyncratic variation or measurement error, that is the matrix Ω in (2)). We assume this share is constant across the four variables (1 parameter). We do this to impose a prior that all 4 series contribute (jointly) to demand and supply shocks. If we do not impose this restriction, the system tends to drive the variance of idiosyncratic factors to zero for the less noisy macro series, in which case the noisier macro series (such as real GDP growth) do not contribute much to the identification of supply and demand shocks.
- The skewness and kurtosis of the supply and demand shocks (4 parameters). Note that we do not assume a parametric model for the distribution of supply and demand shocks at this stage: we simply estimate their skewness/kurtosis coefficients as free parameters.

3.3.2 Empirical results

Table 3 shows that our CMD estimation misses only one moment by more than 1.96 standard errors (the fitted value for real GDP growth skewness is negative, whereas the sample value is positive, though not significantly so, a miss of 2.03 standard errors). Nevertheless, the test of the overidentifying restrictions does reject at the 10 percent level (p-value of 8.63 percent), showing that higher order moments indeed have statistical "bite".

In Table 4, Panel A, we report the supply and demand loadings for the various macro variables. These are generally quite precisely estimated. Our estimates suggest that demand shocks contribute more to the unconditional variance of inflation shocks than supply shocks. Real GDP growth, core inflation, and the unemployment gap all load roughly evenly on supply and demand shocks. We estimate the share of idiosyncratic variation for the four series to be relatively high at 44 percent.

Based on these loadings, we invert the supply and demand shocks from the macro shocks using a constant linear filter:

$$
\begin{aligned}
&u_t^m = K u_t, \\
&K = \Sigma'_{4\times 2}(\Sigma_{4\times 2}\Sigma'_{4\times 2} + \Omega_{4\times 4}\Omega'_{4\times 4})^{-1},
\end{aligned} \tag{8}
$$

where u_t and u_t^m are the vectors of macro and structural shocks, respectively, as in (2), Σ is the 4×2 loading of the macro shocks onto the supply and demand factors, and Ω is a diagonal 4×4 matrix of loadings onto the idiosyncratic shocks (corresponding to the 4 top rows of the matrices Σ and Ω in equation (2)). These loadings are implied by the usual projection formula under multivariate normality or the Kalman filter, which generates minimum root mean squared error (RMSE) estimates among linear filters with constant gain. Table 4, Panel B, reports the K-coefficients, which are all of the intuitive sign.

Finally, in Panel C of Table 4, we report the skewness and kurtosis of the filtered supply and demand shocks. Both shocks are leptokurtic but the demand shock is negatively skewed whereas the supply shock has essentially zero skewness. The departure from the Gaussian distribution of the demand shocks is clearly more pronounced than that of the supply shock. Yet, a standard (small sample corrected) Jarque-Bera test rejects the null of normality with p-values 0.015 and < 0.001, respectively for supply and demand shocks.

While we continue to refer to these shocks as AD/AS shocks, recall that for our purposes the sign restrictions have economic content, whatever label you may want to attach to the shocks. For example, our shock definitions are not consistent with recent micro-founded New Keynesian models (see Woodford, 2003). However, our sign restrictions are present in other classic papers as well, such as Shapiro and Watson (1988) or Gali (1992) but are typically accompanied by a set of additional economic restrictions (e.g., that demand shocks have no long run effect on the level of GDP as in the classic Blanchard and Quah (1989) paper) which we do not need. In Appendix A, we show that our macro shocks indeed have these standard Keynesian short and long-term effects despite our methodology inferring the shocks using only sign restrictions and higher order moments of macro shocks. We can also verify whether NBER recessions are demand or supply driven by inspecting the time series of supply and demand shocks, presented in Figure 3 (NBER defined recessions are shaded). On a relative basis, the first three recessions (1969-1970, 1973-1975, 1980) were predominantly supply driven whereas three of last four were more demand driven (the exception being the 1990-91 recession). For the first five recessions, these results are broadly consistent with Gali's (1992) results. Our results for the Great Recession suggest that AD shocks were slightly larger than AS shocks. A surprisingly large role of supply shocks is not inconsistent with the results in Ireland (2011) or Mulligan (2012), for example. At the same time, recent work by Bils, Klenow and Malin (2012) and Mian and Sufi (2014) stresses lower aggregate demand as the main cause of the steep drop in employment during the Great Recession.

3.4 Estimating Macro Risk Factors

Note that the identification scheme for structural shocks described above is completely model-free, making our methodology applicable with any statistical model which can accommodate non-Gaussian unconditional moments in the data. Given the structural shocks, we are left to identify the BEGE model parameters. We use an estimation and filtering apparatus due to Bates (2006). The methodology is similar in spirit to that of the Kalman filter, but the Bates routine is able to accommodate non-Gaussian shocks. The details of the estimation are in Appendix B.

3.4.1 Parameter Estimates

The parameter estimates for the BEGE model are reported in Table 5. For the demand shock, the parameters governing the "good environment" state variable, p_t, generate behavior similar to that of a Gaussian stochastic volatility model. The unconditional mean of the process, $\bar{p}$, hits an upper bound fixed at 20. Recall that p_t is the shape parameter for one of the two component gamma distributions for demand shocks. With the shape parameter of over 10, the gamma distribution appears nearly Gaussian and further increases in the shape parameter do not substantially change the shape of the distribution.[10] That said, there is substantial variation in the level of the process over time and strong autocorrelation, with a persistence parameter of nearly 0.94. The properties of the bad environment state variable for demand shocks, n_t, contrasts sharply with those of p_t. The unconditional mean of n_t is just 0.34. This implies that the bad environment variable is very non-Gaussian. In particular, its unconditional skewness is $\frac{2}{\sqrt{\bar{n}^d}}$, or 3.45 and its kurtosis is $\frac{6}{\bar{n}^d}$ or 17.86. Recall that because demand shocks load negatively onto the bad-environment shocks by construction, this generates substantial negative skewness for demand shocks. The bad environment shape parameter is also less persistent than the good environment variable, therefore capturing rather short-lived recessionary bursts (0.72 versus 0.94 autocorrelation).

The BEGE parameter estimates for supply shocks are broadly similar to those for demand shocks. The mean of p_t hits the upper bound of 20, suggesting nearly Gaussian innovations, albeit with substantial variation in volatility. Good supply variances are very persistent with an autoregressive coefficient of nearly 0.99. The supply bad-environment distribution is substantially non-Gaussian with the unconditional mean of the shape parameter equal to 4.00. This implies unconditional skewness of 1.00. The shock has similar persistence to the bad environment demand shock, suggesting that supply driven recessions may have similar duration to demand driven recessions.[11]

The bootstrap procedure in Appendix A allows us to verify the small sample properties of our estimation approach. In unreported results, we find that average estimates over the bootstrap samples are essentially unbiased. Bootstrap standard errors in Table 5 confirm that the parameter variation across the samples is reasonable.

3.4.2 Macro Risks

The Bates (2006) procedure also delivers filtered estimates of X_t^{mr}, our 4 macro risk factors. Our model implies that the total conditional variances of demand and supply shocks are the sum of the good and bad components. These are plotted in Figure 4

[10] We also conduct an estimation using an upper bound for $\bar{p}$ of 200. The likelihood values barely change and the macro risks (shape parameters) correlate very highly (usually ≥ 0.99) with the ones obtained from the current estimation.

[11] The astute reader will notice that seven parameters are reported for the supply and demand processes, but there are only six independent parameters required for the estimation, because the unconditional variance of demand and supply shocks is restricted to equal 1. However, the $\bar{n}$-parameters can be expressed as functions of the other model parameters. Their standard errors are calculated using the delta method.

(NBER recessions are shaded). Using the properties of the centered gamma distribution in (7), we define good demand variance as the constant $(\sigma_p^d)^2$ multiplied by the good demand macro risk factor p_t^d and bad demand variance as the constant $(\sigma_n^d)^2$ multiplied by the bad demand macro risk factor n_t^d. The good demand variance (see Panel A) was relatively high in the 70s and the early 80s, and then decreased to low levels consistent with the Great Moderation. The bad demand variance shows much less pronounced low frequency variation but increases in most recessions with notable peaks in the 1981-82, 2001, and the recent Great Recession. It also shows short-lived peaks twice in the decade between 2000 and the beginning of the Great Recession.

Panel B of Figure 4 performs the same exercise for supply variances. Analogously to demand variances, we define the good supply variance as the constant $(\sigma_p^s)^2$ multiplied by the good supply macro risk factor p_t^s and bad supply variance as the constant $(\sigma_n^s)^2$ multiplied by the bad supply macro risk factor n_t^s. The level of good variance does not show much time-variation but is more elevated up until mid-1980s after which it appears to trend down. The bad supply variance appears higher in the stagflationary episodes of the 1970s, but it peaks in most recessions. Its increase in the Great Recession is extreme, starting towards the end of the period and exceeding its unconditional average level of 0.46 until 2012Q1.[12] The secular decline that one might associate with the Great Moderation appears to come from the good variances of both supply and demand shocks.

Panel C of Figure 4 plots together the conditional variances of demand and supply shocks. Given that both supply and demand shocks have unit variance, the graph immediately gives a sense of which variance dominates. In terms of "variance" peaks, the 1981-82, and Great Recession are dominated by demand variances, the other recessions by supply variance peaks.

3.4.3 Conditional Covariances between Macroeconomic Time Series

From the perspective of theoretical asset pricing, an important implication of our structural framework regards the covariance between inflation and real activity. From Equation 3, it is evident that in an environment where demand (supply) variances dominate, the conditional covariance between inflation and real activity is positive (negative). To the extent that variances are persistent, changes in this covariance may have impor-

[12]Campbell, Pflueger, and Viceira (2015) suggest that supply shock volatility decreases after 1980 but its decrease may have been masked by changes in monetary policy, at least until 2000.

tant ramifications for term and bond return premiums, which we examine in Section 5. Surprisingly, to our knowledge, sign-switching macro-correlations have so far only been documented for consumption growth and aggregate inflation (Hasseltoft and Burkhardt, 2012, Ermolov, 2015, Boons, Duarte, de Roon, and Szymanowska, 2017, and Song, 2017).

Figure 5 graphs the conditional covariance between, respectively, inflation and real GDP growth and also between core inflation and the unemployment gap (where the aforementioned signs are reversed). Overall, the covariance is mostly positive (over 90 percent of the time), which is driven by the important contribution of (good) demand variances to all macro variables. For the inflation-GDP growth covariance, there is substantial time variation, but the covariance rarely becomes negative. Early in the Great Recession, demand shocks generate a local peak in the covariance but subsequent large supply shocks then bring the covariance down. A mirror image of this happens for the core inflation-unemployment gap covariance. There, we see more frequent sign switches and the covariance remains positive until 1975, in a supply shocks driven macro-environment. Appendix C confirms that similar covariance patterns also emerge in other conditional covariance models.

An overall covariance of near zero can in fact hide some strong structural non-zero sources of comovement from structural risk factors. To see this more clearly, we also show the good and bad supply and demand covariance components of the total covariance. For example, the near-zero correlation between real GDP and inflation from 2000 up to the onset of the Great Recession (with occasional peaks) is the sum of a sizable positive covariance driven by good and bad demand shocks and a sizable negative covariance driven by supply shocks. In the Great Recession, the conditional bad variance of both kinds of shocks shoots up, with the bad demand shock first ratcheting the covariance upwards, and bad supply variance later bringing it down substantially. Similar movements happen for the core inflation-unemployment covariance with the covariance actually switching signs.

3.4.4 Econometric Concerns

Our empirical methodology consists of several steps and may therefore raise concerns about its ability to deliver trustworthy parameter estimates. While the two first steps in the estimation methodology, the VAR, which delivers the reduced-form shocks, and the GMM procedure, which delivers the identification often supply and demand shocks, do not assume constant conditional moments, they are inefficient, in that they do not exploit the true stochastic nature of the shocks. A one step estimation is computationally infeasible in this context however; and, as is well known, the finite sample properties of "efficient" estimators are not always stellar. Both of our estimators deliver consistent estimates of the true parameters, under the usual stationarity and ergodicity assumptions. However, our macro data are characterized by strong non-Gaussianities, time-variation in second and higher order moments, and high persistence. With such data, standard econometric methodologies may suffer from severe finite sample biases and poorly estimated standard errors. To examine whether our methodology works well in finite samples, we supplement our bootstrap with an additional Monte Carlo exercise. Given the final estimation results, we have a data generating process (DGP) under the null, with time-varying volatility, skewness and kurtosis, following the highly non-linear BEGE dynamics. We draw the BEGE demand and supply shocks from the model and normally distributed measurement errors. Thus, we simulate shocks that constitute residuals for the VAR system and can generate artificial macro data. With these data in hand, we then repeat steps 1 and 2 from the estimation methodology, under the null of the estimated model, and with samples of the size that we use in the original estimation. We conduct 10000 Monte Carlo replications. As a useful comparison, we also generate data using fully normally distributed shocks, to verify whether the non Gaussianities in our DGP worsen the small sample performance of our estimation methodology.

We relegate full tables to the Online Appendix and provide a short summary of the results here. First, there are a grand total of 72 VAR parameters (since it is a second order VAR for 6 variables). The coefficients show, as expected, some biases, but they are not terribly large and not universally worse for the BEGE system as opposed to the Gaussian system. More important in terms of inference, is whether the small sample estimation delivers accurate confidence intervals for the parameters. To examine coverage ratios, we simply calculate how many OLS coefficients fall outside the 90% confidence intervals in the data. If the estimation is well behaved, they should be close to 10%, the nominal level. For Gaussian shocks, the error rate varies between 10.47% and 12.26% (there is only one coefficient for which the error rate is higher than 12%). For the BEGE Monte Carlo, the error rate varies between 10.23% and 12.51%. Clearly, the finite sample performance of the VAR is quite good, and not noticeably worse under BEGE distributed shocks than under Gaussian shocks.

We also examine the performance of Step 2 in the econometric methodology, the GMM step, that identifies the critical loading parameters on the structural shocks (see equation 2). In Table 6, the first column repeats the data estimate; in the second column we show the average estimates across the Monte Carlo replications (with BEGE shocks), and the third column shows the median across the 10000 replications. Compared to the standard errors reported in Table 4, the biases are very small, giving us confidence that the estimation methodology is accurate.

4 Macro Risks and the Term Structure

In this section, we explore the relation of macro factors with the term structure of interest rates. In the preceding sections, we have identified four novel macro-risk factors (p_t^d , n_t^d, p_t^s and n_t^s). These variables can be interpreted as "good" or "bad" conditional volatilities of demand and supply shocks, but their time variation also changes the entire conditional distribution of these shocks. For comparison with the existing literature on explaining bond yields and returns using macro data, we also examine the performance of "level" macro factors, which include expected inflation, expected core inflation and expected real GDP growth (we use the previously described VAR(2) system to compute these expectations). We also use the unemployment gap as a macro level factor. Thus, there are a total of 8 macro-factors we consider.

We address four questions. First, we ask whether macroeconomic factors help explain the yield curve. Second, we investigate the predictive power of our new macro risk factors for bond excess returns. Third, we also explore how the macro risk factors affect term premiums. Finally, we examine the predictive power of the macro factors for realized bond variances.

4.1 Macro Risks and the Yield Curve

We start by computing the classic yield curve financial factors. The "level" factor is the equally weighted average of all yields (from the one year to the 10 year maturity); the "slope" factor is the difference between the 10 year yield and one quarter yield; and finally, the "curvature" factor subtracts twice the two-year rate from the sum of the one quarter rate and the 10 year yield. Taken together, these three factors span the overwhelming majority of variation in yields at all maturities. Thus, to operationalize our test of whether macro factors explain yields, we test whether the macro factors explain variation in these three factors. Table 7 reports R^2 statistics from regressions of the financial factors onto the macro factors. We report the parameter coefficients in the Online Appendix, because the coefficients are difficult to interpret for yields. For example, bad supply volatility should increase the term premium and thus increase yields, but may also decrease yields through a precautionary savings mechanism.

Table 7 reports results regarding the macro level factors and the macro risk factors. First, the explanatory power of the macro level factors alone for the financial factors is substantial, with the adjusted R^2's about 70, 60, and 30 percent respectively for the level, slope, and curvature factors.[13] Second, we proceed to determine the increment in the adjusted R^2 statistics resulting from the addition of the macro risk factors and its statistical significance using the bootstrap test of Bauer and Hamilton (2017). The null hypothesis is that macro risks are unrelated to financial factors. Following Bauer and Hamilton (2017), we simulate 5,000 samples of historical length under the null and compute the p-value as the proportion of samples where the adjusted R^2 increases by at least as much as in the data after the inclusion of macro risks that are, by construction in the bootstrapped samples, unrelated to the yield curve. We find that the macro risks contribute in a statistically significant fashion to all factors, but the statistical and economic significance is much larger for the level (an adjusted R^2 increase of 7.5%) and curvature (an adjusted R^2 increase of 12.5%) factors.

