Procyclicality of the Comovement between Dividend Growth and Consumption Growth

Nancy R. Xu*

January 25, 2018

Abstract

I document that dividend growth and consumption growth comove procyclically. This new stylized fact empirically resolves the “Duffee Puzzle”—stock returns and consumption growth covary procyclically (Duffee, 2005)—but contradicts extant theoretical assumptions in asset pricing models. I then design a new data generating process (DGP) for the joint consumption-dividend dynamics which fits the procyclical comovement and a wide set of other related second moments. Lastly, I solve a variant of Campbell and Cochrane’s habit formation model with this new DGP and the procyclical consumption-dividend growth co-movement as a new state variable. The new procyclical component in the amount of risk induces a more volatile price-dividend ratio at the cost of a lower equity premium due to the now counterbalancing dynamics of the price (countercyclical) and amount (procyclical) of risk. In addition, the new state variable accounts for 13% of the variability of the price dividend ratio in the data and carries a positive price of risk in the cross-section of stock returns.

Keywords: procyclical amount of risk, amount of risk decomposition, consumption-dividend comovement, habit formation, equity premium, price of comovement risk, value premium.

*Xu: nancy.xu@gsb.columbia.edu, 3022 Broadway, Columbia Business School, New York, NY, 10027. I would like to thank Geert Bekaert and an anonymous referee who helped improve the paper greatly. I have also benefited from discussions with Ravi Bansal (discussant), Nina Boyarchenko, Fernando Duarte, Thomas Eisenbach, Eric Engstrom, Andrey Ermolov, Bob Hodrick, Mete Kilic, Lars Lochstoer, Christian Lundblad (discussant), Michaela Pagel, Sung June Pyun (discussant), Carlos Ramírez (discussant), Asani Sarkar, Tano Santos and Xiaoyan Zhang, and from presentations at the 15th Transatlantic Doctoral Conference at London Business School, the Ph.D. Seminar at Columbia Business School, the 28th Australasian Banking and Finance Conference at UNSW (second-place prize winner), the AFA 2017 Session attendees, the Federal Reserve Bank of New York and the 10th Annual SoFiE Conference (Macroeconomics, Finance, and Risks). An earlier version of the paper was circulated with the title “Procyclicality of the Correlation between Dividend Growth and Consumption Growth”. Appendix available at https://www.nancyxu.net/research. All errors are my own.
1 Introduction

Duffee (2005) documents that the amount of consumption risk, that is the conditional covariance between equity returns and consumption growth, is procyclical. Provided that many well-accepted theories imply a constant (e.g., Bansal and Yaron, 2004) or countercyclical (e.g., Campbell and Cochrane, 1999) amount of risk, I identify and coin Duffee (2005)’s finding the “Duffee Puzzle.”

In this article, I first empirically demonstrate that this procyclicality is generated by a procyclical comovement between the cash flow part of the market return and consumption growth, and establish a set of new empirical facts regarding consumption growth, dividend growth, and the amount of risk. Then, I propose a new data generating process (DGP) for the joint dynamics of consumption and dividend growth that makes the empirical findings amenable to consumption-based asset pricing models. Lastly, I explore how incorporating more realistic dynamics into the amount of risk affects the performance of extant dynamic asset pricing models.

Using a flexible empirical framework and a longer sample period (January 1959–June 2014) than the one used in Duffee (2005), the empirical part of this article continues to find robust evidence for the procyclical conditional correlation and conditional covariance between market returns and consumption growth. Then, I decompose the amount of risk into two components,

\[
\text{Cov}_t \left( r_{t+1}^m, \Delta c_{t+1} \right) = \underbrace{\text{Cov}_t \left( \Delta d_{t+1}, \Delta c_{t+1} \right)}_{\text{Exogenous}} + \underbrace{\text{Cov}_t \left( r_{t+1}^m - \Delta d_{t+1}, \Delta c_{t+1} \right)}_{\text{Endogenous}},
\]

where \( r_{t+1}^m \) is the log market return, \( \Delta c_{t+1} \) the log consumption growth, and \( \Delta d_{t+1} \) the log dividend growth. The decomposition of the amount of consumption risk yields a conditional covariance, \( \text{Cov}_t \left( \Delta d_{t+1}, \Delta c_{t+1} \right) \), which is modeled exogenously in the extant consumption-based asset pricing literature, and an endogenous conditional covariance, \( \text{Cov}_t \left( r_{t+1}^m - \Delta d_{t+1}, \Delta c_{t+1} \right) \). Although there is limited research on the cyclicalty of the endogenous component above, both extant theories and empirical evidence suggest that the comovement between changes in the log price dividend ratio—a linear proxy for the non-dividend part of the market return—and consumption growth is not procyclical. Therefore, the covariance between dividend and consumption growth must be strongly procyclical in order to explain the Duffee Puzzle. Such procyclical comovement between dividend and consumption growth could arise from managers’ preference for dividend smoothing. According to Lintner (1956) and a follow-up study by Brav, Graham, Harvey, and Michaely (2005), a “prudent foresighted” manager is reluctant to cut dividends when the economy is in a downturn unless he/she expects the decrease in earnings to be persistent[1]. As a result, changes in financial payouts are expected to be less associated with macroeconomic shocks during bad times than during normal times.

[1] Since Lintner (1956), theories of why firms smooth their dividends are primarily based on either asymmetric information (Kumar, 1988; Brennan and Thakor, 1990; Fudenberg and Tirole, 1995; DeMarzo and Sannikov, 2008; Guttmann, Kadan, and Kandel, 2010) or agency considerations (Allen, Bernardo, and Welch, 2000; DeAngelo and DeAngelo, 2007; Lambrecht and Myers, 2012).
Indeed, I find that the exogenous component of the amount of risk behaves \textit{procyclically} and the endogenous component behaves \textit{countercyclically}. My results are robust to using various measures of the consumption growth innovations, different estimates of the conditional variances, and different proxies for cash flow growth rates (e.g., earnings growth instead of dividend growth). To quantify the significance of the procyclical exogenous component in explaining the procyclical amount of risk, my empirical results reveal that the share, $\frac{\text{Cov}(\Delta d_{t+1}, \Delta c_{t+1})}{\text{Cov}(\Delta r_{m+1}, \Delta c_{t+1})}$, is surprisingly volatile over the business cycle, reaching a peak of 106% during the 1960s expansion and a trough of -89% following the 1969-70 recession. The empirical part of the paper concludes with a comprehensive list of ten stylized facts pertinent to the Duffee Puzzle and its components. In particular, all three comovement statistics (the conditional correlation, covariance and beta of dividend growth to consumption growth) are shown to be procyclical, while the conditional variance of consumption growth is heteroskedastic and countercyclical.

Next, I formulate a parsimonious DGP for the joint dynamics of consumption and dividend growth with a minimum number of state variables that matches the stylized facts. In contrast, as I discuss in detail in Section 3, state-of-the-art consumption-based asset pricing models mostly assume unrealistic joint dynamics between consumption growth (which enters the utility function) and dividend growth (which constitutes the cash flow process of the equity claim to be priced).

The new DGP exhibits two empirically salient features. First, I introduce a new state variable $b$ capturing a time-varying sensitivity of dividend growth to consumption growth (or dividend-consumption beta). In particular, this state variable is procyclical as it comoves positively with consumption growth (the only macroeconomic variable in consumption-based asset pricing framework). The procyclical dividend beta is new to the literature, both empirically and theoretically.

Second, to ensure the simultaneous fit of (1) procyclical correlation and covariance between consumption and dividend growth and (2) countercyclical consumption growth variance, I allow consumption growth to receive two independent shocks each period, a homoskedastic Gaussian shock named the “fundamental shock” and an asymmetric heteroskedastic gamma shock named the “event shock”. To justify the decomposition, I apply the filtration-based maximum likelihood methodology (Bates, 2006) to obtain the fundamental and event shock realizations, and provide economic interpretations of the two consumption shocks:

The (filtered) fundamental shock, behaving procyclically, explains most of the total consumption growth variability during the sample period and has a significant and negative correlation with the detrended consumption-wealth ratio introduced by Lettau and Ludvigson (2001)—which is consistent with my DGP. In my DGP, dividend growth is sensitive to the contemporaneous consumption fundamental shock with a persistent procyclical beta that also comoves positively with the fundamental shock. Suppose a positive fundamental shock arrives this period. The persistent procyclical dividend beta is expected to remain \textit{high} in the future. Because dividend growth variance increases with dividend beta, it is the persistently \textit{high} dividend growth variance in expectations that gets capitalized in financial wealth, driving up the
wealth-consumption ratio in the current period. The event shocks, behaving countercyclically, drive the countercyclicality of the consumption growth variance. The conditional variance of the event shock constitutes the second state variable of the new DGP: macroeconomic uncertainty, denoted as \( n \).

In the last part of the paper, I formally demonstrate the ability of the new DGP to generate realistic dynamics of the amount of risk in a variant of the Campbell and Cochrane model (henceforth, CC)—thus accommodating the Duffee Puzzle—and explore new asset pricing implications of the new state variable \( b \). While the time variation in the price of risk, as in the standard CC model, is driven by the procyclical surplus consumption ratio (\( s \)), the time variation in the amount of risk is now determined by two countervailing sources, which is consistent with the empirical evidence of the Duffee Puzzle decomposition. More precisely, the amount of risk now contains both procyclical and countercyclical sources generated from the dividend-consumption comovement (\( b \)) and the macroeconomic uncertainty (\( n \)), respectively. Hence, in my model, the procyclical comovement risk counteracts the countercyclical volatility risk and the endogenous countercyclical price of risk, rendering the equity claim less risky. In addition, this economy generates a more volatile price-dividend ratio than a standard CC model due to the additional variability introduced by the additional amount-of-risk state variables (\( b \) and \( n \)). The relationship between the price-dividend ratio and the new state variable \( b \), controlling for the pricing effects of the other two state variables in this economy, is positive through a dominant cash flow channel: the persistent procyclical consumption-dividend comovement results in persistent procyclical cash flow volatility, which gets capitalized in equity prices.

As an important byproduct, this theoretical framework allows me to study the relative importance of time-varying price and quantity of risk in price variability. According to both empirical and simulated datasets, I find that the two amount-of-risk state variables (\( b \) and \( n \)) jointly explain about 30% of the fitted price-dividend ratio variance, leaving the only price-of-risk state variable (\( s \)) the dominant source.

To further support the procyclical comovement channel, I provide direct evidence for the pricing of the consumption-dividend comovement risk in the cross section, controlling for market excess returns and innovations to the other two state variables. Using Fama and MacBeth (1973) regressions for the 25 size- and book-to-market–sorted portfolios of Fama and French (1993), I find a significant and positive price of consumption-dividend comovement risk, consistent with the theory. That is, stocks covarying more with aggregate dividend risk are riskier, because the dividend risk is procyclical. Growth stocks exhibit significantly lower (or even negative for the Large-Growth bin) \( b \) loadings than value stocks; this model thus explains 75% of the value premium.

The outline of the paper is as follows. Section 2 replicates the main empirical finding in Duffee (2005) and examines the cyclicality of the exogenous and endogenous amount-of-risk components. Section 3 formulates and estimates the new DGP. Section 4 analyzes a variant of Campbell and Cochrane’s habit formation model that accommodates the Duffee Puzzle. Section 5 provides the cross-sectional evidence. Concluding comments are offered in Section 6.
2 The Duffee Puzzle Revisited, Econometrically

The decomposition of the amount of consumption risk, as shown in Equation (1), yields an exogenous conditional covariance, \( \text{Cov}_t(\Delta d_{t+1}, \Delta c_{t+1}) \), and an endogenous conditional covariance, \( \text{Cov}_t(r^m_{t+1} - \Delta d_{t+1}, \Delta c_{t+1}) \). In this section, I exploit a bivariate dynamic dependence model in the GARCH class in a flexible way to replicate the Duffee Puzzle and identify the cyclical of the two conditional comovements that constitute the puzzle.

2.1 The Model

The empirical analysis uses four variables as follows: consumption growth, the change in the log monthly consumption level, \( \Delta c_{t+1} = \log(C_{t+1}) - \log(C_t) \); dividend growth, \( \Delta d_{t+1} = \log(D_{t+1}) - \log(D_t) \); the log market return, \( r^m_{t+1} = \log \left( \frac{P_{t+1} + D_{t+1}}{P_t} \right) \); and the difference between the log market return and dividend growth (namely, the non-dividend part of the market return), \( r^m_{t+1} - \Delta d_{t+1} \). First, I project each series or its filtered counterpart (see Section 3) onto an exogenous business cycle indicator (1=recession, 0=non-recession) to obtain the series residuals. Consider a bivariate system,

\[
\tilde{\epsilon}_{t+1} = \tilde{\epsilon}_{1,t+1} \tilde{\epsilon}_{2,t+1}',
\]

where, in this paper, \( \tilde{\epsilon}_{1,t+1} \) is the consumption growth residual (denoted by \( \tilde{\epsilon}_{c,t+1} \)) from \( t \) to \( t+1 \), and \( \tilde{\epsilon}_{2,t+1} \) is either the market return residual (\( \tilde{\epsilon}_{rm,t+1} \)), the dividend growth residual (\( \tilde{\epsilon}_{d,t+1} \)), or the non-dividend part residual (\( \tilde{\epsilon}_{rmd,t+1} \)). The conditional variance-covariance matrix of the residuals is defined as,

\[
H_t = E_t \left[ \tilde{\epsilon}_{t+1} \tilde{\epsilon}_{t+1}' \right].
\]

I follow Engle (2002) and express \( H_t \) in a quadratic form in order to estimate the conditional variances (diagonal elements) and the conditional correlation (off-diagonal elements) in two separate steps,

\[
H_t = \Lambda_t \text{Corr}_t \Lambda_t,
\]

where the diagonal terms of \( \Lambda_t \) (2 \times 2) are the square roots of the conditional variances of \( \tilde{\epsilon}_{1,t+1} \) and \( \tilde{\epsilon}_{2,t+1} \) and the off-diagonal terms of \( \Lambda_t \) equal 0, or \( \Lambda_t \Lambda_t' = \begin{bmatrix} h_{1,1} & 0 \\ 0 & h_{2,2} \end{bmatrix} \) (discussed in Section 2.1.A); \( \text{Corr}_t \) (2 \times 2) is the conditional correlation matrix (discussed in Section 2.1.B).