Appendix D reports some additional results, showing that the boost in explanatory power due to the macro risk factors survives the inclusion of the contemporaneous macro level factors constructed by Ang and Piazzesi (2003)[14] as control variables, but becomes statistically insignificant for the slope factors. It also survives the inclusion of realizations (instead of expectations) of macro level factors (in that case, the relative contribution of the macro risk factors is even more substantive).

4.2 Macro Risks and Bond Return Predictability

The literature on bond return predictability is voluminous, but mostly focuses on using information extracted from the yield curve to predict future holding period returns (e.g. Cochrane and Piazzesi, 2005). Ludvigson and Ng (2009) find that "real" and "inflation"

[13] The ability of the factors to explain the variation in the first three principal components is slightly lower (with the decline in R^2 most prominent for the third principal component, relative to curvature).

[14] The factors in Ang and Piazzesi (2003) are 12 lags of a measure of inflation and a measure of real activity (they are the first principal components of a number of empirical measures).

factors, extracted from a large number of macroeconomic time series, have significant forecasting power for future excess returns on nominal bonds and that this predictability is above and beyond the predictive power contained in forward rates and yield spreads. Also, the bond risk premia implied by these regressions have a marked countercyclical component. Bansal and Shaliastovich (2013) show that consumption growth and inflation volatility predict excess bond returns. Cieslak and Pavola (2015) uncover short-lived predictability in bond returns by controlling for a persistent component in inflation expectations. Barillas (2011) shows that the predictability due to macro factors for excess bond returns is economically significant. However, Bauer and Hamilton (2017) have cast a pall over the literature that uses macro factors to explain future bond returns, calling into question the statistical significance of many of these widely-cited results.

In Tables 8 and 9, we explore the link between future bond returns and our macro factors. We focus on excess one-quarter holding period returns relative to the one quarter yield. This avoids the use of overlapping data which can spuriously increase R^2's in predictability regressions due to the high autocorrelation (Bauer and Hamilton, 2017). Nonetheless, all statistical inference is calculated using the small-sample bootstrap of Bauer and Hamilton (2017). To delve into the economic mechanism by which macro risks forecast future bond returns, Table 8 presents the coefficients from forecasting regressions that include both level macro and macro risk factors.[15] Individually, there are few significant coefficients. Of the macro level factors, expected core inflation enters with a positive sign, while expected aggregate inflation enters with a negative significant coefficient of similar magnitude, and is highly significant at all maturities. We find that including expected core inflation is critical to get a significant expected aggregate inflation coefficient in these regressions. This might be related to the results in Cieslak and Pavola (2015) finding a similarly negative coefficient on expected inflation when yields are included in the regression as core inflation is closely related to yields (Ajello, Benzoni, and Chyhruk, 2012). Of the macro risk factors, the bad demand variance has a negative significant coefficient and the bad supply variance a positive (albeit mostly insignificant) coefficient. Therefore, consistent with intuition, being in a risky (that is volatile) demand environment, where bonds are good hedges against general macroeconomic risks, reduces the risk premium on bonds, and the reverse is true in the case of a supply envi-

[15]Including financial factors (level, slope, and curvature) in the regressions does not materially change macro factors and macro risks signs except that the p_t^s-signs switches from being insignificantly positive to being insignificantly negative.

ronment. The effect of bad demand variance is economically large: for example, for the 10 year maturity a one standard deviation increase in the bad demand factor decreases the expected annualized excess bond return by 3.38 percentage points (the risk factors were standardized to a unit variance). The corresponding coefficients increase with bond maturity. The coefficients on the "good" demand risk factors are also negative and significantly different from zero, with coefficients that are even larger than for the bad demand variance factor.[16]

Note from equation (7) that two demand (supply) state variables linearly span the conditional (unscaled) skewness and variance of demand (supply) shocks. Thus, we can conduct the analysis in Table 8 using variances and skewness of demand and supply shocks instead of our macro risk factors. If variances are more persistent than skewness, which is likely, we should expect variances to have relatively larger effects on risk premiums. This is actually what we find (see the Online Appendix). Variances tend to be economically and statistically significant predictors while skewness is statistically insignificant and economically much less significant than variance. Importantly, our main results about demand (supply) variances decreasing (increasing) bond risk premiums remain intact. By construction, the R^2 for the regressions are the same as the ones in Table 8.

Table 9 mimics the regression set up in Ludvigson and Ng (2009) and Bauer and Hamilton (2017) including the financial factors (that is, the level, slope and curvature factors) in the base regression. The adjusted R^2's produced by the financial factors alone are significantly boosted after including both macro level factors and macro risk factors. For maturities from 1 to 10 years, the R^2's from regressions including only financial factors are around 7 percent. Macro level factors only increase the R^2 by 2 to 3 percentage points at short horizons with the increase only significant at the 10 percent level. Macro risks further increase the R^2s by about 4-5 percentage points for short maturities and by about 3 percentage points at the longer maturities. Macro risks alone

[16]To further elaborate on the risk premium intuition, we also added the contemporaneous demand and supply shocks (u^d_{t+1} and u^s_{t+1}) to the bond return regressions. In unreported results, we find that the supply shocks carry positive but economically small and statically non-significant coefficients and the demand shocks carry negative coefficients that are significant at the 5% (short maturities) and 1% (long maturities) levels and become larger in magnitude with maturity. That is, realized bond excess returns are high if a negative demand shock occurred during the holding period. In unreported results, we also explore excess return predictability regressions which include interaction terms between macro risks and the financial factors (the level, slope, and curvature). However, none of the interaction terms is significant at the conventional significant levels.

increase the adjusted R^2 by 6-7 percentage points at short maturities and 3-4 percentage points at long maturities compared to the specification where only financial factors are included. These increases in explanatory power are statistically significant under the Bauer and Hamilton (2017) bootstrap, for testing the null of no predictability coming from macro risks. This is important given that Bauer and Hamilton (2017) have shown that the additional predictive power coming from macro factors over financial factors, as, e.g., in Ludvigson and Ng (2009) and Joslin, Priebsch, and Singleton (2014), often does not survive when p-values are computed using bootstrap procedure. Additionally, while macro risks significantly increase explanatory power for the specification which includes financial and macro level factors, the increase in adjusted R^2 from macro level factors for the specification which already includes financial factors and macro risks is economically small and statistically insignificant. Appendix D reports the return predictability of our macro level and risk factors over Ang-Piazzesi factors, showing an increase in the adjusted R^2's by about 4 percentage points.

Given that previous studies have considered macroeconomic "level" and "risk" factors in isolation and that factors measuring macroeconomic risk have received scant attention in such investigations, the relative predictive power of risk factors is of interest. Table 8 indicates that the adjusted R^2 from macro level factors alone in excess return regressions is around 4-5 percent with macro risk factors contributing additional 2 percentage points.

Ludvigson and Ng (2009) found the bond risk premiums implied by their predictive regressions, which included both yield variables and macro-factors, to be counter-cyclical. It is not difficult to obtain counter-cyclical real bond risk premiums in economic models, e.g., in habit models with counter-cyclical prices of risk (see, e.g., Wachter, 2006). Our framework suggests that not all recessions are equal in this respect. Our predictive regressions indicate that risk premiums are, everything else equal, lower when the macro-environment is primarily demand driven. To verify the cyclicality of bond risk premiums that are implied by our regressions, we use the fitted values of the predictive regressions[17] as estimates of these risk premiums and regress them on a recession indicator, the ratio of the aggregate demand variance, including the good and bad variances, to the corresponding aggregate supply variance, and the interaction of the two. We rescale the demand/supply ratio variable to have a standard deviation of one. While we confirm

[17] Including financial factors (level, slope, and curvature) to construct the expected excess bond returns does not materially change any of the results.

all the reported results using a NBER dummy, NBER recessions are not available in real time. We therefore use the so-called "anxious" index from the Survey of Professional Forecasters as a recession indicator. This index measures the probability of a decline in real GDP from this quarter to the next. We define the recession dummy to be one if the anxious index for the quarter in which the survey is conducted is above 50% and zero otherwise. The results are robust for probabilities between 45% and 55%. Table 10 reports the results. First, coefficients on the recession dummy are positive and increase with maturity. Economically, the effect is rather large: a recession increases the annualized expected excess return on a 1-year bond on average by about 70 basis points. The effect on the 5-year bond premium is much larger, but so is the standard error (which is computed using 20 Newey-West lags). The counter-cyclicality is no longer statistically significant for the 10-year bond premium. In Appendix E, we show that this cyclicality result disappears if we use NBER dummies instead.[18] Second, the demand/supply ratio is indeed negatively associated with risk premiums, and especially so for the 5 and 10 year bonds. Again, these effects are economically very large for the longer maturities and highly statistically significant. For example, for the 1 year bond, if the demand/supply ratio were to increase by 2 standard deviations, the annualized bond risk premium would not increase by 68 basis points in a recession, but decrease by 46 (68.44-2×53.77-2×3.65) basis points. Of course, it is important to recall that supply variances spike up as well in most recessions. In Appendix E, we show that, unlike the cyclicality results above, the coefficient on the demand/supply ratio is robustly negative and statistically significant across across different specifications for the recession dummy.

4.3 Macro Risks and Term Premiums

Most of the literature examining the link between the macroeconomy and bond risk premiums has focused on macro level factors. One important exception is Wright (2011), who does not examine excess holding period returns, but an important and closely related component of bond yields, the term premium. Wright (2011) shows that term premiums are countercyclical and strongly affected by inflation uncertainty in a panel of countries.[19] We compute term premiums for the 5 year and 10 year maturity as the yield for each

[18]This is in line with Bauer, Rudebusch and Wu (2014) and Wright (2014), who find that the cyclicality of bond risk premiums is dependent on the methodology.

[19]Bauer, Rudebusch, and Wu (2014) re-examine Wright's empirical evidence correcting for small sample bias in the VAR he runs to compute the term premium, but his main empirical conclusions remain robust.

maturity minus the average of expected future short-term rates over the life of the bond. To measure the expected short yield, we use Blue Chip survey, which is available semi-annually from 1986Q2.[20]

Results from this exercise are reported in Table 11. They are somewhat similar to the results in Table 8 on excess holding period returns. Expected core inflation, expected inflation and expected GDP growth significantly affect term premiums with the same signs as in the excess holding period return regressions. Whereas the bad demand variance risk factor negatively affects the term premium, consistent with the idea that in such an environment bonds act as a good hedge, the effect is statistically insignificant for the 5 year bond and marginally significant for 10 year bond. Instead, increases in the good demand variance significantly decrease term premiums. We also find that the good supply variance affects term premiums positively. The adjusted R^2 is 69 percent for the 5 and 10 year bonds. The macro risk factors' addition to the explanatory power of the macro level variables is marginally significant.

In Table 12, we examine the cyclicality of the term premiums. In line with Wright (2011) and Bauer, Rudebusch, and Wu (2014), we find that the term premium increases in recessions, by 2.54 percentage points (6.25 percentage points) for the 5-year (10-year) bond. These numbers are economically significant but not statistically significant. The term premium is smaller in demand environments, but the effect is also not significant. The interaction effect with the recession dummy has a negative sign but also fails to be significant. The demand environment effects are substantive; a one standard deviation increase in the demand/supply variance ratio decreases the term premium in a recession by about 2.39 percentage points for the 5 year bond and about 4.83 percentage points for the 10-year bond. Therefore, "demand effects" of this magnitude almost completely offset the usual counter-cyclical term premium increase in recessions. However, we find the magnitude of this effect to vary with the the recession proxy. For instance, if NBER recession dummies are used, we find that the term premium increases in recessions by 0.55 percentage points (0.53 percentage points) for the 5-year (10-year) bond. The impact of the demand environment is also smaller: a one standard deviation increase in the demand/supply variance ratio decreases the term premium in a recession by 56 basis

[20] Our results are similar if we employ the expected short yield computed using Bauer, Rudebusch, and Wu (2014)'s small-sample adjusted VAR(1) including 1 quarter, 1 year, and 10 year yields as the state variables. The correlations between the survey and statistical term premia are 0.7578 and 0.7964 for the 5 and 10 year term premia, respectively.

points for the 5 year bond and 52 basis points for the 10-year bond.

4.4 Macro Risks and the Bond Return Variance

Consider a model of the term structure of interest rates in which macroeconomic factors help determine the levels of bond yields (e.g., habit of Wachter, 2006, or long-run risk of Bansal and Shaliastovich, 2013). Then the conditional variance of the macroeconomic factors, which is captured by our macroeconomic risk factors, should help to determine the conditional variance of bond returns. In the context of a forecasting regression, the macro risk factors should help forecast ex-post bond return variances. In Table 13, we present empirical evidence that such a link between the variance of bond returns and the macro risk factors is indeed present in the data. Specifically, we compute the quarterly realized 10 year zero coupon bond return variances as the sum of squared daily returns inside the quarter. The realized daily returns are computed under the assumption that the 10 year-1 day zero coupon yield is equal to the 10 year zero coupon yield. We regress the quarterly realized variance of returns for the 10-year bond on the lagged values of the macro risk factors and/or the macro level and financial factors. In panel A, we report the adjusted R^2 statistics from such regressions. By themselves, the macro risk factors span about 35 percent of the variation in the ex-post realized variance. In contrast, the macro level factors span only about 19 percent, and the financial factors span less than 14 percent. Further, the macro risk factors always significantly add to the explanatory power of regressions which already use the macro level factors or financial factors as explanatory variables. In contrast, the macro level factors do not significantly add to the explanatory power of regressions that already use the macro risk factors and financial factors as explanatory variables, nor do the financial factors significantly add to the explanatory power of regressions that already use the macro risk factors and the macro level factors. We conclude that the macro risk factors are quite powerful predictors of bond return variances.

Panel B shows the pattern of regression coefficients for one such regression that includes macro level factors and macro risk factors as explanatory variables. The most statistically significant explanatory variable is the bad variance component of demand, which positively affects bond return variance, as expected. Moreover, the coefficients for three out of the four macro risk factors are of the expected positive sign. Among the macro level factors, expected aggregate and core inflation are significant at the 10 and 5 percent level, respectively.

Figure 6 shows the historical pattern of the realized bond return variances (the blue line), and the fitted values from two of the forecasting regressions described above. The regression which uses the macro level factors and macro risk factors shown by the red/circle symbols, captures some of the most prominent features of realized variance, especially the high levels seen in the 1980s and during the 2008-2009 financial crisis. As shown by the line with green/triangle symbols, adding the financial factors to this regression does not significantly alter the patterns of the fitted return variance.

We also test the additional predictive power of macro risks for realized variances over two direct variance forecasts based on bond returns data: the lagged value of the realized variance and the conditional variance from the GJR-GARCH model of Glosten, Jagannathan, and Runkle (1993). We use the GJR-GARCH model, because it fits the data better than a standard GARCH model (Bollerslev, 1987) in terms of the Akaike information criterion. We report detailed estimation results in Appendix F.

Table 14 shows that the past realized variance (Panel A) and the GJR-GARCH model implied conditional variance (Panel B) have strong predictive power for future realized variances, producing adjusted R^2's of over 50%. When we put them together (Panel C), the R^2 increases to over 60% and either one contributes significantly to the predictive power of the other, showing they contain independent predictive components. The remainder of the panels shows, as before, how various combinations of the 3 yield curve factors, the macro level factors and the macros risks add to the predictive power. The overwhelming conclusion is that macro risks always significantly increase the adjusted R^2, whatever the included variables are. Using all variables together leads to an adjusted R^2 of about 72%. Macro level factors and yield factors never contribute significantly to the R^2, once macro risks are included. In fact, a regression with only the return based variances and macro risks generates an adjusted R^2 of almost 71%, barely one percent lower than the regression also including macro level and yield factors.

5 Conclusion

In this article, we document empirical links between "macro risks" and the term structure of interest rates. To do so, we first decompose macroeconomic shocks into "demand" shocks which move inflation and GDP growth in the same direction and "supply shocks" which move inflation and GDP growth in opposite directions. The identification relies on unconditional non-Gaussianities in the macro data. We find aggregate demand shocks to be distinctly negatively skewed and leptokurtic, whereas supply shocks unconditionally show little skewness but are also leptokurtic.