2.1.A Conditional Volatility

The empirical literature lacks consensus about how to model the dynamics of consumption and dividend growth volatility, although researchers have provided empirical evidence for heteroskedasticity (e.g., Kandel and Stambaugh, 1990; Lettau, Ludvidgon, and Wachter, 2008) and non-Gaussianity (e.g., Bekaert and Engstrom, 2017). In contrast, the evidence for heteroskedas-
ticity and non-Gaussianity in market returns is usually found to be strong and robust. Here, I consider four conditional variance models (in the GARCH class) that identify the cyclicality of the conditional variance within the model and test for heteroskedasticity and non-Gaussianity.

The first conditional variance model assumes constant variances that are allowed to be different during recession and non-recession periods. Suppose the residual follows a conditional Gaussian distribution, \( \tilde{\epsilon}_{t+1} \sim N(0, h_t) \) and the conditional variance follows a process,

\[
h_t = \overline{h} (1 + q_t),
\]

where \( \overline{h} \) denotes the predetermined unconditional variance and the process of \( q_t \) is a multiple of the standardized NBER recession indicator (denoted as \( SNBER_t \)) so that the average conditional variance, \( E(h_t) \), is \( \overline{h} \),

\[
q_t = \nu SNBER_t,
\]

where \( \nu \) is a scalar. The zero-mean business cycle indicator, \( q_t \), is the key variable to identify the cyclicality within the model. A positive (negative) coefficient estimate of \( \nu \) suggests a countercyclical (procyclical) conditional variance; a zero estimate fails to reject the null of a constant variance.

In the second and third models, the conditional variances follow autoregressive conditional heteroskedastic processes where the long-run conditional means are allowed to be different during recession and non-recession periods,

\[
h_t = \overline{h} (1 + q_t) + \alpha \left[ \tilde{\epsilon}_{t}^2 - \overline{h} (1 + q_{t-1}) \right] + \beta \left[ h_{t-1} - \overline{h} (1 + q_{t-1}) \right],
\]

where \( \alpha + \beta < 1, \alpha > 0, \beta > 0; \) \( \overline{h} (1 + q_t) \) denotes the long-run conditional mean of the conditional variance and \( q_t \) was introduced in Equation (6). The second and third models impose different distributional assumptions: the second model assumes a conditional Gaussian distribution, \( \tilde{\epsilon}_{t+1} \sim N(0, h_t) \), and the third model assumes a symmetric leptokurtic conditional Generalized Error Distribution, \( \tilde{\epsilon}_{t+1} \sim GED(0, h_t, \tau) \) where the shape parameter \( \tau \) determines the thickness of both tails. In particular, a zero \( \nu \) estimate reduces Equation (7) to a GARCH model (Bollerslev, 1987) in the second model or a GED-GARCH model (Nelson, 1991) in the third model. The first model is a special case of the second model.

The fourth model is introduced to account for conditional asymmetry. I adapt the “Bad Environment-Good Environment” (BEGE) framework in Bekaert, Engstrom, and Er-
Bekaert, Engstrom, and Ermolov (2015) to include a long-run conditional mean that depends on the cycle variable $q_t$. The residual is a composite shock with two centered gamma shocks: 

$$
\tilde{\epsilon}_{t+1} = \sigma_{cp}\tilde{\omega}_{cp,t+1} - \sigma_{cn}\tilde{\omega}_{cn,t+1}
$$

where $\tilde{\omega}_{cp,t+1} \sim \tilde{\Gamma}(\tilde{\nu},1)$ denotes a centered homoskedastic gamma shock governing the right-tail skewness and $\tilde{\omega}_{cn,t+1} \sim \tilde{\Gamma}(cn,1)$ denotes a centered heteroskedastic gamma shock governing the left-tail skewness (given the minus sign); $\tilde{\nu} > 0$ and $cn > 0$ denote the shape parameters of the two independent gamma shocks, respectively; $\sigma_{cp}$ and $\sigma_{cn}$ denote the scale parameters ($> 0$), and the conditional variance and unscaled skewness are $h_t = \sigma_{cp}^2\tilde{\nu} + \sigma_{cn}^2cn_t$ and $skew_t = 2\sigma_{cp}^3\tilde{\nu} - 2\sigma_{cn}^3cn_t$, respectively. Thus, $cn_t$ drives the time variation in the total conditional variance, and has the following process:

$$
cn_t = \overline{\nu}(1 + q_t) + \alpha_{cn}\left[\frac{\tilde{\nu}^2}{2\sigma_{cn}^2} - \overline{\nu}(1 + q_{t-1})\right] + \beta_{cn}\left[cn_{t-1} - \overline{\nu}(1 + q_{t-1})\right], \tag{8}
$$

where $\alpha_{cn} + \beta_{cn} < 1$, $\alpha_{cn} > 0$, $\beta_{cn} > 0$; “$\overline{\nu}(1 + q_t)$” denotes the long-run conditional mean of the downside uncertainty where $\overline{\nu} > 0$ is the long-run unconditional mean. Note that, as in Bekaert, Engstrom, and Ermolov (2015), the squared residual is scaled by the squared scale parameter of the respective gamma distribution. Because the dynamics of the shape parameter $cn_t$ depend on the observed residual $\tilde{\epsilon}_t$ and not the latent gamma shock $\tilde{\omega}_{cn,t}$, the model remains in the GARCH class without requiring filtering these gamma shocks.

### 2.1.B Conditional Correlation

I follow Engle (2002) to model the conditional correlation matrix $\text{Corr}_t$ in Equation (4) with a quadratic form, $(Q_t^*)^{-1}Q_t(Q_t^*)^{-1}$, where $Q_t^*$ is the diagonal matrix with the square roots of the diagonal elements of $Q_t$ on the diagonal (so the diagonal entries of $\text{Corr}_t$ are strictly equal to 1). The off-diagonal element of $\text{Corr}_t$ is the conditional correlation (or equivalently, the conditional covariance) of the standardized residuals, $z_{t+1} = \begin{bmatrix} z_{1,t+1} & z_{2,t+1} \end{bmatrix}' = \Lambda_t^{-1}\tilde{\epsilon}_{t+1}$.

The present model differs from Engle (2002)’s dynamic conditional correlation model (DCC), who assumes a constant long-run mean of the dynamic correlation, and from Colacito, Engle, and Ghysels (2011), who use a weighted average of past correlations to model the long-run conditional mean. Instead, the DCC-$q_t$ model proposed here models the long-run conditional mean as a linear function of an exogenous business cycle indicator. Thus, the DCC-$q_t$ model directly tests the cyclicity of the conditional correlation between the two variables-of-interest:

$$
Q_t = Q_{12}^* \begin{bmatrix} 1 & 1 + q_t \\ 1 + q_t & 1 \end{bmatrix} + \alpha_{12}z_{t}z_{t}' - Q_{12}^* \begin{bmatrix} 1 & 1 + q_{t-1} \\ 1 + q_{t-1} & 1 \end{bmatrix} + \beta_{12}Q_{t-1} - Q_{12}^* \begin{bmatrix} 1 & 1 + q_{t-1} \\ 1 + q_{t-1} & 1 \end{bmatrix}, \tag{9}
$$

where the parameter $Q_{12}^*$ denotes the the predetermined constant conditional correlation between the standardized residuals. Note that Engel (2002)’s DCC model is a special case with $\nu = 0$. To summarize, the bivariate GARCH DCC-$q_t$ model captures correlation clustering as
observed in data, while the long-run conditional mean links the conditional correlation to the business cycle.

2.2 Data

The empirical part of the paper involves four key variables: consumption growth, dividend growth, the equity market return, and the non-dividend part of the market return. I follow Duffee (2005) to use monthly data indexed with \( t \). The sample spans the period from January 1959 to June 2014. Monthly real consumption per capita is defined as the sum of seasonally adjusted real aggregate expenditures on nondurable goods and services divided by monthly estimates of population (source: U.S. Bureau of Economic Analysis, BEA). Note that the realized deflators for aggregate nondurable and services consumption are different (source: BEA). Monthly seasonally adjusted dividend and earnings per market share are collected by Shiller (1989) and available on his website. The number of market shares is obtained by dividing the monthly total market value by the S&P 500 index (source: Center for Research in Security Prices, CRSP). Hence, monthly nominal dividends (earnings) per capita are calculated by multiplying the monthly dividends (earnings) per market share with the number of market shares and then dividing the aggregate dividends (earnings) by the monthly estimates of population. I use changes in the log Personal Consumption Expenditures (PCE) to calculate monthly real dividend (earnings) per capita. Monthly real consumption (dividend or earnings) growth is defined as log-differenced real consumption (dividend or earnings) per capita. The market return is defined as the change in the log market index including dividends (source: CRSP) minus the change in the log PCE. The non-dividend part of the market return is the difference between the market return and dividend growth.

It is well-known that measured aggregate consumption data are flow data which are reported as total consumption over an extended period; this temporal aggregation results in a non-zero autoregressive coefficient of aggregate consumption growth (Working, 1960) even if the true consumption growth is i.i.d.. The temporal aggregation effect could also potentially induce biases in the estimated conditional covariances.

Therefore, I follow Duffee (2005) and construct a measure of monthly consumption growth removes the autoregressive terms up to the third order, 

\[
\Delta c_{t+1} - \sum_{i=1}^{3} \phi_i (\Delta c_{t+1-i} - \bar{c})
\]

where \( \phi_i \) is the \( i^{th} \)-order autoregressive coefficient and \( \bar{c} \) is the unconditional mean.

In earlier versions of the paper, I use a wide set of instruments to approximate business conditions such as output growth, the employment rate and changes in the yield spread. In this version, I keep the NBER recession indicator as the only instrument for simplicity without loss of economic significance.

Duffee (2005), 1691-1694, provides a thorough discussion on why the purely contemporaneous covariance between returns and consumption growth underestimates the true covariance.

In all the Tables and Figures of the current section, “consumption growth” refers to this new measure that controls for the temporal aggregation issue. In Tables OA1, OA4~OA6 of the Online Appendix, I replicate the main results using one-period consumption growth (\( \Delta c_{t+1} \)) and AR(1)-de-meaned consumption growth (\( \Delta c_{t+1} - \phi_1 (\Delta c_t - \bar{c}) \)) and two de-meaned dividend growth measures to provide a comprehensive set of robustness checks.
2.3 Estimation Methodology and Cyclicality Inferences

2.3.A A Two-Stage Procedure

Many dynamic covariance models in the GARCH class (such as the bivariate model in Engle (2002) and the multivariate model in Engle and Kelly (2012)) are estimated using a two-stage quasi-maximum likelihood (QML) estimator. Bollerslev and Wooldridge (1992) and White (1994) show that, under standard regularity conditions, the quasi-maximum likelihood estimator is still consistent and asymptotically Gaussian when a Gaussian log likelihood is maximized even though the distributional assumption of Gaussianity is violated. Thus, the log quasi-likelihood of the dynamic covariance model can be written as the sum of a volatility part and a correlation part (as is true for the log likelihood of a Gaussian model). Therefore, it is customary to maximize the sum of log quasi-likelihoods of individual conditional variance models in the first stage, and maximize the log quasi-likelihood of the bivariate conditional correlation model in the second stage, given the first-stage estimation results of conditional variances.

Here, I modify the two-stage QML estimation methodology and estimate the four conditional variance models of each residual series using the maximum likelihood estimation (MLE) methodology with the actual density functions. For each residual series, the best conditional variance estimate according to the Bayesian Information Criteria (BIC) is selected to standardize the residuals for use in the second-stage estimation. Note that the four conditional variance models impose different time-series and distributional assumptions, and identifying the correct conditional distribution and volatility model is important for the theoretical modeling of the new DGP in Section 3.

Then, the second stage appeals to the QML asymptotic theory to estimate the conditional covariance/correlation of the standardized residuals in two dynamic dependence models: the DCC model and the DCC-$q_t$ model as introduced in Section 2.1.B. The standard errors of the quasi-maximum likelihood estimators are calculated following Engle and Sheppard (2001).

2.3.B Cyclicality Inference

Recall that the primary goal of my empirical framework is to parameterize the dynamic comovement processes in a flexible way so as to simultaneously identify the cyclicality of the comovements between consumption growth and (1) market returns (i.e., the Duffee Puzzle), (2) dividend growth (i.e., the exogenous component), and (3) the non-dividend part of market returns (i.e., the endogenous component).