We then develop a new dynamic model for real economic activity and inflation, where the shocks are drawn from a Bad Environment - Good Environment model, which accommodates time-varying non-Gaussian features with "good" and "bad" volatility. We extract four macro-risk factors, bad and good volatilities for respectively aggregate demand and supply shocks. Until about the mid-seventies conditional supply variances appear to dominate macroeconomic volatility, while afterwards demand variances are more important until the mid-eighties: afterward there are roughly equal contributions of both. The "good" demand variance has decreased markedly over time, but there is no strong evidence that either "bad" demand variances or supply variances have declined. Importantly, recessions continue to be accompanied by temporarily high bad demand and supply variances. We also find that the conditional correlation between inflation and real activity varies through time with occasional sign switches, as the relative importance of demand and supply risk factors varies over time.

We then link the macro factors extracted from the dynamic macro model, expected GDP growth, the unemployment gap, and expected (core) inflation and the macro risk variables represented by the conditional variances (shape parameters) of the demand and supply shocks, to the term structure. The macro variables explain 79 percent of the variation in the levels of yields. While the contribution of the macro risk factors to this R^2 is modest, it is nonetheless statistically significant. When we run predictive regressions of excess bond returns onto the macro variables, the R^2 is around 6 percent, with the macro risk factors contributing one third of the explanatory power. Our macro risk factors resurrect the statistical importance of macro factors for return predictability regressions. We find that increases in both good and bad aggregate demand variance significantly reduce bond risk premiums; the former also significantly decreases term premiums. Macro risks also significantly predict realized bond return variances. They significantly add explanatory power over financial factors, macro level factors, and, importantly, conditional variance proxies computed from bond returns.

It would be useful to be elucidate how variation in risk premiums is accounted for by the various macro risk factors and to decompose risk premiums into real and inflation components. To accomplish this, a term structure model is necessary. In future work, we plan to build a non-Gaussian term structure model where the set of state variables includes both financial factors (as in Feldhütter, Heyerdahl-Larsen, and Illeditsch, 2018) and macro variables (level and risk factors). Despite the non-Gaussianities in their dynamics, the BEGE structure has the advantage that bond prices nonetheless remain affine in the state variables. Lastly, future work should verify whether the predictive power of our macro risk factors survives when real time data are used instead of the final revised data (see Ghysels, Horan, and Moench, 2018).

This article's methodology can also be fruitfully applied to other empirical regularities. For example, the correlation between stock and bond returns shows extreme time variation, with most economic models falling to account for such variation (see Baele, Bekaert and Inghelbrecht, 2010). However, bond and stock returns should be positively correlated in AS environments, and negatively correlated in AD environments. Ermolov (2018) uses a variant of our methodology and an economic model to show that AS and AD macro risks account for 15% of the time variation in stock-bond return correlations.

References

Ajello, Andrea, Benzoni, Luca, and Olena Chyruk, 2012, Core and “Crust”: Consumer Prices and the Term Structure of Interest Rates, Working paper, Federal Reserve Board and Federal Reserve Bank of Chicago.

Ang, Andrew, Geert Bekaert, and Min Wei, 2007, Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?, *Journal of Monetary Economics*, 4, 2007, pp. 1163-1212

Ang, Andrew, and Monika Piazzesi, 2003, A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables, *Journal of Monetary Economics*, 50, 74587.

Baele, Lieven, Geert Bekaert, and Koen Inghelbrecht, 2010, The Determinants of Stock and Bond Return Comovements, *Review of Financial Studies*, 23, pp. 2374-2428.

Bansal, Ravi, and Ivan Shaliastovich, 2013, A Long-run Risks Explanation of Predictability Puzzles in Bond and Currency Markets, *Review of Financial Studies*, 26, pp. 1-33.

Barillas, Francisco, 2011, Can We Exploit Predictability in Bond Markets?, Working paper, Emory University.

Bates, David, 2006, Maximum Likelihood Estimation of Latent Affine Processes, *Review of Financial Studies*, 19, pp. 909-965.

Bauer, Michael, and James Hamilton, 2017, Robust Bond Risk Premia, *Review of Financial Studies*, forthcoming.

Bauer, Michael, and Glenn Rudebusch, 2016, Resolving the Spanning Puzzle in Macro-Finance Term Structure Models, *Review of Finance*, forthcoming.

Bauer, Michael, and Glenn Rudebusch, 2017, Interest Rates under Falling Stars, Working paper.

Bauer, Michael, Rudebusch, Glenn, and Jing Wu, 2014, Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset: Comment, *American Economic Review*, pp. 323-337.

Bekaert, Geert, Eric Engstrom, and Yuhang Xing, 2009, Risk, Uncertainty, and Asset Prices, *Journal of Financial Economics*, 91, pp. 59-82.

Bekaert, Geert, and Eric Engstrom, 2017, Asset Return Dynamics under Habits and Bad Environment-Good Environment Fundamentals, *Journal of Political Economy*, 125, pp. 713-760.

Bekaert, Geert, Engstrom, Eric, and Andrey Ermolov, 2018, Aggregate Demand and Aggregate Supply Macro Risks, Working paper.

Bikbov, Ruslan, and Mikhail Chernov, 2010, No-arbitrage Macroeconomic Determinants of the Yield Curve, *Journal of Econometrics*, 159, pp. 166-182.

Bils, Mark, Peter Klenow, and Benjamin Malin, 2012, Reset Price Inflation and the Impact of Monetary Policy Shocks, *American Economic Review*, 102, pp. 2798-2825.

Blanchard, Olivier, 1989, A Traditional Interpretation of Macroeconomic Fluctuations, *American Economic Review*, 79, pp. 1146-1164.

Blanchard, Olivier, and Danny Quah, 1989, The Dynamic Effects of Aggregate Demand and Supply Disturbances, *American Economic Review*, 79, pp. 655-673.

Bollerslev, Tim, 1987, A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return, *Review of Economics and Statistics*, 69, pp. 542-547.

Boons, Martijn, Duarte, Fernando, de Roon, Frans, and Marta Szymanowska, 2017, Time-varying

Inflation Risk and the Cross Section of Stock Returns, Working paper, New University of Lisbon, Federal Reserve Bank of New York, Tilburg University, and Erasmus University of Rotterdam.

Bureau of Economic Analysis, 1993, Reliability and Accuracy of the Quarterly Estimates of GDP, Survey of Consumer Business, October 1993.

Buraschi, Andrea, and Alexei Jiltsov, 2007, Habit Formation and Macroeconomic Models of the Term Structure of Interest Rates, *Journal of Finance*, 66, pp. 3009-3063.

Burnham, Kenneth, and David Anderson, 2004, Multimodel Inference Understanding AIC and BIC in Model Selection, *Sociological Methods & Research*, 33, pp. 261-304.

Campbell, John, Carolin Pflueger, and Luis Viceira, 2015, Monetary Policy Drivers of Bond and Equity Risks, Working paper, Harvard University and University of British Columbia.

Campbell, John, Adi Sunderam, and Luis Viceira, 2017, Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds, *Critical Finance Review*, 6, pp. 263-301.

Cieslak, Anna, and Pavol Povala, 2015, Expected returns in Treasury bonds, *Review of Financial Studies*, 28, pp. 2859-2901.

Cochrane, John, and Monika Piazzesi, 2005, Bond risk premia, *American Economic Review*, 95,138-160.

Collin-Dufresne, Pierre and Robert Goldstein, 2002, Do Bonds Span the Fixed Income Markets? Theory and Evidence for Unspanned Stochastic Volatility, *Journal of Finance*, 57, pp. 1685-1730.

Cooper, Ilan, and Richard Priestley, 2009, Time-varying Risk Premiums and the Output Gap, *Review of Financial Studies*, 22, pp. 2801-2833.

Duffee, Gregory, 2013, Bond Pricing and the Macroeconomy, *Handbook of the Economics and Finance*, 2, pp. 907-967.

Dufour, Jean-Marie, and Denis Pelletier, 2014, Practical Methods for Modelling Weak VARMA Processes: Identification, Estimation and Specification with a Macroeconomic Application, Working paper, McGill University and North Carolina State University.

Ehling, Paul, Gallmeyer, Michael, Heyerdahl-Larsen, Christian, and Philipp Illeditsch, 2018, Disagreement about Inflation and the Yield Curve, *Journal of Financial Economics*, 127, pp. 459-484.

Engle, Robert, Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models, *Journal of Business and Economic Statistics*, 2002, 20, pp. 339-350.

Engle, Robert, Ghysels, Eric, and Bumjean Sohn, 2013, Stock Market Volatility and Macroeconomic Fundamentals, *Review of Economics and Statistics*, 95, pp. 776-797.

Ermolov, Andrey, 2018, Time-varying Risk of Nominal Bonds: How Important Are Macroeconomic Shocks, Working paper, Fordham University.

Evans, Martin, and Paul Wachtel, 1993, Were Price Changes during the Great Depression Anticipated?: Evidence from Nominal Interest Rates, *Journal of Monetary Economics*, 32, pp. 3-34.

Fama, Eugene, and William Schwert, 1977, Asset Returns and Inflation, *Journal of Financial Economics*, 5, pp. 115-146.

Fama, Eugene, 1981, Stock Returns, Real Activity, Inflation, and Money, *American Economic Review*, 71, pp. 545-565.

Fama, Eugene, and Robert Bliss, 1987, The Information in Long-Maturity Forward Rates, *American Economic Review*, 77, pp. 680-692.

Feldhütter, Peter, Heyerdahl-Larsen, Christian, and Philipp Illeditsch, 2018, Risk Premia and Volatilities in a Nonlinear Term Structure Model, *Review of Finance*, 22, pp. 337-380.

Gali, Jordi, 1992, How Well Does the IS-LM Model Fit Postwar U.S. Data?, *Quarterly Journal of Economics*, 107, pp. 709-738.

Ghysels, Eric, Horan, Casidhe, and Emanuel Moench, 2018, Forecasting through the Rearview Mirror: Data Revisions and Bond Return Predictability, *Review of Financial Studies*, 31, pp. 678-714.

Glosten, Lawrence, Jagannathan, Ravi, and David Runkle, 1993, On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance*, 48, pp. 1779-1801.

Gourieroux, Christian, and Joann Jasiak, 2006, Autoregressive Gamma Processes, *Journal of Forecasting*, 25, pp. 129-152.

Gürkaynak, Refet, Brian Sack, and Jonathan Wright, 2010, The U.S. Treasury Yield Curve: 1961 to the Present, Working paper, Federal Reserve Board.

Gürkaynak, Refet, and Jonathan Wright, 2012, Macroeconomics and the Term Structure, *Journal of Economic Literature*, 50, pp. 331-67.

Hamilton, James, 1989, A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica*, 57, pp. 357-384.

Hannan, Edward, and Jorma Rissanen, 1982, Recursive Estimation of Mixed Autoregressive-Moving Average Order, *Biometrika*, 69, pp. 81-94.

Harvey, Campbell, 1988, The Real Term Structure and Consumption Growth, *Journal of Financial Economics*, 22, pp. 305-333.

Hasseltoft, Henrik, and Dominic Burkhardt, 2012, Understanding Asset Correlations, Working paper, University of Zurich and Swiss Finance Institute.

Hong, Harrison, Sraer, David, and Jialin Yu, 2017, Inflation Bets on the Long Bond, *Review of Financial Studies*, 30, pp. 900-947.

Ireland, Peter, 2011, A New Keynesian Perspecitve on the Great Recession, *Journal of Money, Credit and Banking*, 43, pp. 31-54.

Jorda, Oscar, 2005, Estimation and Inference of Impulse Responses by Local Projections, *American Economic Review*, 95, pp. 161-182.

Joslin, Scott, Priebsch, Marcel, and Kenneth Singleton, 2014, Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks, *Journal of Finance*, 69, pp. 1197-1233.

Joslin, Scott, 2018, Can Unspanned Stochastic Volatility Models Explain the Cross Section of Bond Volatilities?, *Management Science*, 64, pp. 1477-1973.

Joslin, Scott and Yaniv Konchitchki, 2018, Interest Rate Volatility, the Yield Curve, and the Macroeconomy, *Journal of Financial Economics*, 12, pp. 344-362.

Kaul, Gautam, 1987, Stock Returns and Inflation: The Role of the Monetary Sector, Working paper, University of Michigan.

Lanne, Markku, Meitz, Miika, and Pentti Saikkonen, 2017, Identification and Estimation of Non-Gaussian Structural Vector Autoregressions, *Journal of Econometrics*, 196, pp. 288-304.

Ludvigson, Sydney, and Serena Ng, 2009, Macro Factors in Bond Risk Premia, *Review of Financial Studies*, 22, pp. 5027-5067.

McCulloch, Huston, and Heon-Chul Kwon, 1993, U.S. Term Structure Data, 1947-1991, Working paper, Ohio State University.

Mian, Atif, and Amir Sufi, 2014, What Explains the 2007-2009 Drop in Employment?, *Econometrica*, 82, pp. 2197-2223.

Mulligan, Casey, 2012, The Redistribution Recession: How Labor Market Distortions Contracted the Economy, Oxford University Press.

Newey, Whitney K., and Kenneth D. West, 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica*, 55, 703-08.

Piazzesi, Monika, and Eric Swanson, 2008, Futures Prices as Risk-adjusted Forecasts of Monetary Policy, *Journal of Monetary Economics*, pp. 677-691.

Shapiro, Matthew, and Mark Watson, 1988, Sources of Business Cylcle Fluctuations, *NBER Macroeconomics Annual*, 3, pp. 111-156.

Song, Dongho, 2017, Bond Market Exposures to Macroeconomic and Monetary Policy Risks, *Review of Financial Studies*, 30, pp. 2761-2817.

Sugiura, Nariaki, 1978, Further Analysts of the Data by Akaike' s Information Criterion and the Finite Corrections, 1, pp. 13-26.

Wachter, Jessica, 2006, A Consumption-based Model of the Term Structure of Interest Rates, *Journal of Financial Economics*, 79, pp. 365-399.

Wilcox, David, 1992, The Construction of U.S. Consumption Data: Some Facts and Their Implications for Empirical Work, 82, pp. 922-941.

Woodford, Michael, 2003, Interest and Prices: Foundations of a Theory of Monetary Policy, Princeton University Press.

Wooldridge, Jeffrey, 2002, Econometric Analysis of Cross Section and Panel Data, MIT Press.

Wright, Jonathan, 2011, Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset, *American Economic Review*, pp. 1514-1534.

Wright, Jonathan, 2014, Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset: Reply, *American Economic Review*, 104, pp. 338-341.

Figure 1 – Demand and Supply Variance Proxies. Data is quarterly. NBER recessions are shaded.

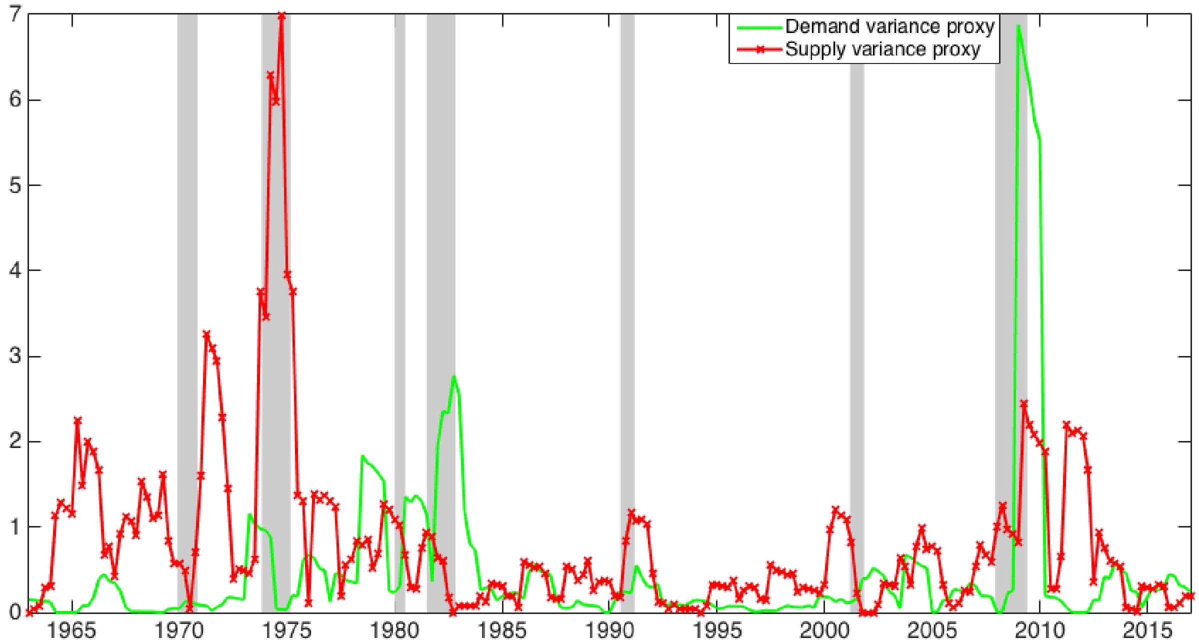

Figure 2 – Bad Environment - Good Environment Distribution. Graphs are probability density functions.