To test the cyclicality of the relevant correlations, I use the estimation results of parameter $\nu$ within the model, and conduct Wald and Likelihood Ratio tests using post estimation inference. By design, the DCC model is the null hypothesis of the DCC-$q_t$ model with the cyclicality coefficient $\nu$ equal to 0. The probability distribution of both test statistics is a $\chi^2$ distribution with a unit degree of freedom. Lastly, given the estimates of the conditional correlations and variances from the empirical model, conditional covariances and betas are obtained ex post. The regression coefficients of the implied covariances and betas on the NBER indicator
provide direct tests on their cyclicality. This inference may differ from the t-test because the variance process itself may induce cyclicality.

### 2.4 Empirical Analysis

In this section, I begin with a discussion of the first-stage conditional variance estimation results. Then, I report the second-stage conditional comovement estimation results on the decomposition of the Duffee Puzzle.

#### 2.4.A Conditional Variance Estimation Results

Table 1 presents the first-stage estimation results of the univariate conditional variance models for each of the four variables in the empirical framework—consumption growth (Panel A), market return (Panel B), dividend growth (Panel C) and the non-dividend part of the market return (Panel D). First, the empirical evidence supports that the conditional variances of all the four variables are heteroskedastic with the conditional models in all panels outperforming the unconditional models (according to both the BIC and AIC criteria). Moreover, in all panels, the best models feature a long-run mean depending on the cyclical indicator \( q_t \), supporting cyclicality.

According to Panel A of Table 1, the conditional variance of consumption growth behaves countercyclically, given the significant and positive coefficient estimates of the standardized NBER recession indicator in all four models. The best \( q_t \)-conditional model (“GED-GARCH, \( q_t \)” with BIC = -5808.60) identifies a smaller cyclicality coefficient estimate \( \hat{\nu} = 0.0428, \ SE = 0.0099 \) than the unconditional model \( \hat{\nu} = 0.1014, \ SE = 0.0066 \) because the GARCH process already captures some cyclical variations in the conditional variance through the squared residuals. This significant cyclicality coefficient estimate indicates that the long-run conditional mean of monthly consumption growth volatility reaches as high as 0.0034 (annualized=0.0117) during recession periods and as low as 0.0032 (annualized=0.0110) during normal periods. The finding of countercyclical consumption growth volatility is in line with the literature (see, Kandel and Stambaugh, 1990; Bansal and Yaron, 2004; Bekaert and Engstrom, 2017). Moreover, given that the GED-GARCH model assumes a symmetric distribution with fat tails and that this model outperforms other conditional variance models, excess kurtosis is a salient feature of the consumption growth innovations.

The conditional variance of market returns also exhibits countercyclical behavior, which is consistent with the literature (see, e.g., Bollerslev, Engle, and Wooldridge, 1988; Schwert, 1989; Hamilton and Lin, 1996). According to Panel B of Table 1, the cyclicality coefficient estimate in the best model (“GED-GARCH, \( q_t \)” is significant and positive \( \hat{\nu} = 0.6935, \ SE = 0.1084 \). In economic terms, the long-run conditional mean of market volatility varies between 0.0616 (annualized=0.2138) during NBER recession periods and 0.0310 (annualized=0.1100) during non-recession periods; the monthly unconditional market volatility in the sample is 0.0374 (annualized=0.1298).

The market return contains a dividend part \( \Delta d \) and a non-dividend part \( r_m - \Delta d \).
The conditional variance of dividend growth is found to be (weakly) procyclical, according to the best model (“BEGE-\(n_t\)-GARCH, \(q_t\)”) in Panel C of Table 1. The cyclicality coefficient is estimated to be -0.1114 (SE=0.0592) in the best model, indicating that the long-run conditional mean of dividend growth volatility varies between 0.0058 (annualized=0.0201) during recession periods and 0.0069 (annualized=0.0241) during normal periods.

On the non-dividend part of the market return, the estimation results in Panel D of Table 1 indicate strong evidence for countercyclical conditional volatility, given that all the cyclicality coefficient estimates in all the four models are significant and positive. In particular, the significant and positive coefficient estimate in the best model (“GED-GARCH, \(q_t\)”), 0.6689 (SE=0.0991), indicates that the long-run conditional volatility is around 0.0611 (annualized=0.2117) during NBER recession periods, and around 0.0311 (annualized=0.1077) during non-recession periods; the sample monthly volatility is 0.0375 (annualized=0.1299).

Table 2 displays the detailed estimation results of the best conditional variance models for each of the four variable residuals. Apart from the main result on cyclicality, the consumption growth variance is found to be highly persistent (\(\alpha + \beta = 0.9985\)). Consumption growth (Panel A), the market return (Panel B) and the non-dividend part of the market return (Panel D) are best fitted with leptokurtic distributions of which the shape parameters (\(\tau\)) are between 1 and 2. Note that the generalized error distributions allow for tails that are either heavier than normal (when \(\tau < 2\)) or lighter than normal (when \(\tau > 2\)). The best model for dividend growth (“BEGE-\(n_t\)-GARCH, \(q_t\)”) features conditional non-Gaussianity including time-varying skewness. This is not surprising as the unconditional scaled skewness of monthly dividend growth is -1.2271 (SE=0.5145).

2.4.B The Duffee Puzzle Revisited

As illustrated in Equation (1), variation in the amount of risk (the conditional covariance between market returns and consumption growth) are driven by variation in either the amount of dividend risk (the conditional covariance between dividend growth and consumption growth) or the amount of non-dividend risk (the conditional covariance between the non-dividend part of market returns and consumption growth). In this section, I identify the source of procyclicality in the Duffee Puzzle by formally examining the cyclicality of the three comovements. I report the core results in Table 3.

I first replicate the Duffee Puzzle using a longer sample. The cyclicality coefficient estimate in the consumption-return correlation is significant and negative (\(\hat{\nu} = -0.1539\), SE=0.0359), according to the third column of Table 3, Panel A. In other words, the long-run conditional mean of the consumption-return conditional correlation is higher during non-recession periods (at 0.2147) and decreases during recession periods (to 0.1251), and the difference is statistically significant. The Likelihood Ratio test indicates that the less restrictive DCC-\(q_t\) model fits the data significantly better than the more restrictive DCC model; namely, a constant long-run conditional correlation is rejected with a p-value of 0.329%. Given the conditional correlation estimates and the conditional volatility estimates, the conditional covariance and the conditional
beta are calculated. To test the cyclicality of the other two comovement measures, Panels B and C of Table 3 report the regression results of the model-implied conditional covariance and beta on the NBER indicator. Both are procyclical and the beta significantly so. This is particularly surprising for the conditional covariance (between consumption growth and market returns) because, as established before, both the conditional volatilities of market returns and consumption growth are countercyclical. Thus, these results support the main findings in Duffee (2005)—procyclical conditional correlation and covariance between consumption growth and market returns—using 13 additional years of data.

Next, I examine the consumption-dividend comovements. According to the fourth and the fifth columns of Table 3, both the Wald test and the LR test reject a constant long-run mean of the consumption-dividend conditional correlation at a significance level smaller than 1%. The cyclicality coefficient estimate is significant and negative ($\hat{\nu}=-0.7999, \text{SE}=0.0987$), which implies procyclical behavior of the conditional correlation—the main result of this article. The magnitude of $\hat{\nu}$ implies that the long-run conditional correlation can drop to close to zero during bad times and increase to around 0.3039 during good times, with the sample correlation of the residuals being 0.2298. Given the estimated conditional correlation from Panel A and the estimated conditional variances from Table 2, Panels B and C of Table 3 show strong evidence that the conditional covariance and the conditional beta are both procyclical, with the procyclicality of the conditional beta stronger due to the countercyclicality of the consumption growth volatility in the denominator.

To support the claim that the exogenous component is the source of procyclicality of the puzzle, I complement the analysis on the exogenous component with an analysis on the endogenous component. The empirical evidence shows that the endogenous component behaves countercyclically. According to the seventh column of Panel A of Table 3, the long-run conditional correlation between the non-dividend part of the market return and consumption growth covaries positively with the countercyclical NBER recession indicator ($\hat{\nu}=0.0445, \text{SE}=0.0221$). In economic terms, the long-run conditional mean of the conditional correlation increases to around 0.2363 during recession periods and decreases to 0.2090 during non-recession periods, with the long-run unconditional mean being 0.2128 during the sample period; this difference is statistically significant. The Wald test and the LR test confirm this result at significance levels of 5% and 10% respectively. Panels B and C show that the conditional covariance and the conditional beta are both strongly countercyclical with this countercyclical endogenous component in the amount of risk counteracting the procyclical exogenous component. The procyclicality of the return-consumption correlation ($\hat{\nu}=-0.1539, \text{SE}=0.0359$) is naturally weaker than the procyclicality of the consumption-dividend correlation ($\hat{\nu}=-0.7999, \text{SE}=0.0987$), comparing the third and the fifth columns of Table 3. Analogously, Panel C shows the market return-consumption beta w.r.t. the NBER indicator to be less negative than the dividend-consumption beta.

One interesting implication of these findings is that the share of the exogenous component in the Duffee Puzzle varies greatly through time. Figure 1 depicts the implied exogenous (solid black line) and endogenous (dashed red line) components in the top plot, and the time-varying
share of the exogenous component in the amount of risk in the bottom plot. The weights of the two counteracting forces vary drastically over the business cycle. In particular, the share of dividend risk in the total amount of risk is procyclical. Given my estimates, the exogenous part explains on average 13.79% (SE=1.12%) of the total risk during the sample period. Its share drops to 0.54% during recession periods, which is statistically significantly lower than the share (16.58%) during non-recession periods (t stat=-6.35); it becomes as high as 106% during the 1960s expansion and as low as -89% following the 1969-70 recession.

2.4. C Robustness

I conduct three robustness checks of the main empirical finding in this paper—the procyclical comovement between consumption and dividend growth. All estimation results can be found in the Online Appendix.

First, I use the best conditional variance estimates without the NBER indicator to standardize the residual series innovations that are used in the second-stage estimation. According to Table OA3 of the Online Appendix, the cyclicality coefficient estimate remains significant and negative ($\hat{\nu}=-0.5544$, SE=0.1298), indicating a procyclical conditional correlation; analogously, the conditional covariance and beta coefficients are also significant and negative, indicating procyclicity.

Second, I use different consumption growth data. In particular, I conduct the same two-stage estimation procedure using (1) the original consumption growth and (2) the AR(1)-de-meaned consumption growth. Again, the procyclicality result is shown to be robust (see Tables OA4~OA6 of the Online Appendix).

Third, while many consumption-based asset pricing papers use dividend data to measure the actual cash flows received by the representative agent investing in equities, Longstaff and Piazzesi (2004) propose to use earnings data to proxy for aggregate economic dividends. Using earnings instead of dividend growth, I continue to find a procyclical comovement with consumption growth ($\hat{\nu}=-1.0875$, SE=0.2314; LR=15.41, p-value=0.009%), according to Table OA7 of the Online Appendix. Earnings growth—which is not smoothed over time—and consumption growth comove procyclically because of the dividend part of earnings growth. Log earnings growth has two components, dividend growth and the change in the log payout ratio. The comovement between the non-dividend part of earnings growth and consumption growth is acyclical ($\hat{\nu}=0.1734$, SE=1.0125). Therefore, my results consistently explain the procyclical comovement between cash flow growth and consumption growth through dividend smoothing.

2.5 Summary of the Empirical Part of the Paper

This section has confirmed Duffee (2005)’s main finding that the amount of risk behaves procyclically using a longer sample, despite several major recessions during this period. I then provide strong and robust evidence that it is the procyclical exogenous component (covariance between consumption and dividend growth) that accounts for the procyclicity in the amount of risk, given the endogenous component being strictly countercyclical. Figure [ ] illustrates the
core of the empirical section. To build a DGP for the joint dynamics of consumption and dividend growth, it is useful to summarize the results of this section in 10 stylized facts:

(a). The conditional variance of $\Delta c$ is countercyclical. \textit{Kandel & Stambaugh (1990)}

(b). The conditional variance of $\Delta d$ is procyclical. \textit{New}

(c). The conditional correlation between $\Delta c$ and $\Delta d$ is procyclical. \textit{New}

(d). The conditional covariance between $\Delta c$ and $\Delta d$ (i.e., dividend risk) is procyclical. \textit{New}

(e). The conditional sensitivity of $\Delta d$ to $\Delta c$ is procyclical. \textit{New}

(f). The conditional variance of $r^m - \Delta d$ is countercyclical. \textit{New}

(g). The conditional variance of $r^m$ is countercyclical. \textit{Schwert (1989)}

(h). The conditional covariance between $\Delta c$ and $r^m - \Delta d$ is countercyclical. \textit{New}

(i). The conditional covariance between $\Delta c$ and $r^m$ (i.e., amount of risk) is procyclical. \textit{Duffee (2005)}

(j). The share of dividend risk in the total amount of risk varies procyclically. \textit{New}

Note that some stylized facts are implied by other two facts; for example, (a) and (d) immediately implies (e), and (d) and (h) immediately implies (j).

3 A New DGP for the Joint Consumption-Dividend Dynamics

State-of-the-art consumption-based asset pricing models tend to assume unrealistic joint dynamics between consumption and dividend growth. In a Lucas tree economy (Lucas, 1978), dividends equal consumption. Most of the literature since then has separated the modeling of consumption from dividends. Often, these two processes are modeled as unit root processes with constant correlations. For example, Campbell and Cochrane (1999) assume constant co-movement (and variances). Bansal and Yaron (2004) assume a zero consumption-dividend conditional co-movement because their model imposes independence among all level and uncertainty shocks. Bansal, Kiku, and Yaron (2012) allow the dividend growth innovations to have a constant exposure to the consumption shock. However, it can be easily shown that all three consumption-dividend co-movement measures (correlation, covariance and beta) are not procyclical in their models.

In Table 4, I report a thorough evaluation of the (in)abilities of seven extant representative consumption-based asset pricing models to match the key empirical facts established in the present research (see Section 2.5): three models fall within the habit-formation workhorse framework, i.e. Campbell and Cochrane (1999), and four long-run risk workhorse framework, i.e. Bansal and Yaron (2004). Note that there is a large literature on improving the performance of these two workhorse models; however, these seven models especially focus on accommodating more realistic dynamics of fundamental shocks, which is in line with the theme of this article. I find that all models fail to match the procyclical consumption-dividend co-movement and the Duffee Puzzle, i.e. Facts (c, b, i).\textsuperscript{8}

\textsuperscript{8}All three habit-formation models show the potential to fit the countercyclical endogenous component (Fact (h)), whereas all four long-run risk models fail to fit this empirical fact. This advantage becomes one of the key
In this section, I formulate a parsimonious DGP for the consumption-dividend joint dynamics with only two state variables that has the potential to fit all stylized facts analytically, which is then verified empirically. In addition, I am able to assign economic interpretations to these state variables and their shocks using actual data and direct evidence, which enhances the plausibility of the new DGP. While the ultimate goal is to accommodate stylized facts into an asset pricing model, the GARCH-class dynamic dependence model in Section 2 is not appealing because it involves at least four state variables to generate realistic consumption-dividend joint dynamics.

3.1 The New DGP

Consumption and dividend growth have the following joint dynamics:

\[
\Delta c_{t+1} = \bar{c} + \sigma_c \tilde{\omega}_{c,t+1} + \sigma_n \bar{n}_{t+1}, \tag{10}
\]

\[
n_{t+1} = (1 - \phi_n) \bar{n} + \phi_n n_t + \sigma_{nn} \tilde{\omega}_{n,t+1}, \tag{11}
\]

\[
\Delta d_{t+1} = \bar{d} + \phi_d (V_{c,t} - \bar{V}_c) + b_t \sigma_c \tilde{\omega}_{c,t+1} + \sigma_d \bar{d}_t, \tag{12}
\]

\[
b_{t+1} = (1 - \phi_b) \bar{b} + \phi_b b_t + \lambda_b \sigma_c \tilde{\omega}_{c,t+1}, \tag{13}
\]

\[
V_{c,t} = \sigma_c^2 + \sigma_n^2 n_t, \tag{14}
\]

\[
\bar{V}_c = \sigma_c^2 + \sigma_n^2 \bar{n}, \tag{15}
\]

where consumption and dividend growth are observables; the two latent state variables in the system are the macroeconomic uncertainty \( n_t \) and the sensitivity of dividend growth to consumption growth \( b_t \) (which is new to the literature). Both state variables are assumed to follow autoregressive processes. \( V_{c,t} \) denotes the conditional variance of consumption growth, which is a linear function of \( n_t \), and \( \bar{V}_c \) represents the average conditional variance. A constant parameter \( \bar{\tau} \) denotes the unconditional mean of process \( x_t \), and \( \phi_c \) denotes the conditional mean feedback of process \( x_t \) to itself or another variable.

This model allows time variation in expected dividend growth that may help generate a negative effect of macroeconomic uncertainty on valuation ratios through a cash flow effect \((\phi_d < 0)\), but assumes a constant mean in the consumption growth equation. According to Table OA1 of the Online Appendix, regressing AR(3)-de-meaned consumption growth on a NBER recession indicator delivers an insignificant coefficient of -0.0008 (SE=0.0005), whereas the corresponding coefficient for dividend growth is significant and negative, -0.0042 (SE=0.0008). Therefore, the conditional mean specifications are consistent with the data. The autoregressive coefficients for the state variables \((\phi_n \text{ and } \phi_b)\) are expected to be positive.

The new DGP features three mutually independent shocks. The consumption “fundamental shock”, \( \tilde{\omega}_{c,t+1} \), is a Gaussian shock with unit standard deviation; the consumption “event

---

The reasons why this paper focuses on a variant of the Campbell-Cochrane model to accommodate the Duffee Puzzle in Section 2

They are (1) consumption and (2) dividend variances, \( h_{c,t} \) and \( h_{d,t} \), (3) the consumption-dividend correlation, \( Corr_t \), and (3) the time-varying long-run conditional mean of the conditional correlation, \( q_t \).
shock”, \( \tilde{\omega}_{n,t+1} \), follows a centered heteroskedastic gamma distribution with a strictly positive shape parameter \( n_t \) (using a subscript to denote that \( n_t \) varies over time) and a unit scale parameter; the dividend-specific shock, \( \tilde{\omega}_{d,t+1} \), follows a centered homoskedastic gamma distribution with a strictly positive shape parameter \( V_d \) and a unit scale parameter. That is,

\[
\tilde{\omega}_{c,t+1} \sim i.i.d. N(0, 1); \quad \tilde{\omega}_{n,t+1} \sim \Gamma(n_t, 1) - n_t; \quad \tilde{\omega}_{d,t+1} \sim \Gamma(V_d, 1) - V_d.
\]

In particular, the probability density function for \( \tilde{\omega}_{n,t+1} \), denoted \( f(\tilde{\omega}_{n,t+1}) \), is given by,

\[
f(\tilde{\omega}_{n,t+1}) = \frac{1}{\Gamma(n_t)}(\tilde{\omega}_{n,t+1} + n_t)^{n_t-1} \exp \left( -\tilde{\omega}_{n,t+1} - n_t \right),
\]

for \( \tilde{\omega}_{n,t+1} > -n_t \) and with \( \Gamma(\cdot) \) representing a complete gamma function. The moment generating function of \( \tilde{\omega}_{n,t+1} \), denoted \( M(\kappa) \equiv E[\exp (\kappa \tilde{\omega}_{n,t+1})] \), is \( \exp \left[ -\kappa n_t - \ln (1 - \kappa) n_t \right] \). It can be easily shown that the conditional mean of \( \tilde{\omega}_{n,t+1} \) is zero, the conditional variance \( n_t > 0 \), the conditional unscaled skewness \( 2n_t > 0 \), and the conditional unscaled excess kurtosis \( 6n_t > 0 \). Similar results hold for the distribution of \( \tilde{\omega}_{d,t+1} \) with \( V_d \) governing the shape of the distribution.

Gamma shocks are not as commonly used as Gaussian shocks in the literature, but they are appealing for my purposes for two reasons. First, they help fit the evidence of thick and skewed tails found in consumption and dividend growth residuals (see Section 2). Second, higher-order moments of gamma-distributed variables can be expressed as linear functions of their shape parameters (see above), which helps produce neat and tractable analytical solutions given that this paper focuses on second (cross) moments. It turns out that, conditional on the constant shock sensitivity parameters having the appropriate signs (\( \sigma_c, \sigma_{nn}, \lambda_b > 0; \sigma_n, \sigma_d < 0 \)), the new DGP has the ability to match all five stylized facts about the consumption-dividend joint dynamics (Facts (a)~(e) as documented in Section 2) and match their distributional properties.

To see this, let me first discuss the consumption system. Each period, consumption growth responds positively to the “Gaussian” fundamental shock (\( \sigma_c > 0 \)) and negatively to the “gamma” event shock (\( \sigma_n < 0 \)). Given the moment generating functions of Gaussian- and gamma-distributed shocks and that the two shocks are assumed to be independent, the total consumption growth variance has an analytical solution, \( \sigma_c^2 + \sigma_n^2 n_t \) in which \( \sigma_c^2 \) captures the “fundamental-shock” variability and \( \sigma_n^2 n_t \) captures the “event-shock” variability. With \( \tilde{\omega}_{n,t+1} \) right-skewed, a negative \( \sigma_n \) implies that the event-shock component of the consumption innovation provides a source of heteroskedasticity that is coming from the left tail of consumption growth (with large negative events). It is a realistic and parsimonious assumption because Bekaert and Engstrom (2017) who model consumption innovations with both positively- and negatively-skewed gamma shocks find that the shape parameter of the positively-skewed gamma shock is large (or, the shock is Gaussian-like) and time-invariant (or, the shock is homoskedastic). In addition, the negative \( \sigma_n \) enables the new DGP to generates countercyclical consumption growth volatility.

\[\text{It is noteworthy that a gamma shock is always right-skewed, and thus the minus gamma shock is left-skewed.}\]
Fact Check (a): $n_{t+1}$ is countercyclical, given that $\text{Cov}[\Delta c_{t+1}, n_{t+1}] = \sigma_n \sigma_{nn} n_t < 0$. Thus, the conditional variance of consumption growth, $\sigma^2_c + \sigma^2_{nn} n_t$, is countercyclical.

Dividend growth loads on the fundamental consumption shock with a time-varying beta ($b_t$), and is further influenced by a homoskedastic left-skewed gamma disturbance (given that $\sigma_d$ is negative and $\tilde{\omega}_{d,t+1}$ is right-skewed). The time-varying $b_t$ is procyclical as $\sigma_c$ and $\lambda_b$ are strictly positive constants:

Fact Check (e): $b_{t+1}$ is procyclical, given that $\text{Cov}[\Delta c_{t+1}, b_{t+1}] = \lambda_b \sigma^2_c > 0$.

The conditional variance of dividend growth, $b_t^2 \sigma^2_c + \sigma^2_d V_d$, increases with $b_t$ (if $b_t > 0$), which immediately implies a procyclical dividend growth variance. Given the countercyclicality of the consumption volatility state variable $n_t$ above, the new dividend growth process is thus modeled conveniently to generate a strictly procyclical comovement between dividend growth and consumption growth (i.e., correlation, covariance, and beta).

Fact Check (b): the conditional variance of dividend growth, $b_t^2 \sigma^2_c + \sigma^2_d$, is procyclical.

Fact Check (c): the conditional correlation between dividend and consumption growth, $\frac{b_t \sigma^2_c}{\sqrt{\sigma^2_c + \sigma^2_{nn} n_t}} \frac{\sqrt{\sigma^2_c + \sigma^2_{nn} n_t}}{b_t^2 \sigma^2_c + \sigma^2_d V_d}$, is procyclical given a countercyclical $n_t$ and a procyclical $b_t$.

Fact Check (d): the conditional covariance between dividend and consumption growth, $b_t \sigma^2_c$, is procyclical.

3.2 DGP Estimation Results

In this section, I first present the estimation results for the new DGP. To enhance the plausibility of the new DGP, I then discuss the economic interpretation of the latent shocks in Section 3.2.B, the evidence on fitting the cyclicality with over-identification in Section 3.2.C, and the possibility of other DGPs that might satisfy the 5 exogenous facts in Section 3.2.D. The parameter estimates are used when exploring the theoretical economy later in Section 4.

3.2.A Parameter Estimation Results of the New DGP

Given that there is no feedback from dividend growth dynamics to consumption growth dynamics, I estimate the consumption growth system ($\{\Delta c, n\}$) and the dividend growth system ($\{\Delta d, b\}$) in a two-step estimation procedure. During the first step, the consumption growth system is estimated using a filtration-based approximate maximum likelihood methodology developed by Bates (2006) to obtain parameter estimates, realizations of the latent macroeconomic uncertainty state variable ($\{\hat{n}_t\}_{t=1}^{T}$) and realizations of the two consumption shocks ($\{\hat{\omega}_{c,t}\}_{t=2}^{T}$, $\{\hat{\omega}_{n,t}\}_{t=2}^{T}$). Given the first-step estimation results, the dividend growth system is estimated using the maximum likelihood methodology to obtain the remaining parameter estimates, estimates of the comovement state variable ($\{\hat{b}_t\}_{t=1}^{T}$) and the dividend-specific shock ($\{\hat{\omega}_{d,t}\}_{t=2}^{T}$). I provide a detailed description of the estimation procedure in Appendix A.

---

11 In this paper, “$\hat{x}$” indicates an estimate of the unknown parameter/variable/shock $x$. 

17
According to Panel A of Table 5, consumption growth depends positively on the symmetric homoskedastic fundamental shock ($\hat{\sigma}_c=0.0029$, SE=0.0001) and negatively on the right-skewed heteroskedastic event shock ($\hat{\sigma}_n=-0.0023$, SE=0.0005). Moreover, the sensitivity of the macroeconomic uncertainty state variable $n_t$ to the event shock is positive, $\hat{\sigma}_{nn}=0.2772$ (SE=0.1027), which analytically implies a countercyclical $n_t$ as it covaries negatively with consumption growth. Therefore, the first-step estimation results deliver countercyclical consumption growth volatility, i.e. Fact (a); I defer its graphical evidence (Figure 2) to Section 3.2.B.