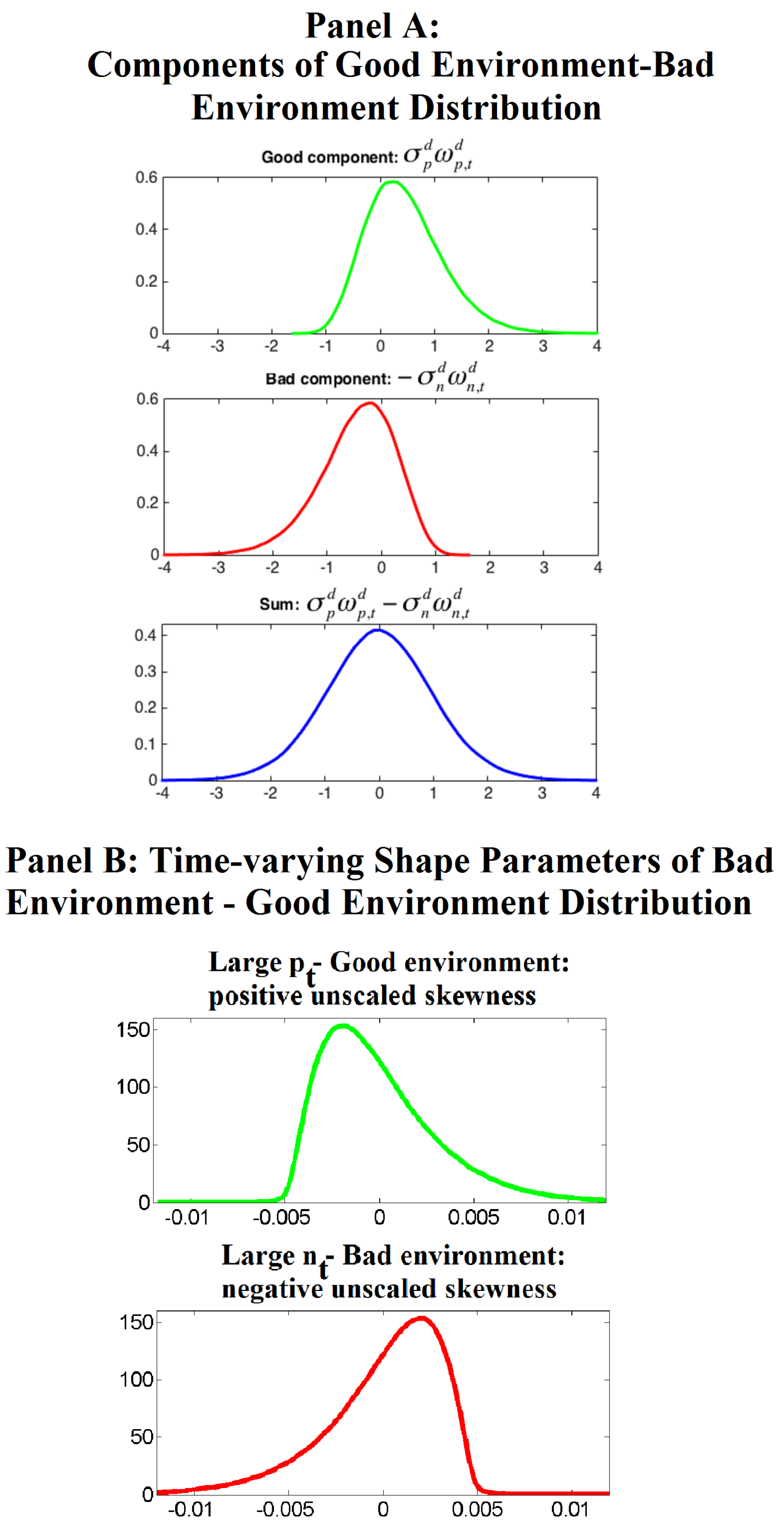

Figure 3 – Filtered Quarterly Demand and Supply Shocks. Shading corresponds to NBER Recessions.

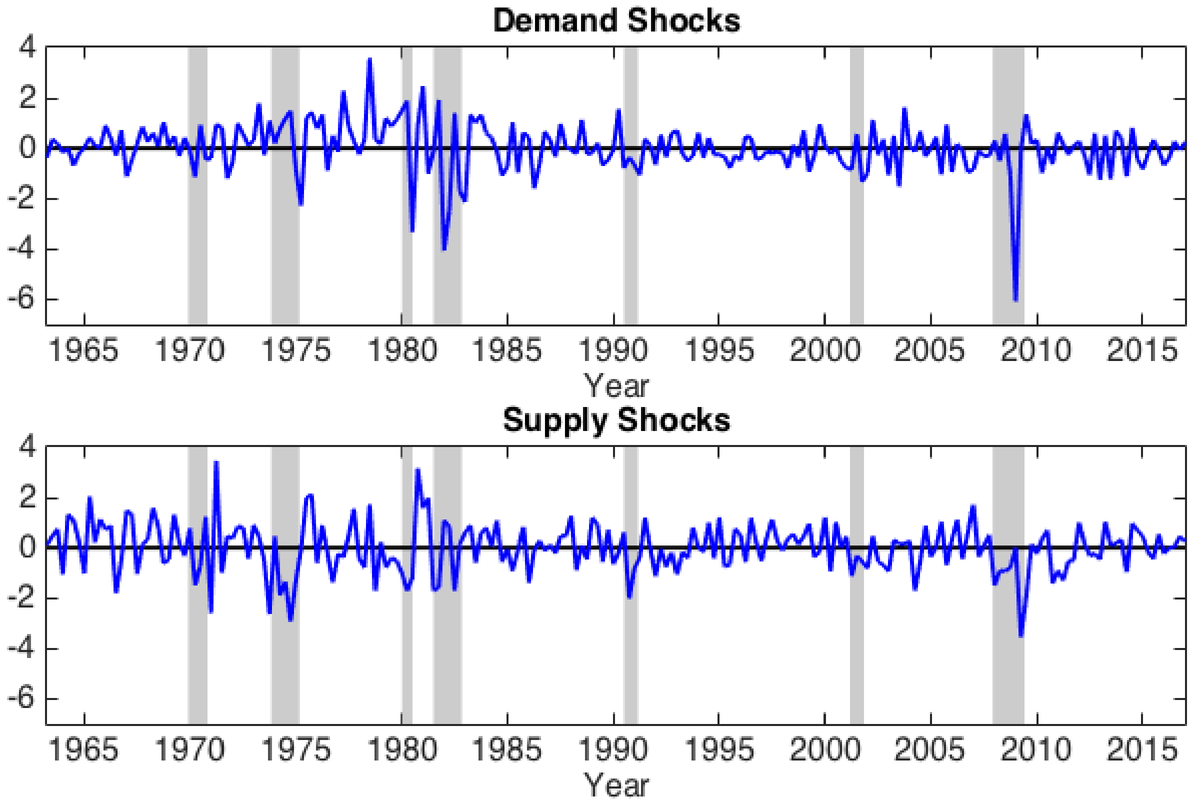

Figure 4 – Filtered Quarterly Demand and Supply Variances. Good demand variance is defined as the constant $(\sigma_p^d)^2$ multiplied by the good demand macro risk factor p_t^d. Bad demand variance is defined as the constant $(\sigma_n^d)^2$ multiplied by the bad demand macro risk factor n_t^d. Good supply variance is defined as the constant $(\sigma_p^s)^2$ multiplied by the good supply macro risk factor p_t^s. Bad supply variance is defined as the constant $(\sigma_n^s)^2$ multiplied by the bad demand macro risk factor n_t^s. Shading corresponds to NBER Recessions.

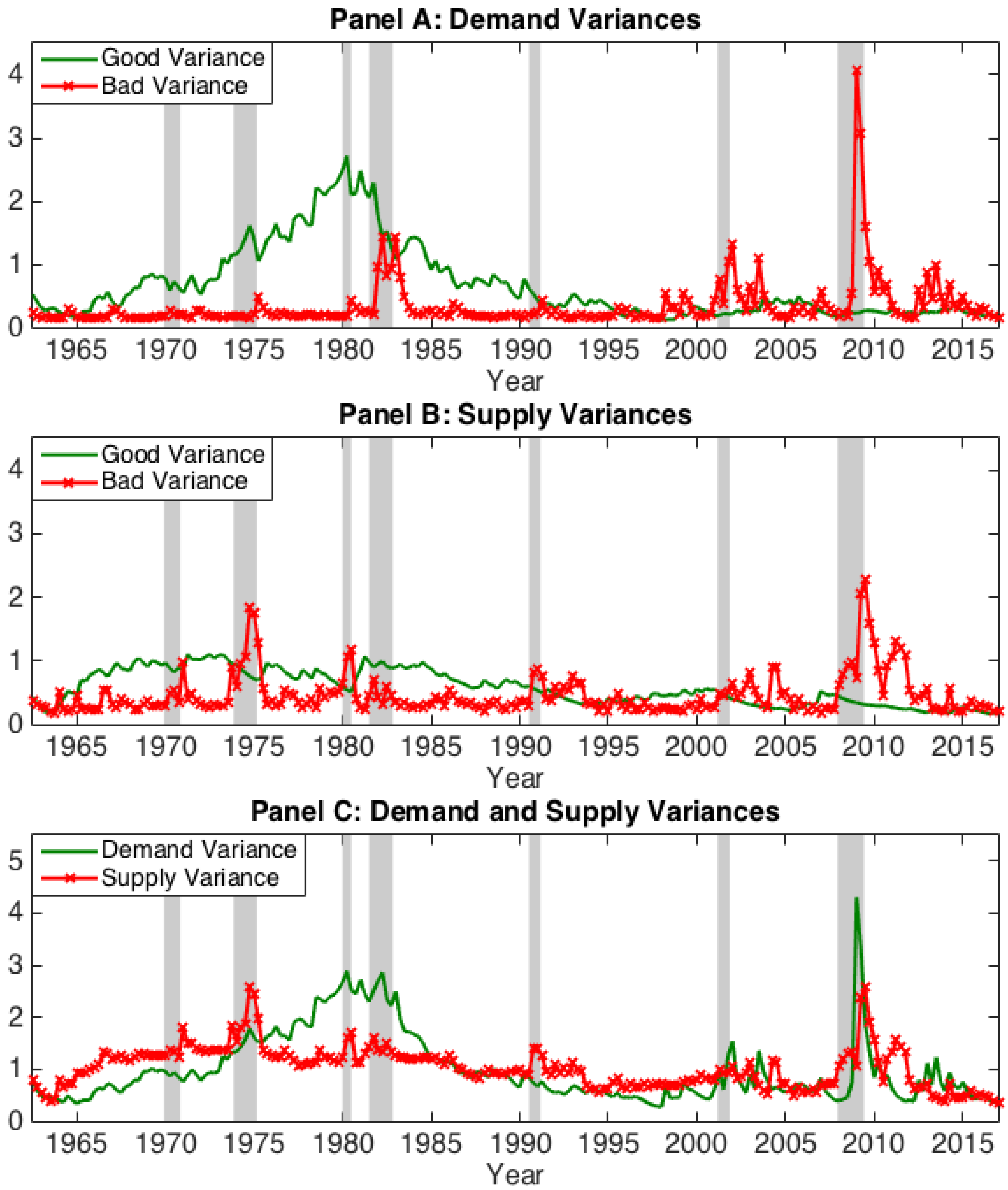

Figure 5 – Quarterly Conditional Covariance between Macroeconomic Variables. Shading corresponds to NBER Recessions.

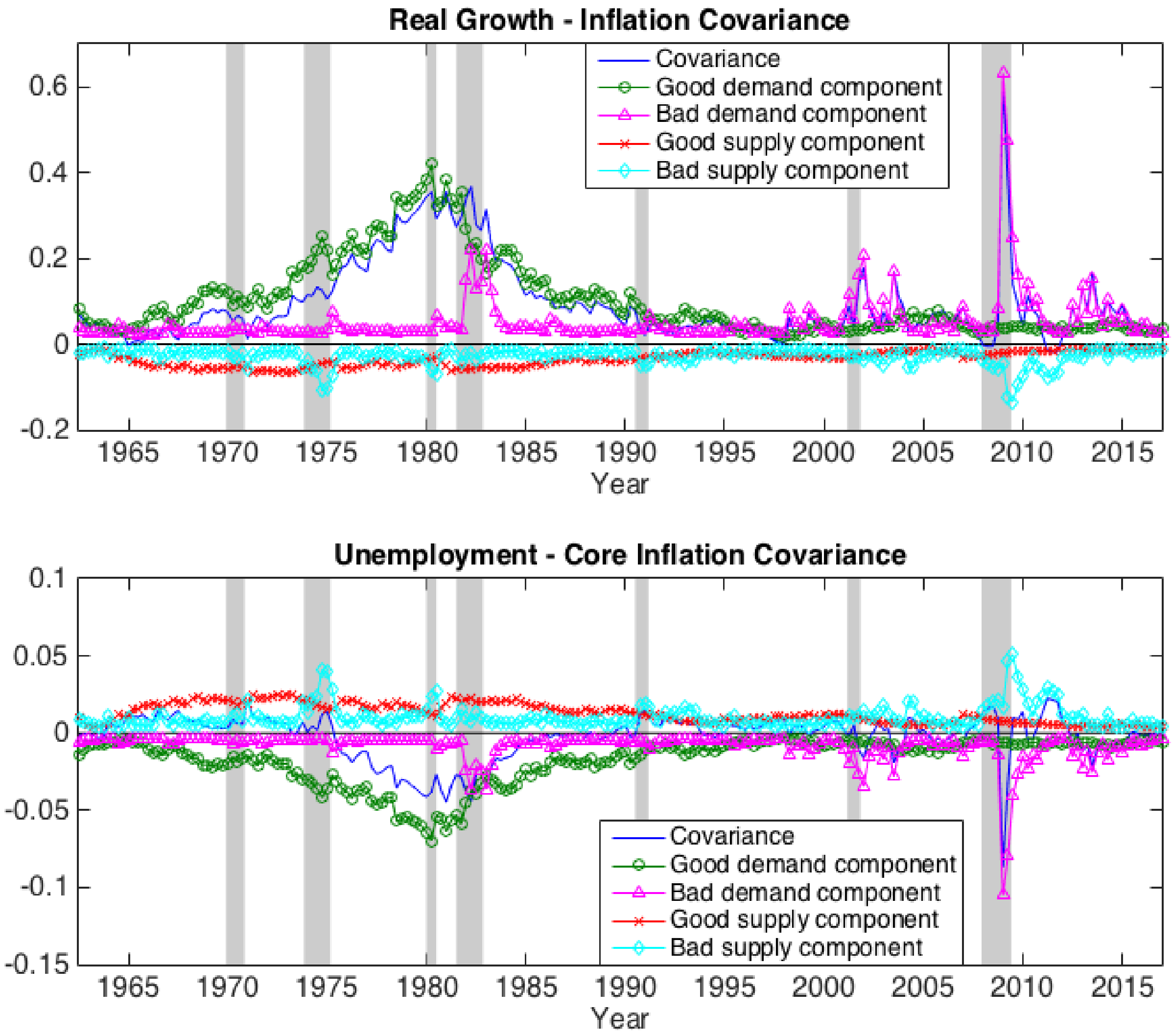

Figure 6 – Explaining Realized 10 Year Bond Return Variance with Macroeconomic and Financial Factors. Realized variances are computed as the sums of squared daily bond returns inside the quarter. The fit is from OLS regressions. Financial factors are the level, slope, and curvature factors. The level factor is the average over 1-10 year yields. The slope factor is the 10 year yield minus the 1 quarter yield. The curvature factor is 10 year yield plus 1 quarter yield minus 2 times the 2 year yield. The macroeconomic level factors are expected inflation, expected core inflation, expected GDP growth and unemployment gap.