The estimation results of the dividend growth system show that the sensitivity of $\hat{b}_t$ to the fundamental shock is significant and positive ($\hat{\lambda}_b=14.0978$, SE=1.3764), which immediately implies that the conditional beta $b_t$ is procyclical given that $\hat{\text{Cov}}_t\left(\Delta c_{t+1}, \hat{b}_{t+1}\right) = \hat{\lambda}_b\hat{\sigma}_c^2 > 0$. Thus, the conditional covariance and correlation between dividend and consumption growth are also procyclical, given the analytical solutions derived before. Next, because the $\hat{b}_t$ estimates are positive (see Figure 3), admitting an one-to-one correspondence between $\hat{b}_t$ and $\hat{b}_t^2$, the conditional variance of dividend growth—that increases with $b_t^2$—is also procyclical. Therefore, the parameter estimation results demonstrate the ability of the new DGP to fit Facts (b)∼(e) analytically.

In addition, I find the unconditional distribution of $\hat{\omega}_{d,t+1}$ to be right-skewed (skewness=0.3070, SE=0.0751; Panel C of Table 5), and the estimate of $\sigma_d$ is significant and negative ($\hat{\sigma}_d = -0.0008$, SE=1.88E-05). Because the consumption fundamental shock is a Gaussian shock, the dividend-specific disturbance ($\sigma_d\hat{\omega}_d$) is designed to capture the strong negative skewness of dividend growth innovations as observed in data (skewness=-1.2271, SE=0.5145) and the conditional evidence in Table 1 featuring time-varying negative conditional skewness. Therefore, the new DGP has the potential to match the distributional properties of the dividend growth innovations.

### 3.2.B Economic Interpretation

Here, I analyze the time variation in the estimated state variables (\(\hat{n}\) and \(\hat{b}\)) in Figures 2 and 3. Moreover, given that the DGP shock structure plays a crucial role in simultaneously matching all five stylized facts, I compare the time variation in the filtered shocks ($\hat{\omega}_n$ and $\hat{\omega}_c$) against various business cycle indicators in Figure 4 to motivate their economic interpretation and thus to enhance the plausibility of the new DGP.

Figure 2 depicts the two estimated components of the total consumption growth variability, $\hat{\sigma}_c^2$ contributed by the fundamental shock and $\hat{\sigma}_n^2\hat{n}_t$ contributed by the event shock. While the fundamental shock clearly plays a dominant role in explaining the total consumption growth variability during normal periods, the event shock accounts for as high as 57.95% of the total variance during NBER recessions. For example, the largest spikes occurred during the 1973 oil crisis and during the oil crisis followed by Volcker’s monetary policy tightening in the early 1980s. The third largest spike occurred in the early 1960s, again coinciding with an NBER recession. During the recent 2007-08 financial crisis, the fraction of the total consumption variance explained by the event shock soared to around 34%, but did not exceed the three largest
spikes mentioned above. These spikes also clearly show up in the quarterly shocks graphed in the bottom panel of Figure 4.

The filtered homoskedastic consumption shock explains (on average) 82.29% of the total consumption growth variability during the sample period. The correlation between the fundamental shock and the NBER recession indicator is -0.182*** at the monthly frequency and -0.2703*** at the quarterly frequency. It immediately follows that this shock is procyclical. The top panel of Figure 4 depicts filtered fundamental shocks aggregated to a quarterly frequency. The shock is consistently negative during NBER recessions, with the largest negative shocks occurring during the 2007-08 recessions.

The fundamental shock is the only shock that determines both the consumption and dividend growth innovations. According to Panel D of Table 5, I find that the detrended quarterly consumption-wealth ratio from Lettau and Ludvigson (2001) (source: Martin Lettau’s website) has a significant and negative correlation (-0.215*** with the filtered fundamental shock, but is uncorrelated with the filtered event shock (0.0561) and the dividend-specific shock (0.0350). The top panel of Figure 4 illustrates the negative correlation between the fundamental shock and the detrended $\hat{cay}$. Therefore, the fundamental shock is not only a procyclical shock but is also positively correlated with the wealth-consumption ratio. Because the dividend-consumption beta, $b_t$, is persistent (with a half life of 5.5 months), a unit fundamental shock at time $t$ increases $b_t$ and will have persistent effects on expected future dividend-consumption comovement and cash flow variance. It is the persistent effects that get capitalized in financial wealth, inducing a higher wealth-consumption ratio.

Several extant models in the consumption-based literature have modeled consumption growth disturbances with two independent shocks. The continuous-time model in Longstaff and Piazzesi (2004) models the consumption growth innovation with a Brownian motion (analogous to the Gaussian fundamental shock here) and a jump process (analogous to the gamma event shock here). My model is also related to the “BEGE” model in Bekaaert and Engstrom (2017) featuring two independent shocks, one associated with the “good” volatility and the other one “bad” volatility. Segal, Shaliastovich, and Yaron (2015) adapt this model in a long-run risk framework. However, my model also ensures the fit of realistic dynamics of dividend volatility and dividend-consumption comovements, which potentially improves all three models above.

Lastly, Figure 3 depicts the heteroskedastic part ($\hat{b}_t^2\hat{\sigma}_c^2$) and the homoskedastic part ($\hat{\sigma}_d^2\hat{V}_d$) of the total dividend growth variance. Consumption shocks explain on average 3.18% of the total dividend variance.

3.2.C Fitting the Cyclicality: Evidence and Over-identification

Even though the empirical evidence regarding cyclicality in Section 2 pertains to dynamic conditional moments, I replicate it here in terms of moments calculated during recession and non-recession periods. Specifically, I simulate the DGP for 100,000 months given the shock

---

12 The quarterly shocks are calculated as the sum of monthly shocks in the same quarter; quarterly aggregation is necessary for computing correlations with Lettau and Ludvigson (2001)’s cay variable which is only available at the quarterly (or lower) frequency (see Panel D of Table 5).
distributional assumptions and the parameter estimates, and then test the closeness between sample moments and simulation moments during recession and non-recession periods.

Because the empirical recession dummy variable uses the NBER recession indicator which (according to the NBER White Paper) is created based on patterns in GDP growth, I therefore develop an algorithm to identify recession patterns in the consumption growth—the only macroeconomic variable in consumption-based asset pricing models (see Appendix B for a detailed description). When applying this algorithm to consumption growth during January 1959 – June 2014, it identifies seven out of the eight NBER recessions. More formally, the consumption-based recession indicator is highly correlated with the actual NBER recession indicator at 0.80, and projecting the consumption-based recession indicator onto the actual NBER recession indicator recessions produces a coefficient of 0.9038 (SE=0.0507), which is insignificantly different from 1.

In Table 6, I calculate cyclical data moments regarding the five stylized facts established in Section 2 on consumption and dividend growth (i.e., Facts (a)∼(e)). The evidence for the procyclicality of the comovement (correlation, covariance, and beta) between dividend and consumption growth remains strong in the data moments, which is consistent with the strong and robust empirical evidence found in the conditional framework in Section 2. Column “M(3)” in Panel A shows that all the model moment point estimates calculated from both recession and non-recession periods are within 95% confidence intervals of the data moments. The DGP-implied moments also look economically very close to the data moments, with a few exceptions. For example, the correlation between dividend and consumption growth is 0.0639 in comparison to 0.0148 in the data. Given an imperfect recession identification scheme, such gap is understandable.

In Table 7, I investigate the fit of the DGP with respect to a number of other data moments as an over-identification test. These moments include the means, standard deviations, skewness, and excess kurtosis of consumption and dividend growth (8), the heteroskedastic nature of their innovations (2), and their unconditional comovement moments (3). According to Column “M(3)” of Table 7, these moments are matched statistically well except for the skewness and excess kurtosis of dividend growth, which are underestimated.

3.2.D On the Uniqueness of DGPs Accommodating the Duffee Puzzle

Duffee (2005) uses a simplified two-asset model with constant discount rate and growth rates to show that consumption growth is more positively correlated with stock returns when stock market wealth is relatively more important in determining consumption than other asset wealth—this is what he calls the “composition effect”. Therefore, his framework essentially suggests a procyclical component in the consumption variance.

Inspired by the Duffee model, one intuitive alternative (to the new DGP) is to assume a “constant” exposure of dividend growth to the consumption fundamental shock but assign this consumption shock a “procyclical” conditional variance, whereas the current model assumes

\[3^{13}\text{In short, “M(3)” denotes the model in this paper; I defer further explanation to Section 4.3.}\]
a procyclical exposure and a homoskedastic fundamental shock. Then, each period, the consumption growth innovation in this alternative DGP receives a heteroskedastic Gaussian shock with procyclical volatility and a heteroskedastic gamma shock with countercyclical volatility. This alternative model can in theory generate a procyclical consumption-dividend covariance, correlation and beta.

However, this alternative DGP generates two problems. First, the identification of consumption growth variance becomes harder. Analytically, the consumption growth conditional variance is now the sum of a procyclical component (through the fundamental shock) and a countercyclical component (through the event shock), which makes the ultimate cyclicality unclear. However, both the empirical evidence in Section 2 and the recession/non-recession sample moments in this section suggest strongly countercyclical consumption growth variance. From the viewpoint of an econometrician, given that a Gaussian distribution is symmetric and not bounded, the heteroskedastic fundamental shock might act as the event shock to try to fit the left-tail events in the estimation. Hence, the estimation results might generate countercyclical fundamental shock volatility and thus countercyclical consumption-dividend comovement, which contradicts the empirical findings. Restricting the fundamental shock volatility to be procyclical (e.g., restricting the signs of certain parameters) could resolve the technical problem; then, the estimation results obtained from a constrained estimation usually become harder to interpret, which makes this alternative DGP less appealing.

Second, and more importantly, this alternative DGP likely generates a positive correlation between consumption and dividend growth variances because the heteroskedastic fundamental shock now positively explains the dynamics of both variances. However, the best GARCH-class conditional variance estimation results suggest that the consumption growth variance is weakly negatively correlated with dividend growth variance ($\rho = -20.41\%$, SE = 3.82\%). This negative correlation is consistent with Fact (a), countercyclical consumption growth variance, and Fact (b), procyclical dividend growth variance. The new DGP nails this moment: according to the estimation results, my model implies a correlation of -13.11\%, which is within the 95\% confidence interval of the data moment.

4 An External Habit Model

In this section, I explore how incorporating more realistic dynamics into the amount of risk affects the performance of the extant dynamic asset pricing models. In Section 4.1 I describe the assumptions regarding the preferences of the representative agent. In Section 4.2 I provide approximate analytical solutions of asset prices that demonstrate various features of the model. In Section 4.3 I confront the numerical solution of the model with a wide range of empirical moments.

The key moments to be matched by the model are the cyclicity of the two components of the Duffee Puzzle. The procyclical conditional covariance between dividend and consumption growth (or the exogenous component) is immediately satisfied given the new DGP (see
Section 3). However, different consumption-based asset pricing models have different implications regarding the cyclicality of the conditional covariance between the non-dividend part of the market return and consumption growth (or the endogenous component). The (external) habit-formation framework naturally entails a countercyclical endogenous component through countercyclical risk aversion; the effects of consumption shocks on the valuation ratio are amplified during bad times when risk aversion is higher. On the other hand, the long-run risk framework assumes that shocks (i.e., consumption shock, dividend shock, volatility shock, long-run expected growth shock) are independent of one another, resulting in a zero covariance between the price dividend ratio and consumption growth and thus a zero endogenous component. In addition, I show later that an endowment economy with procyclical dividend risk requires a countercyclical price of risk to generate realistic cyclical behavior for the price dividend ratio and the equity premium. As a result, in this paper, I consider the more natural external habit paradigm as developed in Campbell and Cochrane (1999).

4.1 Habit-Based Preferences

In this economy, the representative agent maximizes:

$$E_0 \left[ \sum_{t=0}^{\infty} \beta^t \frac{(C_t - X_t)^{1-\gamma} - 1}{1-\gamma} \right],$$

(17)

where $C_t$ is the real consumption level and $X_t$ the external habit level at time $t$ ($C_t > X_t$). The parameter $\beta$ is the time discount factor, and $\gamma$ is the curvature parameter. The relative risk aversion (the local curvature of the utility function) is $\gamma \frac{C_t}{C_t - X_t} \equiv \gamma \tilde{S}_t$, where the surplus consumption ratio $S_t$ is defined as the percentage gap between consumption and habit level. When the consumption level is closer to the habit level (i.e., lower surplus consumption ratio), the agent becomes more risk averse (instantaneously). The log surplus consumption ratio $s_t$, log($S_t$), follows an AR(1) process with its shock structure perfectly spanned by the two aforementioned consumption shocks:

$$s_{t+1} = (1 - \phi_s)\bar{s}_t + \phi_s s_t + \lambda_t (\sigma_c \tilde{\omega}_{c,t+1} + \sigma_n \tilde{\omega}_{n,t+1}),$$

(18)

where $\phi_s$ is the persistence coefficient of $s_t$, $\bar{s}_t$ the time-varying long-run mean, and $\lambda_t$ the sensitivity function. Note that $\bar{s}_t$ is assumed constant in the CC model, but countercyclical in my model (see Section 4.2.A). The stochastic discount factor (SDF) is the ratio of marginal utilities, $M_{t+1} = \beta \frac{U_{C_t}(C_{t+1}, X_{t+1})}{U_C(C_t, X_t)} = \beta \left( \frac{C_{t+1}S_{t+1}}{C_tS_t} \right)^{-\gamma}$; the log real pricing kernel is,

$$m_{t+1} = \ln(\beta) - \gamma \Delta c_{t+1} - \gamma \Delta s_{t+1}$$

$$= \ln(\beta) - \gamma (1 - \phi_s)(\bar{s}_t - s_t) - \gamma (1 + \lambda_t) (\sigma_c \tilde{\omega}_{c,t+1} + \sigma_n \tilde{\omega}_{n,t+1}).$$

(19)
4.2 Asset Prices

In this section, I present salient features of asset prices implied by the model with (quasi) analytical solutions.

4.2.A The Risk Free Rate and the Sensitivity Function

The risk free rate, \( rf_t \), is solved from the usual first-order condition for the consumption-saving ratio, \( rf_t = \ln \{ E_t[\exp(m_{t+1})]\}^{-1} \). Given the moment generating functions of the two independent shocks in the kernel process (Equation (19)), the risk free rate has an exact closed-form solution,

\[
rf_t = -\ln \beta + \gamma \tau + \gamma (1 - \phi_s)(\pi_t - s_t) - \frac{1}{2} \gamma^2 (1 + \lambda_t)^2 \sigma_c^2 - [\gamma (1 + \lambda_t) \sigma_n - \ln (1 + \gamma (1 + \lambda_t) \sigma_n)] n_t.
\]

(20)

Similar to the Campbell and Cochrane model, one source for time variation in the risk free rate is the “intertemporal substitution effect”. When risk appetite is low (\( s_t < \pi_t \)) and thus expected to be high in the future, the agent borrows to smooth marginal utility, driving up the interest rate. Another source is the “precautionary savings effect”, when the agent wants to save more during volatile periods, driving down the interest rate.

The literature proposes various ways of specifying the sensitivity function, which plays an important role in the Campbell and Cochrane (1999) model as it represents the price of consumption risk. First, in Campbell and Cochrane (1999), the two effects perfectly cancel out, rendering the risk free rate constant; however, a constant short rate is counterfactual. Second, in Wachter (2005, 2006), the intertemporal substitution effect dominates in order to generate an upward sloping real yield curve and a positive bond risk premium. This framework results in a countercyclical risk free rate—which is counterfactual given that Ang, Bekaert, and Wei (2008) find the U.S. real risk free rate to be procyclical. Third, Bekaert and Engstrom (2017) propose a time-varying risk free rate such that the relative importance of the two effects varies over time. Their risk free rate also depends on \( n_t \).

I propose a fourth way delivering a risk free rate that is strictly procyclical (i.e., the precautionary savings effect dominates). The sensitivity function is solved such that the second-order Taylor approximation of Equation (20) is a constant, which is referred to as the \( rf_{CC} \) component of the exact risk free rate (see Equation (24) below or Appendix C):

\[
\lambda_t = \begin{cases} 
\frac{1}{s_t} \sqrt{1 - 2(s_t - \pi_t)} - 1, & s_t \leq s_{\text{max},t} \\
0, & s_t > s_{\text{max},t}
\end{cases}
\]

(21)
where \( \eta_t = \log(\mathcal{S}_t) \) and \( s_{\text{max},t} \) are derived as functions of the free parameters and \( n_t \),

\[
\mathcal{S}_t = \sqrt{\left(\sigma_c^2 + \sigma_n^2 n_t\right) \frac{\gamma}{1 - \phi_s}}, \tag{22}
\]

\[
s_{\text{max},t} = \eta_t + \frac{1}{2}(1 - \mathcal{S}_t^2). \tag{23}
\]

Note that \( \mathcal{S}_t \) is defined endogenously to CC’s \( \mathcal{S} \) with the only difference that the consumption growth variance varies through time. The \( s_{\text{max},t} \) variable is likewise the time-varying equivalent to the expression in CC. The dynamics of the sensitivity function are thus determined by \( s_t \) and \( n_t \). The surplus consumption ratio state variable \( s_t \) has an intuitive negative effect on \( \lambda_t \) (as in the Campbell and Cochrane model); when the consumption level is closer to the habit level, the price of risk increases. The uncertainty state variable \( n_t \) has a negative effect on \( \lambda_t \) through \( s_t \) (as also assumed in Bekaert and Engstrom, 2017) and a positive effect through \( s_t \).

With this sensitivity function, a third order Taylor approximation of the risk free rate is given by:

\[
rf_t \approx -\ln \beta + \gamma e - \frac{(1 - \phi_s)\gamma}{2} + \frac{1}{3} \gamma^3 (1 + \lambda_t)^3 \frac{\sigma_n^3}{\sigma_n} n_t. \tag{24}
\]

In this expression, the constant term “\( rf_{CC} \)” is identical to the risk free rate in Campbell and Cochrane (1999). The appended “precautionary savings” term, \( \frac{1}{3} \gamma^3 (1 + \lambda_t)^3 \frac{\sigma_n^3}{\sigma_n} n_t \), is determined by the countercyclical sensitivity function (ultimately driving the price of risk) in Equation (21) and the countercyclical macroeconomic uncertainty. Because \( \sigma_n \) is negative according to Table 5, the appended term and the risk free rate are strictly procyclical. The actual risk free rate has a negative and procyclical nonlinear term appended to \( rf_{CC} \) instead of the approximate cubic term above (see Appendix C).

4.2.B Approximate Analytical Solution for Equity Prices

The endowment economy features three state variables: the procyclical risk appetite (\( s_t \)), the countercyclical macroeconomic uncertainty (\( n_t \)), and the procyclical consumption-dividend comovement (\( b_t \)). Campbell and Cochrane (1999)’s state variable, \( s_t \), ensures variation in the price of risk; the two new state variables, \( n_t \) and \( b_t \), introduce dynamics into the amount of risk. Note that the model does not have exact closed-form solutions and thus is formally solved with numerical methods (see Section 4.3). Nevertheless, approximation analytical solutions help

---

14The cyclicity of the state variables can be easily proved. Risk appetite (introduced in Section 4.1) is procyclical because \( \text{Cov}(s_{t+1}, \Delta c_{t+1}) = \lambda(s_t)(\sigma_c^2 + \sigma_{n}^2 n_t) > 0 \). As discussed in Section 3, macroeconomic uncertainty is countercyclical because \( \text{Cov}(n_{t+1}, \Delta c_{t+1}) = \sigma_n \sigma_m n_t < 0 \), and consumption-dividend comovement is procyclical because \( \text{Cov}(b_{t+1}, \Delta c_{t+1}) = \sigma_b^2 b_t > 0 \).
provide economic intuitions.

First, I conjecture an approximate process for the log valuation ratio $pd_t \equiv \ln \left( \frac{P_t}{D_t} \right)$,

$$pd_t = A_0 + A_1 s_t + A_2 b_t + A_3 b_t^2 + A_4 n_t. \quad (25)$$

The analytical solutions rely on two additional approximations. First, I apply the Campbell and Shiller linearization to the log market return, $r_{mt+1} = \ln \left( \frac{P_{t+1}}{P_t} + D_{t+1} \right) \approx \Delta d_{t+1} + a_1 pd_{t+1} + a_0$ where $a_0$ and $a_1$ are linearization constants that only depend on the average level of $pd$. The approximate market return is used in the Euler equation. Second, given the shock assumptions and the $pd$ conjecture, there are three types of shocks in the log market return: a Gaussian shock, a $\chi^2(1)$ shock, and a gamma shock. I then use a quasi quadratic Taylor approximation to the Euler equation: $1 = E_t \left[ \exp(m_{t+1} + r_{mt+1}) \right] \approx \exp \left[ E_t(m_{t+1} + r_{mt+1}) + \frac{1}{2} V_t(E(m_{t+1} + r_{mt+1}^2)) \right]$ (see Appendix D). Conditional on these three approximations, the coefficients in the conjectured price dividend ratio are solved in closed form in Appendix E. The state variables affect the price dividend ratio via a discount rate (DR) channel and/or via a cash flow (CF) channel:

[1]. The risk aversion effects: $A_1 > 0$.

When risk appetite (risk aversion) is low (high), the required compensation per unit of consumption risk to investing in risky assets increases; hence, the price dividend ratio decreases as the risk compensation demanded increases.

Risk aversion has another DR channel, operating through the interest rate. As shown in Equation (24), the precautionary savings effect on the interest rate is amplified when risk aversion is high. Thus, interest rate decreases with risk aversion, driving down the total return demanded; hence, price the dividend ratio increases. This effect is, however, dominated by the risk premium effect above, given the parameter choices and numerical solutions in Section 4.3.

[2]. The comovement effects (New): $A_2, A_3 > 0$.

Through a pure CF channel, the price dividend ratio can be interpreted as reflecting the outlook on future dividend growth. The persistent procyclical $b_t$ induces a persistent procyclical dividend growth variance, which gets capitalized in equity prices. Analytically, the expected value of the exponential of dividend growth increases with both the expected growth and conditional variance$^{15}$ The conditional variance component has a closed-form solution that strictly increases with $b_t^2$. Therefore, this pure CF channel suggests a positive relationship between $b_t$ and $pd_t$.

However, there is a potentially countervailing risk premium effect. The total risk premium to compensate changes in dividend growth can be intuitively approximated with $-Cov_t(m_{t+1}, \Delta d_{t+1}) = \gamma(1 + \lambda_t) b_t \sigma_c^2$. The compensation for cash flow risk increases with both $b_t$ and $\lambda_t$. When a positive fundamental shock arrives, $b_t$ and $s_t$ increase and $\lambda_t$ decreases simultaneously. If $\lambda_t$ was not countercyclical, the model would generate a higher risk premium, potentially resulting in a counterintuitive negative relationship between the procyclical $b_t$ and the procyclical $pd_t$. Given

$^{15}$The Gaussian analogue is, $E_t[\exp(\Delta d_{t+1})] = \exp \left[ E_t(\Delta d_{t+1}) + \frac{1}{2} V_t(\Delta d_{t+1}) \right]$. 25
the parameter choices and numerical solutions, this habit formation model has the ability to generate a positive \( pd_t \sim b_t \) relationship. This is another reason why the habit formation paradigm is preferred in comparison with the long-run risk framework (which assumes a constant price of risk).

[3]. The macroeconomic uncertainty effects: \( A_4 > 0 \).

The CF channel of uncertainty is well-recognized. When macroeconomic uncertainty \( (n_t) \) is higher, future dividend growth is expected to be lower, driving down the current price—which is in the spirit of the long-run risk story. On the other hand, a higher \( n_t \) induces more precautionary savings, driving down the interest rate and lowering the total return demanded; this DR channel of \( n_t \) also appears in Bekaert and Engstrom (2017). Given my parameter choices, the DR channel dominates the CF channel in the current model during almost all periods.

Given these results, I verify Facts (f) and (g) regarding the cyclicity of price-dividend ratio and market return variances.

**Fact Check (f) and (g):** The log market return in the approximate analytical solution is \( \Delta d_{t+1} + a_1 pd_{t+1} - pd_t + a_0 \). Given the dividend growth dynamics in the new DGP and the price dividend ratio conjecture, the conditional variances of the log price dividend ratio and the log market return have the following approximate expressions,

\[
\begin{align*}
\text{Var}_t(pd_{t+1}) & \approx \varsigma_{pd} + \varsigma_1 \lambda_t + \varsigma_2 b_t + \varsigma_3 n_t + \varsigma_4 \lambda^2_t + \varsigma_5 b^2_t + \varsigma_6 \lambda_t b_t + \varsigma_7 \lambda_t n_t + \varsigma_8 \lambda^2_t n_t, \\
\text{Var}_t(r^n_{t+1}) & \approx \varsigma_{rm} + a^2_1 \varsigma_1 \lambda_t + [a^2_1 \varsigma_2 + 2 a_1 \lambda_b \varsigma^2_c (A_2 + 2 A_3(1 - \phi_b) \bar{b})] b_t + a^2_1 \varsigma_3 n_t + a^2_1 \varsigma_4 \lambda^2_t + (a^2_1 \varsigma_5 + 2 a_1 \varsigma_2) \lambda^2_t b_t + (a^2_1 \varsigma_6 + 2 a_1 \varsigma_1) \lambda^2_t n_t + a^2_1 \varsigma_7 \lambda^n_t n_t + a^2_1 \varsigma_8 \lambda^2_t n_t, 
\end{align*}
\]

where \( \varsigma_{pd}, \varsigma_{rm}, \varsigma_1, \varsigma_2, \varsigma_3, \varsigma_4, \varsigma_5, \varsigma_6, \varsigma_7, \varsigma_8, a_1, \lambda_b, \sigma_c, \phi_b \) and \( \bar{b} \) are strictly positive constants, and \( \varsigma_7 = 2 A_1 A_4 \sigma_n \sigma_{nn} \) is negative when the discount rate effect of \( n_t \) dominates its cash flow effect, and positive vice versa. Thus, the model has the potential to generate a countercyclical price dividend ratio (a linear proxy for the variance of the non-dividend part of the market return) and market return variances.

4.2.C The Duffee Puzzle Revisited, Theoretically

The empirical evidence in Section 2 finds that the exogenous component \( \text{Cov}_t(\Delta d_{t+1}, \Delta c_{t+1}) \) is procyclical, the endogenous component \( \text{Cov}_t(r^n_{t+1} - \Delta d_{t+1}, \Delta c_{t+1}) \) countercyclical, and the amount of risk \( \text{Cov}_t(r^n_{t+1}, \Delta c_{t+1}) \) procyclical. In particular, the latter two empirical facts are important testable hypotheses in evaluating the theoretical model.