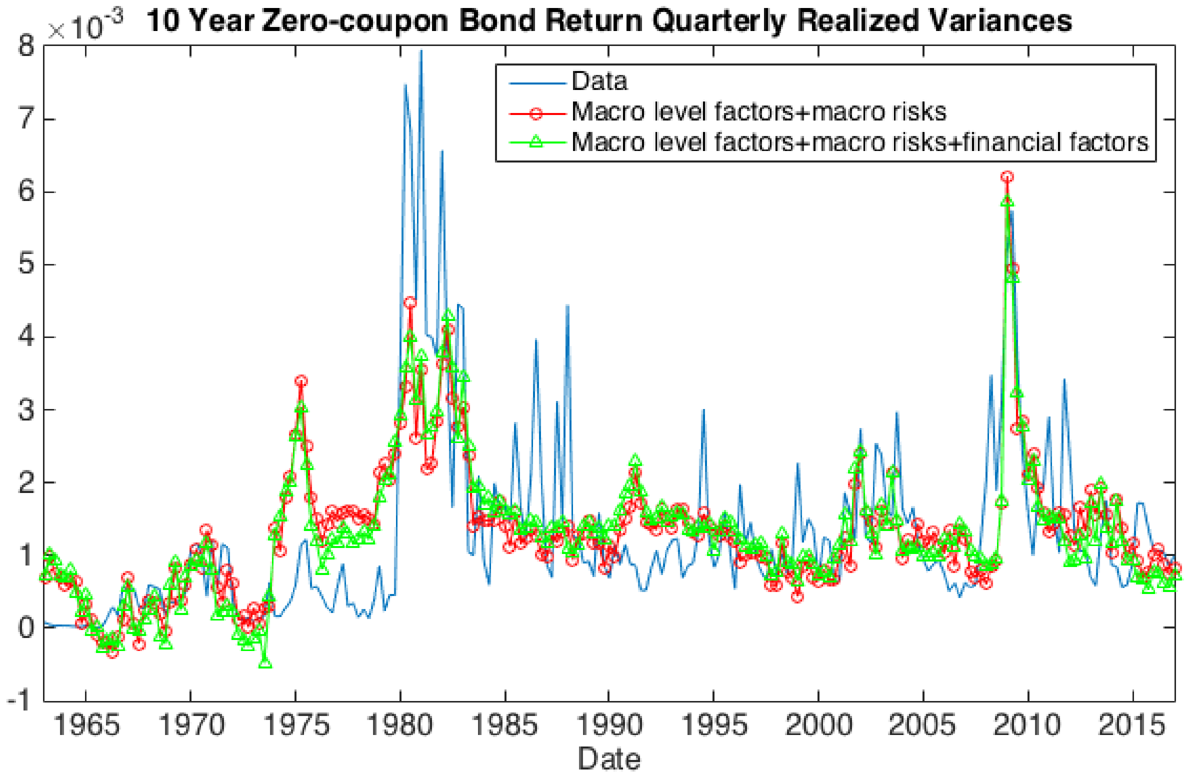

Table 1 – Explaining Quarterly Excess Bond Returns with Macro Factors: Partial Results. The sample is quarterly 1962Q4 to 2016Q4. The excess returns are annualized 1 quarter holding period returns on zero coupon US Treasuries. $E_t\pi_{t+1}^{core}$ is the expected core inflation. $E_t\pi_{t+1}$ is the expected aggregate inflation. E_tg_{t+1} is the expected GDP growth. $ugap_t$ is the unemployment gap. Macro risks (demand proxy$_t$ and supply proxy$_t$) are scaled to have unit variance. All regressions also include a constant and expected aggregate inflation, expected core inflation, expected GDP growth and unemployment gap as control variables. The value in parentheses is the proportion out of 10,000 Bauer and Hamilton (2017) bootstrap runs where the t-stat for the coefficient is smaller than in data. The asterisks, * and **, correspond to statistical significance at the 10 and 5 percent levels, respectively.

	1 year bond	2 year bond	5 year bond	10 year bond
Demand proxy$_t$	-0.3455*	-0.8139**	-1.5945**	-3.5636**
	(0.0378)	(0.0207)	(0.0114)	(0.0079)
Supply proxy$_t$	0.1580	0.1569	0.3963	0.8010
	(0.5444)	(0.5393)	(0.6261)	(0.5795)
Adjusted R^2	0.0498	0.0541	0.0537	0.0598

Table 2 – Model Selection for Expectations of Macro Variables. The sample is quarterly from 1962Q4 to 2016Q4. Dependent variables are the log-difference of the CPI, log real GDP growth, the log difference of core CPI, and the unemployment rate gap. The predictive variables are the macro variables mentioned above and the 90-day T-bill and the 10-year zero-coupon Treasury yield. AIC and BIC are Akaike and Bayesian information criteria, respectively. The models are sorted by AIC.

Model	Number of parameters	Log-likelihood	AIC	BIC
VAR(2)	93	-798.9	1801.8	2097.7
VAR(3)	129	-755.5	1802.8	2204.5
VARMA(2,2)	105	-785.9	1804.5	2136.3
VARMA(2,1)	99	-794.7	1807.6	2121.5
VARMA(3,1)	135	-752.9	1812.7	2231.4
VARMA(2,3)	111	-787.2	1821.5	2171.0
VARMA(3,2)	141	-749.9	1822.1	2257.7
VARMA(3,3)	147	-743.8	1825.5	2277.9
VARMA(1,3)	75	-856.8	1875.6	2116.8
VARMA(1,1)	63	-879.5	1893.8	2097.7
VAR(1)	57	-888.6	1898.4	2083.5
VARMA(1,2)	69	-875.7	1899.7	2122.3

Table 3 – Higher Order Moments of Macroeconomic Shocks Used for Classical Minimum Distance Estimation. u_t^g, u_t^π, $u_t^{\pi^{core}}$, and u_t^u are the shocks to real GDP growth, aggregate inflation, core inflation and unemployment gap, respectively. The data is quarterly from 1962Q4 to 2015Q2. The covariance matrix for moments is a diagonal matrix calculated via a block-bootstrap with a block length of 20 quarters. Asterisks, *, **, and ***, correspond to statistical significance of individual moments at the 10, 5, and 1 percent levels, respectively.

			Volatility			
	u_t^π	u_t^g	$u_t^{\pi^c}$	u_t^{ue}		
data	0.5655***	0.7078***	0.3252***	0.2658***		
standard error	(0.0867)	(0.0781)	(0.0531)	(0.0228)		
fitted	0.5655	0.7078	0.3252	0.2658		
			Skewness			
	u_t^π	u_t^g	$u_t^{\pi^c}$	u_t^{ue}		
data	-1.3570	0.4956	0.1144	0.3745**		
standard error	(1.0067)	(0.3714)	(0.3808)	(0.1879)		
fitted	-0.4456	-0.2585	-0.2264	0.2308		
			Excess kurtosis			
	u_t^π	u_t^g	$u_t^{\pi^c}$	u_t^{ue}		
data	11.2751**	2.5052**	2.0640**	1.0528***		
standard error	(5.7197)	(1.0656)	(0.8233)	(0.4056)		
fitted	1.9051	1.1046	0.9798	1.0160		
			Correlations			
	$u_t^\pi u_t^g$	$u_t^\pi u_t^{\pi^c}$	$u_t^\pi u_t^{ue}$	$u_t^g u_t^{\pi^c}$	$u_t^g u_t^{ue}$	$u_t^{\pi^c} u_t^{ue}$
data	0.1392	0.5400***	-0.2058***	0.0626	-0.5615***	-0.1630*
standard error	(0.1197)	(0.0726)	(0.0733)	(0.1281)	(0.0534)	(0.0969)
fitted	0.2415	0.5274	-0.2204	0.0604	-0.5587	-0.0372
			Co-skewness			
	$(u_t^\pi)^2 u_t^g$	$(u_t^\pi)^2 u_t^{\pi^c}$	$(u_t^\pi)^2 u_t^{ue}$	$(u_t^g)^2 u_t^\pi$	$(u_t^g)^2 u_t^{\pi^c}$	$(u_t^g)^2 u_t^{ue}$
data	-0.9790*	-0.4251	0.9978*	-0.2876	-0.1337	-0.1683
standard error	(0.5588)	(0.3519)	(0.5623)	(0.3977)	(0.2386)	(0.3941)
fitted	-0.3714	-0.3544	0.3579	-0.3144	-0.2514	0.2489
	$(u_t^{\pi^c})^2 u_t^\pi$	$(u_t^{\pi^c})^2 u_t^g$	$(u_t^{\pi^c})^2 u_t^{ue}$	$(u_t^{ue})^2 u_t^\pi$	$(u_t^{ue})^2 u_t^g$	$(u_t^{ue})^2 u_t^{\pi^{core}}$
data	-0.0814	-0.2427	0.2308	-0.4526*	-0.0987	-0.2621**
standard error	(0.2620)	(0.1813)	(0.1901)	(0.2513)	(0.3258)	(0.1180)
fitted	-0.2826	-0.2311	0.2225	-0.2926	-0.2397	-0.2342
			Excess co-kurtosis			
	$(u_t^\pi)^2(u_t^g)^2$	$(u_t^\pi)^2(u_t^{\pi^c})^2$	$(u_t^\pi)^2(u_t^{ue})^2$	$(u_t^g)^2(u_t^{\pi^c})^2$	$(u_t^g)^2(u_t^{ue})^2$	$(u_t^{\pi^c})^2(u_t^{ue})^2$
data	2.8288*	0.9001**	2.5459	0.8804***	1.1683**	0.7172**
standard error	(1.7353)	(0.4307)	(1.7067)	(0.2841)	(0.5452)	(0.2931)
fitted	1.3899	1.2650	1.3041	1.0355	1.0571	0.9972
	$(u_t^\pi)^3 u_t^g$	$(u_t^\pi)^3 u_t^{\pi^c}$	$(u_t^\pi)^3 u_t^{ue}$	$(u_t^g)^3 u_t^\pi$	$(u_t^g)^3 u_t^{\pi^c}$	$(u_t^g)^3 u_t^{ue}$
data	5.4690*	2.3743	-5.3776*	1.6048*	0.9830	-1.6559*
standard error	(3.3311)	(1.6502)	(3.1267)	(0.9644)	(0.7055)	(0.6289)
fitted	1.5255	1.5383	-1.4667	0.9894	0.6839	-1.0801
	$(u_t^{\pi^c})^3 u_t^\pi$	$(u_t^{\pi^c})^3 u_t^g$	$(u_t^{\pi^c})^3 u_t^{ue}$	$(u_t^{ue})^3 u_t^\pi$	$(u_t^{ue})^3 u_t^g$	$(u_t^{ue})^3 u_t^{\pi^c}$
data	1.0483	0.5848	-0.7485**	-1.1668	-0.9086*	-0.3166
standard error	(0.4346)	(0.5241)	(0.3655)	(0.7724)	(0.5445)	(0.2325)
fitted	1.0780	0.5661	-0.5272	-0.8572	-1.0357	-0.5635
J-stat	29.6525					
p-value	(0.0819)					
Joint significance of 3^{rd} and 4^{th} order moments	299.43					
p-value	(<0.0001)					

Table 4 – Loadings of Macroeconomic Shocks on Demand and Supply Shocks. The coefficients are from Classical Minimum Distance estimation matching unconditional higher order moments of 4 macroeconomic shocks time series: real GDP growth (u_t^g), aggregate (u_t^π) and core inflation ($u_t^{\pi^{core}}$) and unemployment gap (u_t^u). Standard errors in parentheses account for sampling error in the higher-order moments and the VAR(2) parameters.

Panel A: Loadings of Macro Shocks on Supply and Demand Shocks				
Shock	Supply loading	Demand Loading		
u_t^π	-0.1736	0.3856		
	(0.0555)	(0.1012)		
u_t^g	0.3414	0.4044		
	(0.0888)	(0.0950)		
$u_t^{\pi^c}$	-0.1678	0.1760		
	(0.0438)	(0.0678)		
u_t^{ue}	-0.1344	-0.1464		
	(0.0334)	(0.0264)		
idiosyncratic variance share	0.4408			
	(0.0473)			
Panel B: Kalman Gain of Macro Shocks for Supply and Demand				
Shock	u_t^π	u_t^g	$u_t^{\pi^c}$	u_t^{ue}
Supply	-0.4553	0.5069	-1.2790	-1.4202
	(0.1744)	(0.1038)	(0.3453)	(0.2772)
Demand	0.6758	0.4233	0.9561	-1.0825
	(0.1312)	(0.1066)	(0.2000)	(0.2848)
Panel C: Unconditional moments of supply and demand				
Shock	Skewness	Excess Kurtosis		
Supply	0.0289	3.3186		
	(0.8770)	(1.7417)		
Demand	-1.4030	8.6770		
	(0.9987)	(4.8979)		

Table 5 – Bad Environment - Good Environment Parameter Estimates for Demand and Supply Processes. Parameter estimates are obtained using Bates (2006) approximate maximum likelihood methodology. Standard errors in parentheses are approximate maximum likelihood asymptotic standard errors. As demand and supply shocks are assumed to have variances exactly equal to 1, $\bar{n}$-parameters can be solved as functions of other model parameters, and their standard errors are calculated using the delta method.

	Supply shock	Demand shock
$\bar{p}$	20.0000	20.0000
	–	–
$\bar{n}$	4.0030	0.3359
	(7.1293)	(0.2177)
σ_p	0.1644	0.1801
	(0.0193)	(0.0107)
σ_n	0.3389	1.0229
	(0.2879)	(0.3271)
ρ_p	0.9881	0.9392
	(0.0177)	(0.0279)
ρ_n	0.6737	0.7243
	(0.2046)	(0.1551)
σ_{pp}	0.5524	0.9834
	(0.4162)	(0.3434)
σ_{nn}	1.2502	0.5723
	(1.1114)	(0.3905)

Table 6 – Monte Carlo Estimates of Macroeconomic Shocks Loadings on Demand and Supply Shocks. The estimates use 10,000 replications of simulated samples of historical length.

	Data estimate	Mean Monte Carlo estimate	Median Monte Carlo estimate
Supply loading of u_t^{π}	-0.17	-0.20	-0.19
Demand loading of u_t^{π}	0.39	0.35	0.36
Supply loading of u_t^{g}	0.34	0.29	0.33
Demand loading of u_t^{g}	0.40	0.40	0.41
Supply loading of $u_t^{\pi^c}$	-0.17	-0.17	-0.18
Demand loading of $u_t^{\pi^c}$	0.18	0.15	0.16
Supply loading of u_t^{ue}	-0.13	-0.12	-0.13
Demand loading of u_t^{ue}	-0.15	-0.15	-0.15

Table 7 – Explanatory Power (Adjusted R^2) of Macro Risk Factors for Yield Curve Factors. The sample is quarterly from 1962Q4 to 2016Q4. Macro level factors are expected real GDP growth, expected aggregate and core inflation, and unemployment gap. Financial factors are the level, slope, and curvature factors. The level factor is the average over 1-10 year yields. The slope factor is the 10 year yield minus the 1 quarter yield. The curvature factor is 10 year yield plus 1 quarter yield minus 2 times the 2 year yield. The increase in adjusted R^2 significance, which is always tested over the specification in the previous row, is Bauer and Hamilton (2017) small-sample adjusted significance using 5000 bootstrap runs. The asterisks, *, **, and ***, correspond to statistical significance at the 10, 5, and 1 percent levels, respectively.

	Level	Slope	Curvature
Macro level factors	0.7146	0.5713	0.2808
Macro level factors+macro risks	0.7902***	0.5975*	0.4072***

Table 8 – Explaining Quarterly Excess Bond Returns with Macro Factors. The sample is quarterly from 1962Q4 to 2016Q4. The excess returns are annualized 1 quarter holding period returns on zero coupon US Treasuries. Macro risks (p_t^d, n_t^d, p_t^s and n_t^s) are scaled to have unit variance. The value in parentheses is the proportion out of 5,000 Bauer and Hamilton (2017) bootstrap runs where the t-stat for the coefficient is smaller than in data. The asterisks, * , **, and *** correspond to statistical significance at the 10, 5, and 1 percent levels, respectively.

	1 year bond	2 year bond	5 year bond	10 year bond
Constant	0.0533	0.7436	2.3547	5.1106
	(0.0698)	(0.3742)	(0.5058)	(0.5942)
$E_t\pi_{t+1}^{core}$	5.5115	11.6445	22.5331	38.2388
	(0.6818)	(0.6766)	(0.7170)	(0.7468)
$E_t\pi_{t+1}$	-5.1162***	-11.0131***	-21.7026***	-36.4865***
	(0.0014)	(0.0016)	(0.0016)	(0.0018)
$E_t g_{t+1}$	0.7092	0.9958	3.0505	7.5204
	(0.6442)	(0.5416)	(0.5672)	(0.5918)
$ugap_t$	0.2131	0.5477	1.2754	2.1034
	(0.6058)	(0.6228)	(0.7056)	(0.6346)
p_t^d	-0.8742***	-1.5057***	-3.1487***	-5.2105***
	(0.0020)	(0.0014)	(0.0014)	(0.0016)
n_t^d	-0.2270***	-0.6327***	-1.6587***	-3.3794***
	(0.0008)	(0.0010)	(0.0010)	(0.0008)
p_t^s	0.3998	0.5255	0.8686	0.7653
	(0.8600)	(0.6622)	(0.5794)	(0.3338)
n_t^s	0.3359	0.6965	1.4538	2.9693*
	(0.8668)	(0.8844)	(0.9296)	(0.9514)
Adjusted R^2 without macro risks	0.0416	0.0475	0.0471	0.0469
Adjusted R^2 with macro risks	0.0604	0.0610	0.0613	0.0685

Table 9 – Explanatory Power (Adjusted R^2) of Macro Risk Factors for Quarterly Excess Bond Returns over Macro Level and Financial Factors. The sample is quarterly from 1962Q4 to 2016Q4. Macro level factors are expected real GDP growth, expected aggregate and core inflation, and unemployment gap. Financial factors are the level, slope, and curvature factors. The level factor is the average over 1-10 year yields. The slope factor is the 10 year yield minus the 1 quarter yield. The curvature factor is the 10 year yield plus the 1 quarter yield minus 2 times the 2 year yield. The increase in adjusted R^2 significance, which is tested over the specification without the last set of factors (e.g., "3 financial factors+macro level factors+macro risks" row tests the incremental contribution of macro risks for the specification already including 3 financial factors and macro level factors), is Bauer and Hamilton (2017) adjusted significance using 5000 bootstrap runs. The asterisks, *, **, and ***, correspond to statistical significance at the 10, 5, and 1 percent levels, respectively.

	1 year bond	2 year bond	5 year bond	10 year bond
3 financial factors	0.0666	0.0657	0.0708	0.0796
3 financial factors+macro level factors	0.0962*	0.0932*	0.0774	0.0749
3 financial factors+macro risks	0.1338***	0.1292***	0.1101**	0.1164*
3 financial factors+macro level factors+macro risks	0.1429**	0.1370**	0.1065*	0.1051*
3 financial factors+macro risks+macro level factors	0.1429	0.1370	0.1065	0.1051

Table 10 – Cyclicality of Expected Excess Bond Returns. The sample is quarterly 1969Q4-2016Q4. The dependent variable is the expected annualized quarterly excess return computed from the OLS regressions of realized annualized quarterly excess returns on 4 macro level factors (expected aggregate and core inflations, expected real GDP growth, and unemployment gap) and 4 macro risks (good and bad demand and supply variances). The recession dummy is set to 1 if the anxious index for the quarter (as surveyed during that quarter) is above 50% and 0 otherwise. Demand/supply-ratio is the ratio of aggregate demand variance (good+bad) to aggregate supply variance (good+bad). Demand/supply-ratio is scaled to have the standard deviation of 1. Standard errors are Newey-West standard errors computed with 20 lags. The asterisks, ** and ***, correspond to statistical significance at the 5 and 1 percent levels, respectively.

	1 year bond	5 year bond	10 year bond
constant	1.2227***	6.2253***	13.1634***
	(0.2579)	(0.9967)	(2.1864)
recession-dummy	0.6844**	4.4037***	1.5002
	(0.3471)	(1.5038)	(3.6225)
demand-supply ratio	-0.5377***	-2.4920***	-5.2721***
	(0.1444)	(0.4925)	(0.9748)
recession-dummy×demand-supply ratio	-0.0365	-0.7412	0.2911
	(0.1571)	(0.5901)	(1.3283)
Adjusted R^2	0.3655	0.4086	0.3596

Table 11 – Explanatory Power (Adjusted R^2) of Macro Factors for Term Premiums. The dependent variable is annualized term premium computed as the observed US Treasury long yield minus the expected 1 quarter US Treasury yield over the life of the long yield. The expectations of 1 quarter yield over the life of the long yield are from Blue Chip survey and are available semi-annually. The sample is 1986Q2-2016Q4. The standard deviation of each macro risk factor is scaled to 1. The value in parentheses is the proportion out of 5,000 Bauer and Hamilton (2017) bootstrap runs where the t-stat for the coefficient is smaller than in data. The significance of the increase in adjusted R^2 is computed using 5,000 bootstrap runs of Bauer and Hamilton (2017) bootstrap. The asterisks, * , **, and *** correspond to statistical significance at the 10, 5, and 1 percent levels, respectively.

	5 year bond	10 year bond
constant	0.1852	0.5254*
	(0.9370)	(0.9604)
$E_t\pi_{t+1}^{core}$	6.7811***	8.0065***
	(0.9978)	(0.9994)
$E_t\pi_{t+1}$	-5.0956***	-6.5618***
	(0.0026)	(0.0008)
E_tg_{t+1}	0.8876*	1.0378*
	(0.9720)	(0.9608)
$ugap_t$	0.0769	0.1164
	(0.5100)	(0.6018)
p_t^d	-0.0236*	-0.1107**
	(0.0412)	(0.0206)
n_t^d	-0.0121	-0.0887*
	(0.2678)	(0.0318)
p_t^s	0.5720***	0.6415***
	(0.9998)	(0.9996)
n_t^s	-0.2629	-0.1723
	(0.2614)	(0.2928)
Adjusted R^2 without macro risks	0.6513	0.6543
Adjusted R^2 with macro risks	0.6914*	0.6941*

Table 12 – Cyclicality of the Term Premium. The dependent variable is annualized term premium computed as the observed US Treasury long yield minus the expected 1 quarter US Treasury yield over the life of the long yield. The expectations of 1 quarter yield over the life of the long yield are from Blue Chip survey and are available semi-annually. The sample is 1986Q2-2016Q4. The recession dummy is set to 1 if the anxious index for the quarter (as surveyed during that quarter) is above 50% and 0 otherwise. Demand/supply-ratio is the ratio of aggregate demand variance (good+bad) to aggregate supply variance (good+bad). Demand/supply- ratio is scaled to have the standard deviation of 1. Standard errors are Newey-West standard errors computed with 20 lags. The asterisks, *, **, and *** correspond to statistical significance at the 10, 5, and 1 percent levels, respectively.

	5 year	10 year
constant	0.6540	1.2404**
	(0.6496)	(0.6075)
recession-dummy	2.5410	6.2530
	(3.1010)	(3.8839)
demand-supply ratio	-0.2274	-0.2677
	(0.2030)	(0.1839)
recession-dummy×demand-supply ratio	-2.1670	-4.5687
	(2.3204)	(2.7858)
R^2	0.0282	0.0362

Table 13 – Explanatory Power of Macro Factors for Realized 10 Year Bond Return Variances. The sample is quarterly from 1962Q4 to 2016Q4. Realized variances are computed as the sums of squared daily bond returns inside the quarter. The standard deviation of each macro risk factor is scaled to 1. Financial factors are the level, slope, and curvature factors. The level factor is the average over 1-10 year yields. The slope factor is the 10 year yield minus the 1 quarter yield. The curvature factor is 10 year yield plus 1 quarter yield minus 2 times the 2 year yield. The signs in Panel B are from the OLS regression. The value in parentheses is the proportion out of 5,000 Bauer and Hamilton (2017) bootstrap runs where the t-stat for the coefficient is smaller than in data. The increase in adjusted R^2 significance, which is tested over the specification without the last set of factors (e.g., "3 financial factors+macro level factors+macro risks" row tests the incremental contribution of macro risks for the specification already including 3 financial factors and macro level factors), is Bauer and Hamilton (2017) small-sample adjusted significance using 5000 bootstrap runs. The asterisks, * , **, and *** correspond to statistical significance at the 10, 5, and 1 percent levels, respectively.

Panel A: Adjusted R^2's	
Macro risks	0.3473
Macro level factors	0.1890
3 financial factors	0.1390
Macro level factors + macro risks	0.4200***
3 financial factors +macro risks	0.4267***
3 financial factors+macro level factors	0.2937***
3 financial factors+macro level factors+macro risks	0.4408***
3 financial factors+macro risks+macro level factors	0.4408
macro risks+macro level factors+financial factors	0.4408
Panel B: Regression coefficients	
constant	0.0015*
	(0.9574)
$E_t\pi_{t+1}^{core}$	0.0026**
	(0.9822)
$E_t\pi_{t+1}$	-0.0016*
	(0.0348)
E_tg_{t+1}	3.14E-05
	(0.5970)
$ugap_t$	1.61E-04
	(0.8098)
p_t^d	8.48E-05
	(0.6284)
n_t^d	4.92E-04***
	(0.9998)
p_t^s	-3.50E-04*
	(0.9592)
n_t^s	3.15E-05
	(0.5452)

Table 14 – Explanatory Power of Macro Factors for Realized 10 Year Bond Return Variances over Lagged Realized Variances and GARCH models. The sample is quarterly from 1962Q4 to 2016Q4. Realized variances are computed as the sums of squared daily zero-coupon bond returns inside the quarter. All predictive variables are from the previous quarter. GJR-GARCH refers to conditional volatility from Glosten, Jagannathan, and Runkle (1993) model. GJR-GARCH outperforms GARCH in terms of Akaike Information Criterion. Financial factors are the level, slope, and curvature factors. The level factor is the average over 1-10 year yields. The slope factor is the 10 year yield minus the 1 quarter yield. The curvature factor is 10 year yield plus 1 quarter yield minus 2 times the 2 year yield. The increase in adjusted R^2 significance, which is tested over the specification without the last set of factors (e.g., "3 financial factors+macro level factors+macro risks" row tests the incremental contribution of macro risks for the specification already including 3 financial factors and macro level factors), is Bauer and Hamilton (2017) small-sample adjusted significance using 5000 bootstrap runs. The asterisks, * , **, and *** correspond to statistical significance at the 10, 5, and 1 percent levels, respectively.

Set of predictors	Adjusted R^2
Panel A: Predictability over lag 1 realized variance	
Lag 1 realized variance	0.5174
Lag 1 realized variance+macro risks	0.5813***
Lag 1 realized variance+macro level factors	0.5439*
Lag 1 realized variance1+3 financial factors	0.5223
Lag 1 realized variance+3 financial factors+macro risks	0.5916***
Lag 1 realized variance+macro risks+3 financial factors	0.5916
Lag 1 realized variance+macro level factors+macro risks	0.5977**
Lag 1 realized variance+macro risks+macro level factors	0.5977
Lag 1 realized variance+macro level factors+3 financial factors	0.5471
Lag 1 realized variance+3 financial factors+macro level factors	0.5471*
Lag 1 realized variance+3 financial factors+macro level factors+macro risks	0.5949**
Lag 1 realized variance+macro level factors+macro risks+3 financial factors	0.5949
Lag 1 realized variance+macro risks+3 financial factors+macro level factors	0.5949
Panel B: Predictability over GJR-GARCH variance	
GJR-GARCH	0.5191
GJR-GARCH+macro risks	0.7057***
GJR-GARCH+macro level factors	0.5734***
GJR-GARCH+3 financial factors	0.6082***
GJR-GARCH+3 financial factors+macro risks	0.7055***
GJR-GARCH+macro risks+3 financial factors	0.7055
GJR-GARCH+macro level factors+macro risks	0.7109***
GJR-GARCH+macro risks+macro level factors	0.7109
GJR-GARCH+macro level factors+3 financial factors	0.6290**
GJR-GARCH+3 financial factors+macro level factors	0.6290
GJR-GARCH+3 financial factors+macro level factors+macro risks	0.7155***
GJR-GARCH+macro level factors+macro risks+3 financial factors	0.7155
GJR-GARCH+macro risks+3 financial factors+macro level factors	0.7155
Panel C: Predictability power over lag 1 realized variance and GJR-GARCH variance	
Lag 1 realized variance+GJR-GARCH	0.6046***
GJR-GARCH+Lag 1 realized variance	0.6046***
Lag 1 realized variance+GJR-GARCH+macro risks	0.7086***
Lag 1 realized variance+GJR-GARCH+macro level factors	0.6134
Lag 1 realized variance+GJR-GARCH+3 financial factors	0.6375*
Lag 1 realized variance+GJR-GARCH+3 financial factors+macro risks	0.7076***
Lag 1 realized variance+GJR-GARCH+macro risks+3 financial factors	0.7076
Lag 1 realized variance+GJR-GARCH+macro level factors+macro risks	0.7148***
Lag 1 realized variance+GJR-GARCH+macro risks+macro level factors	0.7148
Lag 1 realized variance+GJR-GARCH+macro level factors+3 financial factors	0.6515**
Lag 1 realized variance+GJR-GARCH+3 financial factors+macro level factors	0.6515
Lag 1 realized variance+GJR-GARCH+3 financial factors+macro level factors+macro risks	0.7179**
Lag 1 realized variance+GJR-GARCH+macro level factors+macro risks+3 financial factors	0.7179
Lag 1 realized variance+GJR-GARCH+macro risks+3 financial factors+macro level factors	0.7179

Appendix A - Macroeconomic Impulse Responses and Bootstrapped Standard Errors

We characterize the long run effects of the structural shocks using standard impulse response analysis. For the purposes of calculating impulse response functions for the macro data, we use our estimated VAR(2) parameters. To compute the response of the four macroeconomic series at various horizons to the supply and demand shocks, we need the contemporaneous response of all the variables to supply and demand shocks. For the four macroeconomic series, these responses are the row elements of the Σ matrix corresponding to macro data in equation (2). For the two yield variables, we extract the time series for reduced-form shocks from the VAR(2)-estimation and simply regress these shocks onto the filtered supply and demand shocks. The responses of the six endogenous variables to the two structural shocks, supply and demand, of unit size at horizon h, are given by the expression:

$$IR(h) = (A_1^h + A_2^{\max(h-1,0)})\Sigma i,$$

where A_1 and A_2 are lag 1 and 2 AR matrices from the VAR(2)-model and i is the impulse.

Note that the standard error for the impulse response coefficients must account not only for the estimation of the VAR(2) parameters but also for the error incurred in identifying supply and demand shocks, which involves the higher order moments of VAR residuals. These sources of error affect the distribution of the sampling error of the loadings of the endogenous variables onto supply and demand shocks, the time series estimates of the supply and demand shocks, and the impulse response functions.

To account for all of these sources of error, we use a bootstrapping routine. We begin by sampling, with replacement, the reduced-form shocks from the estimated VAR(2) model. We assemble synthetic samples using 22 randomly chosen blocks with a length of 20 quarters. This results in synthetic samples of approximately the same length as our data (220 for bootstraps, 225 for the data). We use these shocks and the estimated VAR(2) parameters to build up synthetic samples of the endogenous variables. Note that we do not need any estimates of the covariance matrix of shocks to do this. Beginning from these synthetic samples, we follow the same procedures for each bootstrap sample that we do for the actual sample to calculate all the statistics of interest:

- Estimate VAR(2) parameters on the synthetic sample.
- Estimate higher-order moments of the reduced form shocks and their covariance matrix
- Estimate loadings of the macro variables onto supply and demand using the GMM procedure on the higher order moments
- Invert supply and demand shocks using the Kalman filter procedure
- Estimate the loadings of the yield variables onto the supply and demand shocks by OLS
- Estimate the impulse responses

We use the same procedure to estimate standard errors for model parameters.

The results are as follows with standard errors in parentheses (recall that these shocks have unit variance by construction):

Panel A: Contemporaneous (Quarter 0) Responses		
	Demand Shock	Supply Shock
Real GDP level	0.40%	0.34%
	(0.10%)	(0.08%)
Price level	0.39%	-0.17%
	(0.10%)	(0.06%)
Panel B: Cumulative (20 Quarters) Responses		
	Demand Shock	Supply Shock
Real GDP level	0.09%	0.52%
	(0.27%)	(0.27%)
Price level	2.15%	-0.05%
	(0.66%)	(0.54%)

The effects are consistent with the standard Keynesian interpretation. Demand shocks have large short run effects on real GDP growth (with the initial shock being 0.40 percent) but their cumulative effect on output is small (0.09 percent) and insignificantly different from zero. Supply shocks generate smaller short run GDP growth effects but their cumulative effect is 0.52 percent which is significantly different from zero. Demand and supply shocks have very different effects on the price level, with the cumulative effects close to +2 percent in the case of demand shocks, but the supply shock effect peters out to zero. In sum, our identification scheme yields shocks whose long-run effects are consistent with a well-established macroeconomic literature.