According to the approximate analytical solution (see Appendix E), the model-implied amount of risk contains a procyclical exogenous component as assumed in the new DGP; the model-implied endogenous component contains three parts from the three state variables re-
spectively:

\[
\begin{align*}
\text{1. Exogenous Component} & \quad b_t\sigma_c^2 \\
\text{2. Endogenous Component: benchmark amount of risk in Campbell and Cochrane (1999)} & \quad a_1 A_1 \lambda_t \sigma_c^2 \\
& \quad + \left[a_1 A_2 \lambda_t + 2a_1 A_3 (1 - \phi_b) b_t \lambda_t + 2a_1 A_3 \phi_b \lambda_t b_t \right] \sigma_n^2 \\
\text{3. Endogenous Component: additional amount of risk induced by comovement} & \quad + a_1 \left[A_1 \lambda_t \sigma_n^2 + A_4 \sigma_m \sigma_n \right] n_t \\
\text{4. Endogenous Component: additional amount of risk induced by uncertainty} & \quad + a_1 \left[A_1 \lambda_t \sigma_n^2 + A_4 \sigma_m \sigma_n \right] n_t
\end{align*}
\]

where the parameters $\sigma_c, \sigma_n, \phi_b, \lambda_b$ and $a_1$ are positive and $\sigma_n$ is negative according to the DGP estimation results in Table 5. The procyclical exogenous component, Term (1), is directly obtained from the new DGP. The next three terms constitute the endogenous component. The strictly countercyclical Term (2) captures the amount of risk implied from linearizing the original Campbell and Cochrane model. As discussed in Section 4.2.B, it reflects the positive effect of risk appetite on the valuation ratio. Term (3) captures the procyclical amount of “comovement risk” that emerges from the positive sensitivity of the valuation ratio to the time-varying consumption-dividend comovement. The last term captures the amount of risk that is associated with the countercyclical macroeconomic uncertainty; the cyclicality of Term (4) is parameter-dependent, but has the potential to generate a countercyclical process.

With these expressions, I examine the model’s ability to deliver Facts (h)∼(j).

- **Fact Check (h):** The model has the potential to generate a countercyclical endogenous component if the procyclical part(s) is counteracted by the countercyclical part(s).
- **Fact Check (i):** The procyclical terms (Terms (1) and (3)) in the amount of risk expression (Equation (28)) must counteract the countercyclical sources in order to obtain procyclical amount of risk, thus resolving the Duffee Puzzle.
- **Fact Check (j):** The share of the procyclical exogenous component in the total amount of risk has the potential to be procyclical if the implied endogenous component is not strongly procyclical.

### 4.2.D The Equity Premium

The equity premium in this revised habit formation model is approximately the product of a countercyclical price of risk, $\gamma(1 + \lambda_t)$, and a time-varying amount of risk that comprises both procyclical and countercyclical sources according to Equation (28). The detailed derivation is included in Appendix E.

The interactions between the two additional risk sources from the two new state variables (the countercyclical volatility risk and the procyclical dividend risk) and the countercyclical price of risk have direct implications for the magnitude of the equity risk premium. On the one hand,
the introduction of the countercyclical volatility risk makes the asset riskier as the volatility risk is higher when risk aversion is higher during economic turmoil; from this perspective, a higher unconditional equity premium is expected. On the other hand, the introduction of the procyclical dividend risk—the core contribution of the present research—potentially lowers the level of the unconditional equity premium. This is because the amount of risk now contains a procyclical component which counteracts the countercyclical price of risk; in other words, the asset becomes less risky. Moreover, the conditional equity premium no longer monotonically increases with the price of risk.

4.3 Numerical Solutions and Results

The final model features three state variables: the habit state variable, \( s \), and two new state variables, \( n \) and \( b \). Therefore, in the numerical analysis, I introduce two intermediate models to analyze the effects of the two new state variables in the final model.

The first intermediate model “M(1)” is an adapted Campbell and Cochrane model which features homoskedastic fundamentals. To be more precise, consumption growth innovations depend on a homoskedastic Gaussian fundamental shock and a homoskedastic gamma event shock (i.e., shape parameter = \( \bar{n} \)), and dividend growth has a constant exposure to the consumption fundamental shock (i.e., the dividend beta = \( \bar{b} \)). M(1) features only one state variable, as present in Campbell and Cochrane (1999), the surplus consumption ratio. The second intermediate model “M(2)” is an adapted Bekaert and Engstrom (2017) model that builds on M(1) but incorporates countercyclical macroeconomic uncertainty as the second state variable. Thus, the final model “M(3)” can be viewed as a generalization of M(2) where the consumption-dividend comovement is now procyclical. The comparison between the numerical solutions of M(3) and M(2) (M(2) and M(1)) reveals the pricing implications of \( b_t \) (\( n_t \)). All three models price dividend claims. Appendix C presents M(1) and M(2) in detail.

In the remainder of the section, Section 4.3.A describes the calibration of the preference (non-DGP) parameters. Then, I confront the three models with a wide range of asset price statistics. Specifically, in Section 4.3.B I evaluate the model fit in terms of conventional asset price statistics and over-identify the models with additional unconditional fundamental and cross moments. In Section 4.3.C, I focus on the implications of the uncertainty and comovement state variables on fitting the 10 stylized facts revolving the Duffee Puzzle. I conclude by revisiting how the three state variables affect price dividend ratios and then their relatively importance in driving equity prices in Section 4.3.D.

4.3.A Calibration and Simulation

There are four non-DGP parameters to be determined: \( \gamma \), \( \phi_s \), \( r_{fCC} \), and \( \beta \). I fix the utility curvature parameter \( \gamma \) at 2. As commonly assumed in the literature, the AR(1) coefficient of the \( s_t \) process, \( \phi_s \), equals the AR(1) coefficient of monthly log price dividend ratio. The benchmark constant risk free rate, \( r_{fCC} \), as appeared in Equation (20), is chosen to match the average monthly real short rate proxied by the difference between the change in log nominal 90-day
Treasury index constructed by CRSP and the continuously compounded inflation rate. \( \beta \) is the time discount parameter inferred from the \( rf_{CC} \) equation. Table 8 summarizes the non-DGP parameters.

The log valuation ratios are solved numerically using the “series method” from Wachter (2005). \( M(1) \) is solved using a one-dimensional grid \((20 \times 1)\) for the one state variable: the log surplus consumption ratio; \( M(2) \) is solved over a two-dimensional grid \((20 \times 20)\) for the two state variables: the log surplus consumption ratio and macroeconomic uncertainty. The final model \( M(3) \) uses a three-dimensional grid \((20 \times 20 \times 20)\) for all three state variables.

For each model, I draw 100,000 months of fundamental shocks and then construct the state variable processes according to their data generating processes. Given the grid solutions, I apply the piecewise polynomial cubic interpolation for \( M(1) \), and the piecewise polynomial spline interpolation for \( M(2) \) and \( M(3) \) to obtain the log price dividend ratio for each simulated month given the state variable values. All the reported theoretical moments in this paper are calculated using the second half of the simulated dataset, i.e., 50,001-100,000. As mentioned in Section 3.2.C, the recession periods in the simulated dataset are identified using patterns in simulated consumption growth (see Appendix B for details).

### 4.3.B Unconditional Moments

Table 9 reports the fit of the models with respect to the unconditional moments of fundamentals and asset prices. The equity risk premium increases a lot from 4.4520% in \( M(1) \) to 5.9374% in \( M(2) \) because the time-varying macroeconomic uncertainty introduces additional countercyclical dynamics into the amount of risk, which also results in a lower price dividend ratio. The average price dividend ratio implied from \( M(2) \) is below the sample 95% confidence interval. However, the log valuation ratio variability implied from \( M(2) \), 0.2882, is not statistically significantly different from the data moment, which is an improvement to \( M(1) \) and is consistent with Bekaert and Engstrom (2017). The standard CC model and its adapted version fail to generate sufficient price dividend ratio variability.

Adding procyclical dividend risk, in \( M(3) \), the unconditional equity premium is slightly lower compared to that in \( M(2) \) with a constant dividend risk, which is consistent with economic intuition. In \( M(3) \), the amount of risk now contains a procyclical component which counteracts the countercyclical macroeconomic uncertainty in the amount of risk and the countercyclical risk aversion in the price of risk. Thus, equities are less risky in \( M(3) \). Quantitatively, the annualized equity premium drops from 5.9374% in \( M(2) \) to 5.3537% in \( M(3) \)—both not rejected by the sample counterpart (4.8780%). With procyclical dividend risk affecting the price dividend ratio positively, \( \sigma(pd) \) implied from \( M(3) \) increases to 0.3037, which is the closest to the data moment (0.3946) in this row.

The variability of the market excess return in \( M(1) \) is below the sample 95% confidence interval of the data moment (14.8816%). The fit significantly improves when introducing heteroskedasticity into consumption growth in \( M(2) \), because of the more volatile price dividend ratio as discussed earlier. Although \( M(3) \) incorporates one more state variable, the market
return implied from M(3) is less variable, \( \sigma(r^m - r_f) = 14.5441\% \), than that implied from M(2). Here is the interpretation. The unconditional variance of market return can be approximately decomposed into three components: the variance of \( pd \), the variance of \( \Delta d \), and the covariance between \( pd \) and \( \Delta d \). M(1) and M(2) by construction impose a strictly positive covariance between \( pd \) and \( \Delta d \) through their positive exposures to the consumption fundamental shock. The unconditional covariance between the two in M(3) contains an additional negative term, because \( \text{Cov}(\lambda_t, b_t) < 0 \), reflecting the procyclical dividend risk serving as an internal “buffer” against the countercyclical risk aversion in the economy\(^{16}\).

Despite failing to fit return variability, the M(1)-implied Sharpe Ratio is within the sample 95% confidence interval of the data point estimate (0.3278). The implied Sharpe Ratio from the M(3) economy, 0.3681, has the best fit among the three models. The kurtosis moment is also matched statistically well by all three models. However, the equity return skewness implied by all three models is indifferent from zero, although M(2) and M(3) do generate negative skewness.

M(2) and M(3) generate the same risk free rate dynamics because they only differ in the dividend growth processes. Their risk free rate is slightly higher than the one in M(1), but both risk free rates have means that are within the 95% confidence interval of the data point estimate, 1.4209%.

In addition, I over-identify the models with a set of 18 unconditional first, second, and cross moments of economic fundamentals and market returns. Table 7, as discussed before in Section 3 when I evaluated the fit of the new DGP, also shows that M(1) and M(2) can fit reasonably well with respect to these unconditional higher moments, with the exception of return-based volatilities.

### 4.3.C Duffee Puzzle

In Table 10, I evaluate the fit of the theoretical models with respect to the cyclical moments related to the endogenous part of the Duffee Puzzle, i.e. Facts (f)\(~\) (j). To begin with, the risk aversion channel does not suffice to explain the strongly countercyclical volatility dynamics of asset prices as I observe in data; thus, M(1) fails to match Facts (f) and (g). Nevertheless, M(1) matches the endogenous part of the puzzle, the countercyclical \( C(r^m - \Delta d, \Delta c) \) or Fact (h), which further demonstrates the advantage of a habit formation framework.

The countercyclical uncertainty \( n_t \) introduces additional countercyclicality into the variances of \( r^m \) and \( r^m - \Delta d \), which are now significantly higher during recession periods and

\(^{16}\)As a simplified version of Equation (27), the conditional covariance between the dividend growth and price changes (ignoring the quadratic term) in M(3) is

\[
\text{Cov}(pd_{t+1}, \Delta d_{t+1}) \approx A_1 \text{Cov}(n_{t+1}, \Delta d_{t+1}) + A_2 \text{Cov}(b_{t+1}, \Delta d_{t+1}) + A_4 \text{Cov}(n_{t+1}, \Delta d_{t+1}) \\
= A_1 \lambda_t b_t \sigma^2_c + A_2 \lambda_t b_t \sigma^2_c + 0. \\
\tag{29}
\]

Then the unconditional covariance becomes,

\[
E[\text{Cov}(pd_{t+1}, \Delta d_{t+1})] \approx A_1 \sigma^2_c \mathbb{E}(\lambda_t) \mathbb{E}(b_t) + A_1 \sigma^2_c \mathbb{E}(\lambda_t, b_t) + A_4 \lambda_t \sigma^2_c E(b_t). \\
\tag{30}
\]
statistically close to the data moments. Moreover, M(2) improves M(1) towards generating a more realistic endogenous component. The difference between the two $C(r^m - \Delta d, \Delta c)$ values calculated during recession and non-recession periods is $2.54 \times 10^{-5} - 2.23 \times 10^{-5} = 0.31 \times 10^{-5}$ in M(2) and $2.45 \times 10^{-5} - 2.25 \times 10^{-5} = 0.20 \times 10^{-5}$ in M(1), whereas the difference calculated using data is $2.36 \times 10^{-5} - 2.06 \times 10^{-5} = 0.30 \times 10^{-5}$ (one-sided p-value=0.039). The difference in M(2) is within the 95% confidence interval of the data moment.

Next, I discuss the implications of the comovement state variable $b_t$. The simulation results further demonstrate the ability of this procyclical exogenous component (driven by $b_t$) to counterbalance the countercyclical endogenous component in the total amount of risk and thus resolve the Duffee Puzzle. To be more specific, Table 10 Fact (i), shows that M(3) generates a higher covariance between market returns and consumption growth during non-recession periods $(C(r^m, \Delta c)(I_{rece.}=0) = 2.7672 \times 10^{-5})$ than during recession periods $(C(r^m, \Delta c)(I_{rece.}=1) = 2.6743 \times 10^{-5})$; both point estimates are not rejected by their sample counterparts. In the same row, M(1) and M(2) without the procyclical comovement state variable $b_t$ generate a countercyclical amount of risk.