Following Jorda (2005), we also calculate the model-free impulse responses using OLS regressions of the form:

$$Y_{t+h} = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \beta_3 \hat{u}_{t-1}^{supply} + \beta_4 \hat{u}_{t-1}^{demand} + \epsilon_{t+h},$$

where $\hat{u}^{supply}$ and $\hat{u}^{demand}$ are the inverted supply and demand shocks. Standard errors are computed as above.

The results are as follows with standard errors in parentheses:

Cumulative (20 Quarters)		
	Demand Shock	Supply Shock
Real GDP level	0.37%	0.81%
	(0.46%)	(0.42%)
Price level	2.15%	-0.05%
	(0.31%)	(0.30%)

Appendix B - Maximum likelihood estimation of demand and supply shock dynamics

We restrict attention to the demand shock estimation, as the supply shock estimation is identical. The system to estimate is:

$$
\begin{aligned}
&u_{t+1}^d = \sigma_p^d \omega_{p,t+1}^d - \sigma_n^d \omega_{n,t+1}^d, \\
&\omega_{p,t+1}^d \sim \Gamma(p_t^d, 1) - p_t^d, \\
&\omega_{n,t+1}^d \sim \Gamma(n_t^d, 1) - n_t^d, \\
&p_{t+1}^d = \bar{p}^d + \rho_p^d (p_t^d - \bar{p}^d) + \sigma_{pp}^d \omega_{p,t+1}^d, \\
&n_{t+1}^d = \bar{n}^d + \rho_n^d (n_t^d - \bar{n}^d) + \sigma_{nn}^d \omega_{n,t+1}^d,
\end{aligned}
$$

where only the time series of demand shock realizations, $\{u_t^d\}_{t=1}^T$ is observed.

The following notation is defined:

$U_t^d \equiv \{u_1^d, ..., u_t^d\}$ is the sequence of observations up to time t.

$F(i\phi, i\psi^1, i\psi^2 | U_t^d) \equiv E(e^{i\phi u_{t+1}^d + i\psi^1 p_{t+1}^d + i\psi^2 n_{t+1}^d} | U_t^d)$ is the next period's joint conditional characteristic function of the observation and the state variables.

$G_{t|s}(i\psi^1, i\psi^2) \equiv E(e^{i\psi^1 p_t^d + i\psi^2 n_t^d} | U_s^d)$ is the characteristic function of the time t state variables conditioned on observing data up to time s.

The estimation procedure is an application of Bates (2006)'s algorithm for the component model of two gamma distributed variables and consists of the time 0 initialization and 3 steps repeated for each observation in $\{u_t^d\}_{t=1}^T$. At time 0, the characteristic function of the state variables $G_{0|0}(i\psi^1, i\psi^2)$ is initialized. The distribution of p_0^d and n_0^d is approximated with gamma distributions. Note that the unconditional mean and variance of p_t^d are $E(p_t^d) = \bar{p}^d$ and $Var(p_t^d) = \frac{\sigma_{pp}^2}{1-\rho_p^{d2}} \bar{p}^d$, respectively. The approximation by the gamma distribution with the shape parameter k_0 and the scale parameter σ_0^p is done by matching the first two unconditional moments. Using the properties of the gamma distribution, $k_0^p = \frac{E^2 p_t^d}{Var(p_t^d)}$ and $\theta_0^p = \frac{Var(p_t^d)}{E(p_t^d)}$. Thus, p_0^d is assumed to follow $\Gamma(k_0^p, \theta_0^p)$ and n_0^d is assumed to follow $\Gamma(k_0^n, \theta_0^n)$, where k_0^n and θ_0^n are computed in the same way. Using the properties of the expectations of the gamma variables, $G_{0|0}(i\psi^1, i\psi^2) = e^{-k_0^p \ln(1-\theta_0^p i\psi^1) - k_0^n \ln(1-\theta_0^n i\psi^2)}$. Given $G_{0|0}(i\psi^1, i\psi^2)$, computing the likelihood of U_T^d is performed by repeating the steps 1-3 below for all subsequent values of t.

Step 1. Computing the next period's joint conditional characteristic function of the observation and the state variables:

$$
\begin{aligned}
&F(i\Phi, i\psi^1, i\psi^2|U_t^d) = E(E(e^{i\Phi(\sigma_p^d\omega_{p,t+1}^d-\sigma_n^d\omega_{n,t+1}^d)+i\psi^1(\bar{p}^d+\rho_p^d p_t^d+\sigma_{pp}^d\omega_{p,t+1}^d)+i\psi^2(\bar{n}^d(1-\rho_n^d)+\rho_n^d n_t^d+\sigma_{nn}^d\omega_{n,t+1}^d)}|U_t^d) \\
&= E(e^{i\psi^1\bar{p}^d(1-\rho_p^d)+i\psi^2\bar{n}^d(1-\rho_n^d)+(i\psi^1\rho_p^d-\ln(1-i\Phi\sigma_p^d-i\psi^1\sigma_{pp}^d)-i\Phi\sigma_p^d-i\psi^1\sigma_{pp}^d)p_t^d+(i\psi^2\rho_n^d-\ln(1+i\Phi\sigma_n^d-i\psi^2\sigma_{nn}^d)+i\Phi\sigma_n^d-i\psi^2\sigma_{nn}^d)n_t^d}|U_t^d) \\
&= e^{i\psi^1\bar{p}^d(1-\rho_p^d)+i\psi^2\bar{n}^d(1-\rho_n^d)}G_{t|t}(i\psi^1\rho_p^d-\ln(1-i\Phi\sigma_p^d-i\psi^1\sigma_{pp}^d)-i\Phi\sigma_p^d-i\psi^1\sigma_{pp}^d, i\psi^2\rho_n^d-\ln(1+i\Phi\sigma_n^d-i\psi^2\sigma_{nn}^d)+i\Phi\sigma_n^d-i\psi^2\sigma_{nn}^d).
\end{aligned}
$$

Step 2. Evaluating the conditional likelihood of the time $t+1$ observation:

$$
p(u_{t+1}^d|U_t^d) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(i\Phi, 0, 0|U_t^d)e^{-i\Phi u_{t+1}^d})d\Phi,
$$

where the function F is defined in step 1 and the integral is evaluated numerically.

Step 3. Computing the conditional characteristic function for the next period, $G_{t+1|t+1}(i\psi^1, i\psi^2)$:

$$
G_{t+1|t+1}(i\psi^1, i\psi^2) = \frac{\frac{1}{2\pi}\int_{-\infty}^{\infty} F(i\Phi, i\psi^1, i\psi^2|U_t^d)e^{-i\Phi u_{t+1}^d}d\Phi}{p(u_{t+1}^d|U_t^d)}.
$$

As above, the function $G_{t+1|t+1}(i\psi^1, i\psi^2)$ is also approximated with the gamma distribution via matching the first two moments of the distribution. The moments are obtained by taking the first and second partial derivatives of the joint characteristic function:

$$
\begin{aligned}
E_{t+1}p_{t+1}^d &= \frac{1}{2\pi p(u_{t+1}^d|U_t^d)}\int_{-\infty}^{\infty} F_{\psi^1}(i\Phi, 0, 0|U_t^d)e^{-i\Phi u_{t+1}^d}d\Phi, \\
Var_{t+1}p_{t+1}^d &= \frac{1}{2\pi p(u_{t+1}^d|U_t^d)}\int_{-\infty}^{\infty} F_{\psi^1\psi^1}(i\Phi, 0, 0|U_t^d)e^{-i\Phi u_{t+1}^d}d\Phi - E_{t+1}^2 p_{t+1}^d, \\
E_{t+1}n_{t+1}^d &= \frac{1}{2\pi p(u_{t+1}^d|U_t^d)}\int_{-\infty}^{\infty} F_{\psi^2}(i\Phi, 0, 0|U_t^d)e^{-i\Phi u_{t+1}^d}d\Phi, \\
Var_{t+1}n_{t+1}^d &= \frac{1}{2\pi p(u_{t+1}^d|U_t^d)}\int_{-\infty}^{\infty} F_{\psi^2\psi^2}(i\Phi, 0, 0|U_t^d)e^{-i\Phi u_{t+1}^d}d\Phi - E_{t+1}^2 n_{t+1}^d,
\end{aligned}
$$

where F_{ψ^i} denotes the derivative of F with respect to ψ^i. The expressions inside the integral are obtained in closed form by derivating the function $F(i\Phi, i\psi^1, i\psi^2|U_t^d)$ in step 1, and integrals are evaluated numerically. Using the properties of the gamma distribution, the values of the shape and the scale parameters are $k_{t+1}^p = \frac{E_{t+1}^2 p_{t+1}^d}{Var_{t+1}p_{t+1}^d}$ and

$\theta^p_{t+1} = \frac{Var_{t+1} p^d_{t+1}}{E_{t+1} p^d_{t+1}}$, respectively. The expressions for k^n_{t+1} and θ^n_{t+1} are similar.

The total likelihood of the time series is the sum of individual likelihoods from step 2: $L(Y_T) = \ln p(u^d_1 | k^p_0, \theta^p_0) + \sum_{t=2}^{T} \ln p(u^d_{t+1} | U^d_t)$.

Appendix C - Conditional Covariances between Macroeconomic Time Series: Evidence from Other Models

In order to validate the BEGE patterns, we consider three alternative conditional covariance models. Our first model is simply a rolling covariance. We report the results using a 10 quarters rolling covariance, but the pattern is very similar for values between 4 and 20 quarters. Our second model is DCC-GJR-GARCH. Under this model, variances follow the GJR-GARCH model of Glosten, Jagannathan, and Runkle (1993) and the conditional correlation follows DCC model of Engle (2002). Our third model is a bivariate Gaussian regime-switching model in the spirit of Hamilton (1989). We use a 2 state model, because it is the model preferred by the Akaike information criterion. Each of the models above is estimated twice: to compute conditional covariances between VAR(2) shocks to real GDP growth and aggregate inflation and between VAR(2) shocks to the unemployment gap and the core inflation, respectively.

Figure 7 plots the implied conditional real GDP growth and aggregate inflation shocks covariances. Note that economically all models deliver largely the same pattern: the covariance spikes around 1980 and then again during the Great Recession while remaining rather stable around 0 outside these two periods. Table 15 confirms the similarity of patterns across models quantitatively by showing that conditional covariances across different models are positively correlated (correlations >0.40 for the BEGE model with the other models).

Figure 8 plots the implied conditional unemployment gap and core inflation shocks covariances from different models. Economically, the message is similar to Figure 7: all models imply a covariance around 0 with a prolonged drop around 1980 and a sharp but short-lived drop during the Great Recession. Table 16 confirms the similarity of patterns across models quantitatively by showing that conditional covariances across different models are positively correlated. The BEGE model covariances now correlate at least 48% with the covariances produced by the other models.

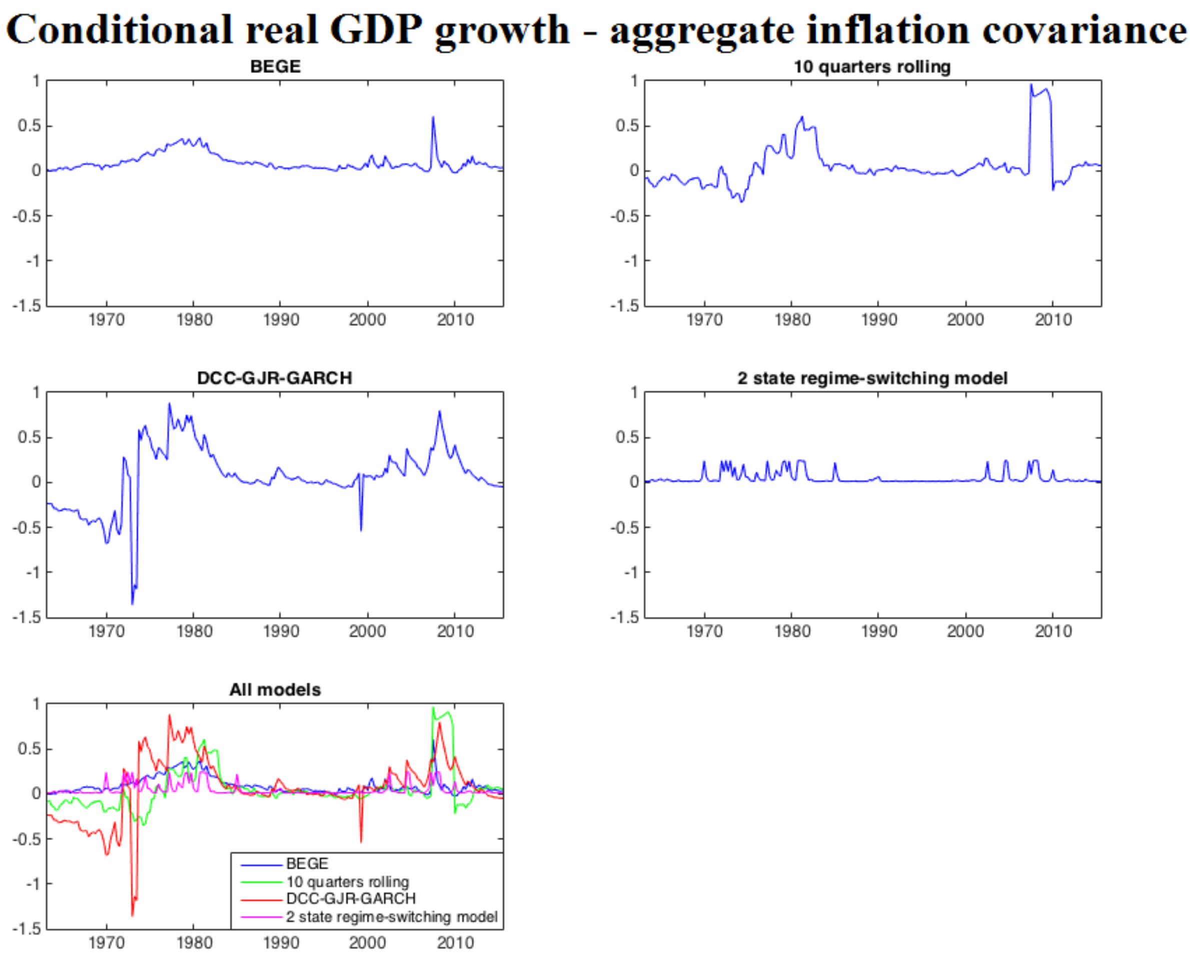

Figure 7 – Conditional Covariances between VAR(2) Shocks to GDP Growth and Aggregate Inflation. Data is quarterly.

Table 15 – Correlation of Conditional Covariances between GDP Growth and Aggregate Inflation Shocks from Different Models. Data is quarterly.

	Rolling window	DCC-GJR-GARCH	2-state regime-switching
BEGE	0.4534	0.5163	0.4350
Rolling window		0.5301	0.2750
DCC-GJR-GARCH			0.3122

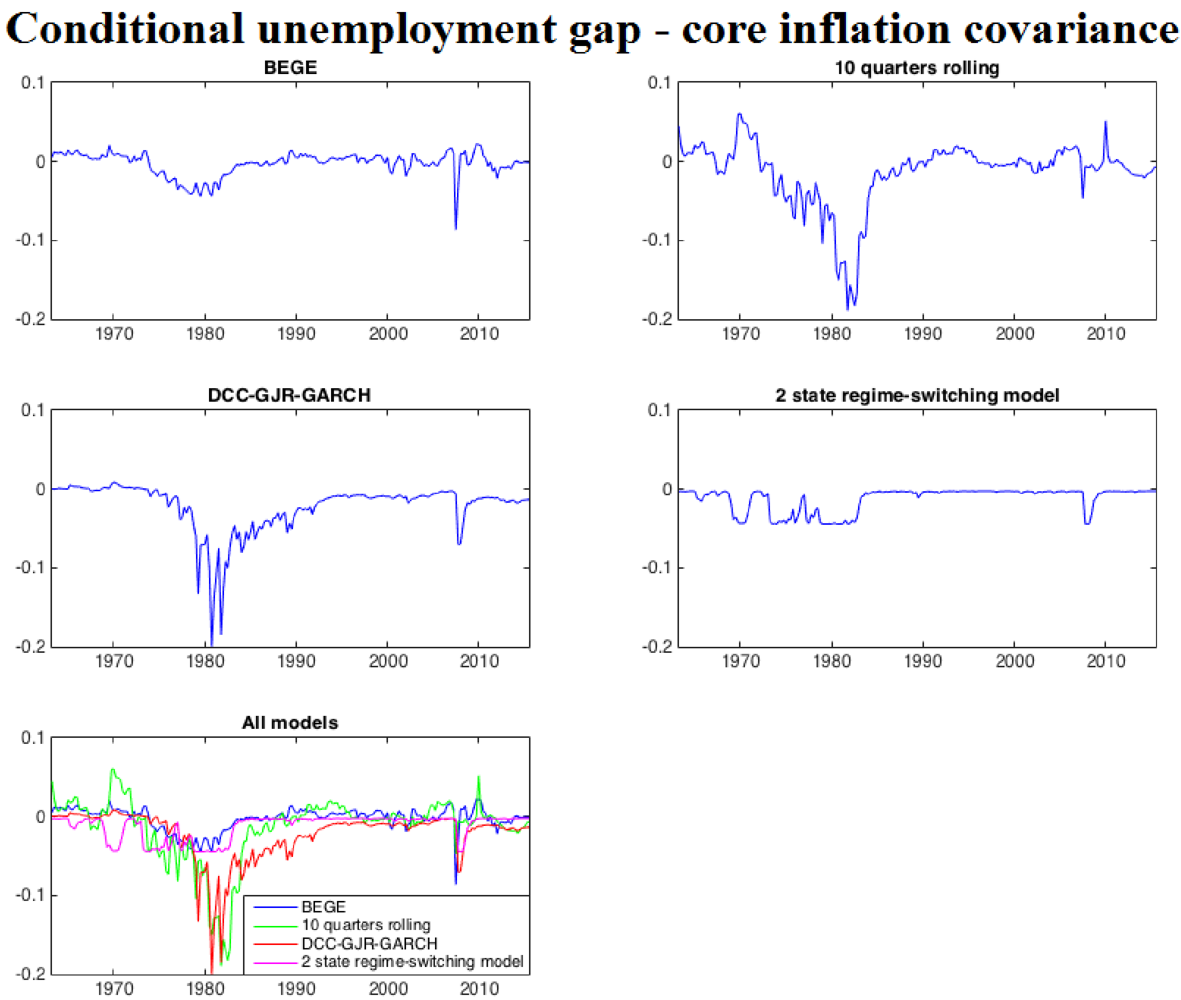

Figure 8 – Conditional Covariances between VAR(2) Shocks to Unemployment Gap and Core Inflation. Data is quarterly.

Table 16 – Correlation of Conditional Covariances between Unemployment Gap and Core Inflation Shocks from Different Models. Data is quarterly.

	Rolling window	DCC-GJR-GARCH	2-state regime-switching
BEGE	0.6454	0.4871	0.4788
Rolling window		0.7713	0.5173
DCC-GJR-GARCH			0.4312

The issue with the analysis in Figures 7-8 and Tables 15-16 is that it does not provide standard errors. Indeed, it is not straightforward to come up with standard errors for conditional moments. As a rough approximation, we generate bootstrap standard errors for rolling window covariance estimates as follows. At each time point we use the same window which we used to to compute the rolling covariance estimate for that time point (10 quarters). Then we resample observations in that window 10,000 times and compute the covariance for each replication. This gives us the distribution of covariances for each point of time. We use this distribution to construct the 95% confidence interval of the covariance values for that window.

Figure 9 plots the outcome illustrating that BEGE estimates generally fall inside the 95% confidence interval for the 10 quarters rolling covariance estimates. For the real GDP growth-aggregate inflation covariance the BEGE estimate is inside the 95% confidence interval of the rolling estimate 77.24% of the time. For the unemployment gap-core inflation covariance this number is 93.33%. While these numbers are below 95%, the difference is often economically non-significant. To see this, note from the top panel of Figure 9 that the BEGE estimate falls outside the rolling covariance 95% confidence interval for a long time, for example, during 1967-1971. During this period the rolling covariance estimate is slightly negative while the BEGE estimate is essentially 0, and although the confidence interval is statistically narrow, economically the difference between the covariances generated by the two models is rather small. Other longer periods where the BEGE estimate for the real GDP growth-aggregate inflation covariance falls outside the rolling estimate 95% confidence interval are 1987-1989 and 1994-1996. Again these are the outcomes of very narrow confidence intervals: economically, it is difficult to see the differences between BEGE and rolling estimates from the top panel of Figure 9.

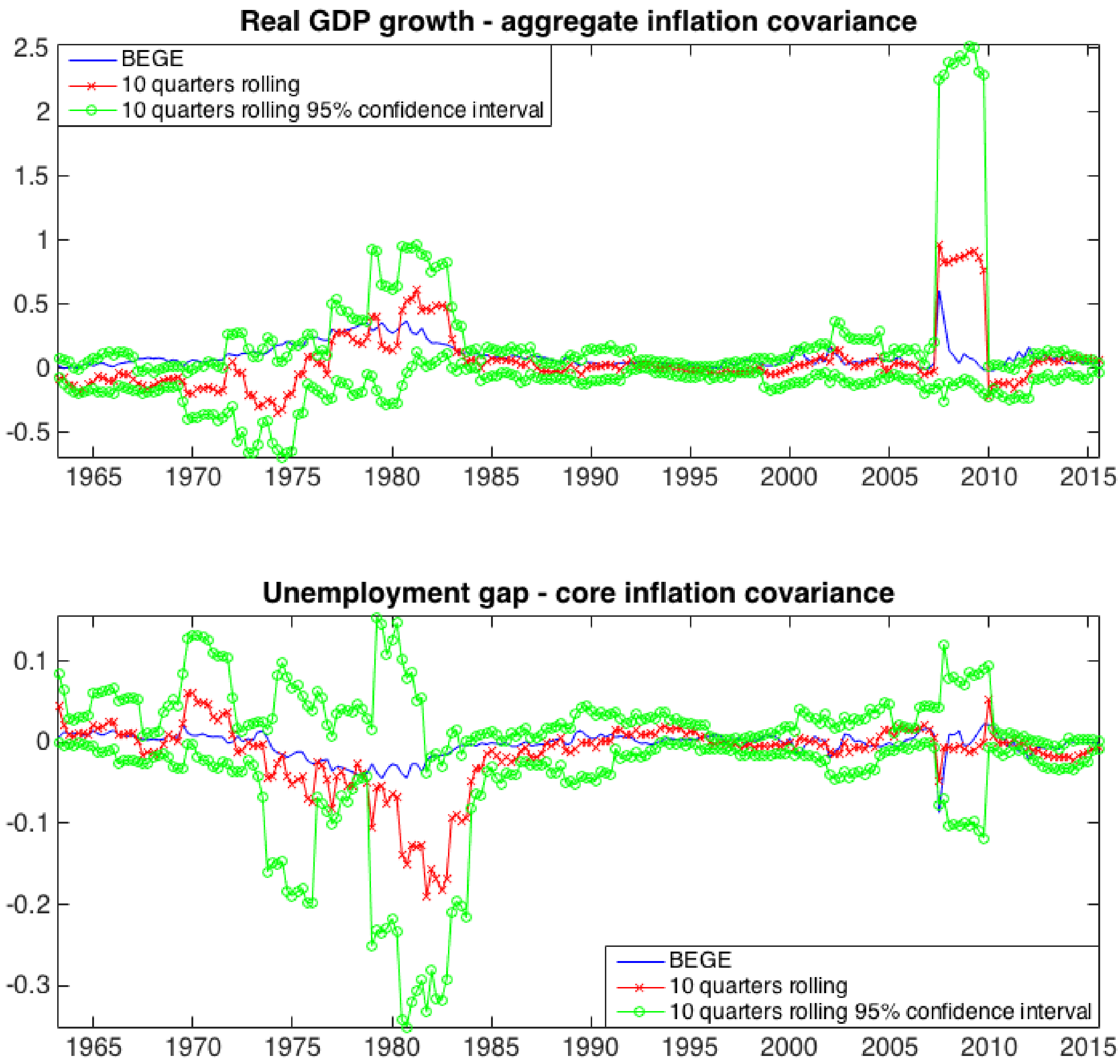

Figure 9 – Conditional BEGE Covariances and 10 Quarters Rolling Windows Covariance Estimates with 95% Confidence Intervals. 10 quarters rolling windows covariance estimates are obtained from 10,000 bootstrap runs at each time point. Data is quarterly.

Appendix D - Additional Results on Explanatory Power of Macro Risks

Explanatory Power (Adjusted R^2) of Macro Risk Factors over Ang-Piazzesi Factors for Yield Curve Factors. The sample is quarterly from 1962Q4 to 2016Q4. Ang-Piazzesi factors are contemporaneous Ang and Piazzesi (2003) real and nominal factors. Macro level factors are expected real GDP growth, expected aggregate and core inflation, and unemployment gap. Financial factors are the level, slope, and curvature factors. The level factor is the average over 1-10 year yields. The slope factor is the 10 year yield minus the 1 quarter yield. The curvature factor is 10 year yield plus 1 quarter yield minus 2 times the 2 year yield. The increase in adjusted R^2 significance, which is always tested over the specification in the previous row, is Bauer and Hamilton (2017) small-sample adjusted significance using 5000 bootstrap runs. The asterisks, *, **, and ***, correspond to statistical significance at the 10, 5, and 1 percent levels, respectively.

	Level	Slope	Curvature
Ang-Piazzesi (2003) factors	0.2555	0.3126	0.1229
Ang-Piazzesi (2003) factors + macro level factors	0.7122***	0.5906***	0.2918***
Ang-Piazzesi (2003) factors + macro level factors + macro risks	0.7974***	0.6078	0.4086***

Explanatory Power (Adjusted R^2) of Macro Risk Factors for Yield Curve Factors over Realizations of Macroeconomic Time Series. The sample is quarterly from 1962Q4 to 2016Q4. Macro level factors are real GDP growth, aggregate and core inflation, and unemployment gap. Financial factors are the level, slope, and curvature factors. The level factor is the average over 1-10 year yields. The slope factor is the 10 year yield minus the 1 quarter yield. The curvature factor is 10 year yield plus 1 quarter yield minus 2 times the 2 year yield. The increase in adjusted R^2 significance, which is always tested over the specification in the previous row, is Bauer and Hamilton (2017) adjusted significance using 5000 bootstrap runs. The asterisks, *, **, and *** correspond to statistical significance at the 10, 5, and 1 percent levels, respectively.

Realizations of Macroeconomic Level Factors and Macro Risks			
	Level	Slope	Curvature
Realization of macroeconomic level factors	0.4795	0.5277	0.2168
Realization of macroeconomic level factors + macro risks	0.7151***	0.5675*	0.4038***

Explanatory Power (Adjusted R^2) of Macro Risk Factors for Quarterly Excess Bond Returns over Ang-Piazzesi (2003) and Financial Factors. The sample is quarterly from 1962Q4 to 2016Q4. Ang-Piazzesi factors are lag 1-12 Ang and Piazzesi (2003) real and nominal factors. Macro level factors are expected real GDP growth, expected aggregate and core inflation, and unemployment gap. Financial factors are the level, slope, and curvature factors. The level factor is the average over 1-10 year yields. The slope factor is the 10 year yield minus the 1 quarter yield. The curvature factor is the 10 year yield plus the 1 quarter yield minus 2 times the 2 year yield. The increase in adjusted R^2 significance, which is tested over the specification in the previous row, is Bauer and Hamilton (2017) adjusted significance using 5000 bootstrap runs. The asterisks, *, **, and *** correspond to statistical significance at the 10, 5, and 1 percent levels, respectively.

Predictors	1 year bond	2 year bond	5 year bond	10 year bond
3 financial factors	0.0663	0.0653	0.0638	0.0795
3 financial factors+Ang-Piazzesi	0.1549**	0.1415**	0.1325*	0.1295
3 financial factors+Ang-Piazzesi+macro level factors	0.1734	0.1537	0.1432	0.1471
3 financial factors+Ang-Piazzesi+macro level factors+macro risks	0.1903	0.1870**	0.1710*	0.1622

Appendix E - Expected Excess Bond Return Cyclicality

Cyclicality of Expected Excess Bond Returns. The sample is quarterly 1969Q4-2016Q4. The dependent variable is the expected annualized quarterly excess return computed from the OLS regressions of realized annualized quarterly excess returns on 4 macro level factors (expected aggregate and core inflations, expected real GDP growth, and unemployment gap) and 4 macro risks (good and bad demand and supply variances). The recession dummy is set to 1 if the corresponding recession indicator is on and 0 otherwise. Demand/supply-ratio is the ratio of aggregate demand variance (good+bad) to aggregate supply variance (good+bad). Demand/supply-ratio is scaled to have the standard deviation of 1. Standard errors are Newey-West standard errors computed with 20 lags. The asterisks, ** and ***, correspond to statistical significance at the 5 and 1 percent levels, respectively.

	1 year bond			5 year bond			10 year bond		
recession indicator	NBER	30% anxious index	50% anxious index	NBER	30% anxious index	50% anxious index	NBER	30% anxious index	50% anxious index
constant	1.4523***	1.1801***	1.2227***	7.0844	5.9878***	6.2253***	13.3929***	13.2660***	13.1634***
	(0.2862)	(0.2929)	(0.2579)	(1.3336)	(1.6030)	(0.9967)	(2.6969)	(2.3571)	(2.1864)
recession-dummy	-0.1619	0.5817	0.6844**	-0.5175	3.6825*	4.4037***	1.5612	0.1260	1.5002
	(0.3776)	(0.4183)	(0.3471)	(2.0585)	(2.0213)	(1.5038)	(3.7882)	(3.0674)	(3.6225)
demand-supply ratio	-0.6599***	-0.5456***	-0.5377***	-2.8687***	-2.5133***	-2.4920***	-5.3758***	-5.5066***	-5.2721***
	(0.1339)	(0.1552)	(0.1444)	(0.6183)	(0.7079)	(0.4925)	(1.1351)	(0.9946)	(0.9748)
recession-dummy*demand-supply ratio	0.3664	0.0313	-0.0365	1.1137	-0.3172	-0.7412	0.6516	1.3812	0.2911
	(0.2309)	(0.1906)	(0.1571)	(0.7273)	(0.7413)	(0.5901)	(1.1785)	(1.2381)	(1.3283)
Adjusted R^2	0.4394	0.4025	0.3655	0.4178	0.4487	0.4086	0.3804	0.3771	0.3596

Appendix F - GARCH Models of Bond Returns

The GARCH model (Bollerslev, 1987) is:

$$r_{t+1} = \mu + \sigma_t \epsilon_{t+1},$$
$$\epsilon_{t+1} \sim \mathcal{N}(0,1),$$
$$\sigma_t^2 = \bar{\sigma}^2 + \rho(\sigma_{t-1}^2 - \bar{\sigma}^2) + \phi \epsilon_t^2.$$

The GJR-GARCH model (Glosten, Jagannathan, and Runkle, 1993) is:

$$r_{t+1} = \mu + \sigma_t \epsilon_{t+1},$$
$$\epsilon_{t+1} \sim \mathcal{N}(0,1),$$
$$\sigma_t^2 = \bar{\sigma}^2 + \rho(\sigma_{t-1}^2 - \bar{\sigma}^2) + \phi \epsilon_t^2 \mathbb{1}_{\epsilon_t \geq 0} + \phi_n \epsilon_t^2 \mathbb{1}_{\epsilon_t < 0},$$

where $\mathbb{1}$ is the indicator function.

Both models are estimated through maximum likelihood. Parameter standard errors are computed as the square roots of the diagonal elements of the inverse information matrix. Models are estimated using quarterly excess bond returns on 10 year zero-coupon bonds from 1962Q4 to 2016Q4.

Parameter estimates are as follows (standard errors are in parentheses):

	GARCH	GJR-GARCH
μ	2.0397	1.6732
	(1.2916)	(1.3083)
$\bar{\sigma}^2$	164.9249	173.8736
	(59.3819)	(53.4069)
ρ	0.7241	0.7740
	(0.0844)	(0.1070)
ϕ	0.1808	0.07592
	(0.0710)	(0.0839)
ϕ_n		0.2217
		(0.1021)
Akaike Information Criterion	1929.9761	1929.6805