Lastly, and more importantly, the share of the amount of risk explained by the exogenous component is already weakly procyclical in M(1) and M(2). This is true because the endogenous components, which enter the denominator, are countercyclical through the risk aversion channel. However, without a procyclical exogenous component, M(1) and M(2) do not fit Fact (j). More specifically, the implied shares during non-recession periods in all three models are statistically close to the sample counterparts: 15.980% in data, 14.136% in M(1), 14.248 in M(2), and 13.443% in M(3). However, the implied shares during recession periods are statistically significantly higher in M(1), 13.120%, and M(2), 12.717%, than in the data, 1.322%. Only model M(3) generates the share, 4.196%, that is within the 95% confidence interval of the data. This in turn demonstrates the economic significance of the procyclical comovement documented in this article.

4.3.D Price Dividend Ratio Dynamics

In this section, I analyze the dependence of the price dividend ratio on the three state variables both economically and quantitatively. In Figure 5 lines with squares illustrate the dependence of the price dividend ratio on $s$ ($n$) given the numerical solution of M(3), conditional on different combinations of the other two state variables $n$ and $b$ ($s$ and $b$); regarding different combinations, I consider mean and critical values for each of the other two variables. Since M(2) has only two dimensions, $s$ and $n$, I fix one when evaluating the relationship between $pd$ and the other variable as shown in solid lines with triangles. M(1) with only one dimension, $s$, is depicted in solids lines with circles.

All lines in the top plot of Figure 5 demonstrate the positive relationship between the price dividend ratio and the surplus consumption ratio, consistent with the literature and the analytical prediction earlier. Next, the dotted line with squares shows that the price dividend ratio is higher in states of high $n$ (1.4246, the 95% quantile value in the $n_t$ simulation), con-
trolling for the same $b$. This “high $n$”–“high PD” line indicates that the DR effect of $n_t$ on the valuation ratio via the risk free rate dominates the CF effect via the expected dividend growth in this model. However, both analytical (see Appendix E) and numerical (here) solutions confirm that the price dividend ratio does not have a monotonic relationship with $n$. The bottom plot of Figure 5 shows that, after certain cutoff point around $n = 4.5$, the price dividend ratio decreases with $n$; namely, the CF effect of macroeconomic uncertainty becomes dominant during periods with “extremely” bad events (i.e., 0.1% according to the simulation). This non-monotonic relationship precisely reflects the countervailing DR and CF effects of the macroeconomic uncertainty state variable, as discussed in the analytical part (see Section 4.2.B). In the top plot of Figure 5, the DR effect dominates because the 95% quantile value (1.4246) is less than the cutoff value (4.5).

Figure 6 analyzes the relationship between the price dividend ratio and $b$, conditional on $s$ and $n$. The three lines with squares show the positive relationship between the valuation ratio and the comovement state variable in M(3) by fixing $s$ and $n$ at mean or critical values, confirming the dominant CF channel mentioned in Section 4.2.B. In addition, the price dividend ratio implied from M(3) at the “low $s$”–“π” plane ($s = -3.5281, n = 0.3742$; dotted line) is lower than at the “$E(s)$”–“π” plane ($s = -2.6595, n = 0.3742$; solid line), indicating a positive relationship between $PD$ and $s$. Analogously, the price dividend ratio is higher at the “$E(s)$”–“high $n$” plane ($s = -2.6595, n = 1.4246$; dashed line) than at the $E(s)$–π plane (solid line), indicating a positive relationship between $PD$ and $n$. Both observations are consistent with Figure 5. Lastly, the M(1) and M(2) horizontal lines intersect the $E(s)$–π plane of M(3) at around $b = 0.45$, which is expected because $\hat{b} = 0.4447$.

Next, I examine the log price dividend ratio dynamics quantitatively. In Table 11, I conduct univariate and multivariate contemporaneous regressions of the log price dividend ratio on the state variables, and confront the parameter estimates and variance decomposition results in each model with their sample counterparts. In the last row, I note the $R^2$ of the univariate or multivariate regression within each model.

I obtain the empirical proxies for the three state variables as follows. I follow Wachter (2006) to construct a monthly empirical proxy for $s_t$, $\sum_{i=1}^{108} \phi_s \Delta c_{t-i}$ where $\phi_s$ at the monthly frequency is 0.9957 (Table 8) and $\Delta c_t$ is the AR(3)-de-meaned consumption growth. Because of the cumulative sum, the empirical proxy for $s_t$ (thus the sample regressions) starts in March 1968. I then scale the empirical proxy to match the mean and volatility of $s_t$ implied from the theoretical model. The monthly empirical proxies for $n_t$ and $b_t$ are obtained from the DGP estimation results in Table 5.

The univariate coefficient estimate of $pd_t$ on $b_t$ in the empirical regression is 1.2807 (SE=0.1292); economically, that means that a unit standard deviation increase in $b_t$ (0.10) is associated with $1.3$ unit standard deviation increase in $pd_t$ (0.14). The univariate model explains

\[15\] Note that the long-run mean of $s_t$ is by design time-varying in M(2) and M(3), as discussed in Section 4.1. The long-run mean of $s$ in M(1) is constant, $\pi = -2.6677$, as shown in Table 8. In Figure 6, I use the unconditional mean of $s$ simulated from M(3) as the $s$ intercept value for all five lines for consistency.

\[18\] “$\pi$” (0.3742) and “high $n$” (1.4246) are within the lower region in the bottom plot of Figure 5 where the DR effect still dominates the CF effect; thus, a positive relationship is expected.
12.900% of the total variability of $pd_t$. The multivariate model has a higher $R^2$ of 49.587%. The log surplus consumption ratio accounts for 72.779% of the fitted log price dividend ratio variability (or variance decomposition, $\text{VARC}(s_t) \equiv \hat{b}(s_t) \text{cov}(s_t, \hat{pd}_t) / \text{var}(\hat{pd}_t) = 72.779\%$ where $\hat{pd}_t$ denotes the fitted value), the macroeconomic uncertainty for 14.102%, and the consumption-dividend comovement for 13.386%. The log surplus consumption ratio $s_t$ and the procyclical comovement $b_t$ positively predict $pd_t$ in the multivariate regression, while $n_t$ has a negative coefficient. Hence, empirically, the relationship between the log valuation ratio and $b_t$ remains significant and positive after controlling for other state variables, providing empirical evidence of the CF channel as discussed in Section 4.2.B.

With the simulated dataset of M(1), the $s_t$ coefficient of 0.6693 is higher than the sample 95% confidence interval. The regression $R^2$ is 96.933% given that the M(1) economy is spanned by only one state variable; it is not 100% because the CC channel builds in a non-linearity through the sensitivity function. The bivariate regression using the simulated dataset of M(2) shows a significant and positive coefficient estimate of $s_t$ (0.5564) with a large variance decomposition percentage (69.091%). The bivariate regression of M(2) appears parsimonious to capture the additional non-linearity introduced by the uncertainty state variable, thus obtaining a lower $R^2$ of only 62.012%. However, the linear models are potentially useful to evaluate first order effects.\footnote{Note that in the long-run risk literature, it is common to use linear approximations.}

The M(3) model is the only model with the new state variable $b_t$. With the simulated dataset of M(3), the univariate coefficient estimate of $pd_t$ on $b_t$ is 1.1146, which is statistically close to the empirical coefficient estimate. Given that M(3) is a non-linear model with no exact closed-form solutions, I use the approximate linear conjecture (Equation (25)) underlying the approximate analytical solution in Section 4.2.B for the multivariate analysis. Thus, the multivariate regression model has four explanatory variables, \{s_t, n_t, b_t, b_t^2\}. The multivariate regression delivers a significant and positive coefficient estimate for $s_t$ (0.5368) with a 64.317% variance decomposition percentage; the estimate is within two standard errors of the empirical coefficient. The $b_t$ coefficient estimate is 0.8951, which is higher than the empirical coefficient (0.5480) but would not be rejected at the 1% significance level. In contrast to the empirical finding on the dominating CF effect of $n_t$ on $pd_t$ (i.e., a negative $n_t$ coefficient), the multivariate regression results using the simulated datasets of both M(2) and M(3) show a dominating DR effect (i.e., positive $n_t$ coefficients) as also seen in Figures 5 and 6. The M(3) model also underestimates the significance of $b_t$ in explaining price variability: the VARC is 6.657% using the simulated dataset of M(3), but 13.386% using actual data. The $R^2$ of the M(3) approximate model is 65.255%, suggesting that higher order terms are important in explaining price dividend ratio.

Lastly, Table 11 reveals two more economic insights. First, controlling for macroeconomic uncertainty and dynamic cash flow comovement, changes in risk aversion play a significant and dominant role (data VARC = 72.799%) in explaining the price variability even with a rough proxy. Thus, risk aversion appears to be a more economically important factor in explaining...
risky asset prices than are second moment state variables, which is a testable hypothesis for future research. Second, the data VARCs of the procyclical consumption-dividend comovement (13.386%) and the countercyclical uncertainty (14.102%) are relatively close. Thus, while the countercyclical uncertainty is well-acknowledged in the literature, the procyclical dividend risk introduced in this article appears to also have nontrivial economic significance.

5 The Cross Section of Expected Returns

Macroeconomic variables are widely-acknowledged candidates for systematic risk factors that are correlated with consumption and investment opportunities and thus maybe priced in the cross-section of expected returns (see Maio and Santa-Clara, 2015). There is a small but burgeoning literature using macro factors to explain the cross-section of expected returns (see e.g. Lettau and Ludvigson, 2001; Bansal, Dittmar, and Lundblad, 2005; Bali, Brown, and Tang, 2017).

To enhance the plausibility of the procyclical comovement state variable introduced in this article, I now examine the cross-sectional pricing abilities of the three state variables. A four-factor pricing model with market returns and the innovations to the three state variables is estimated using the Fama and MacBeth (1973) methodology and the 25 size– and book-to-market–sorted portfolios constructed by Fama and French (1993). The sample period is March 1968–June 2014. Market returns are included to acknowledge systematic financial asset risk factors that are orthogonal to macroeconomic shocks in a parsimonious way. The first four panels in Table 12 report the portfolio loadings on the four factors, the 5th bin–1st bin (“5-1”) differences along each dimension and their significance. The last panel reports the second-stage cross-sectional regression results for the prices of risk and their significance.

I find that the price of volatility risk is negative ($\hat{\lambda}_{\text{innov}}=-0.0697$, SE=0.0413; Panel E of Table 12). Investors pay for insurance against increases in macroeconomic uncertainty. In the cross section, value stocks and small stocks comove more negatively with changes in volatility than growth and large stocks and thus require higher risk premiums, whereas large growth stocks exhibit positive betas and thus provide volatility risk insurance. Note that the macroeconomic uncertainty in this paper is simply proxied by consumption growth volatility; its role in cross-sectional pricing supports the recent findings in Bali, Brown, and Tang (2017) who use the well-known economic uncertainty index of Jurado, Ludvigson, and Ng (2015). In addition, the price of comovement risk is positive ($\hat{\lambda}_{\text{innov}}=0.0474$, SE=0.0216; Panel E of Table 12)—which is consistent with the theory in the current paper. Investors consider stocks that covary more positively with aggregate dividend risk riskier, because dividend risk is procyclical. This is new to the cross-section of expected-return literature.

Given the cross-sectional findings in Table 12, value stocks exhibit higher $b_1$ loadings than growth stocks, and small stocks exhibit higher loadings than large stocks. As the extreme example in the 5×5 panel, the loading of the VALUE-SMALL portfolio on the $b_1$ innovations in the four-factor pricing model is significant and positive ($\hat{\beta}_{1,\text{innov}}=17.1760$, SE=6.3492). Fur-
thermore, I test the significance of the loading differences between the 5th bin ("VALUE") and the 1-st bin ("GROWTH"), conditional on different sizes. For portfolios with above-medium sizes, the "5-1" differences are significant and positive (8.5369* for the "Size 4" stocks, and 19.6829*** for the "LARGE" stocks). The empirical value premium is calculated as the average spread in returns between the 5th bin ("VALUE") and the 1-st bin ("GROWTH") stocks across time and across all sizes. The explained value premium through the comovement channel is calculated as the average loading difference ("5-1") multiplied by the price of comovement risk across all sizes. As an immediate implication, the comovement channel explains 75% of the value premium, that is 0.39% out of 0.51%, through the lens of this model.

This paper provides a framework to evaluate the explanatory power of the economic state variables for both time-varying price variation and cross-sectional variation in expected returns. Among the three economic factors motivated from the theoretical model, while in Section 4.3.D I find that changes in risk aversion is a dominant factor explaining aggregate price-dividend ratio variability (VARC=72.79% according to Table 11), but changes in macroeconomic uncertainty and cash flow comovement (higher-order-moment factors) account for 94.65% of the explained cross-sectional variation in expected returns.

6 Conclusion

This paper contributes to the literature in the following ways. First, it recognizes and replicates the puzzling finding in Duffee (2005) in which he provides empirical evidence for the procyclicality of the amount of risk, \( \text{Cov}_t (\Delta d_{t+1}, \Delta c_{t+1}) \). To resolve the Duffee Puzzle, I decompose the covariance into two components, providing strong and robust empirical evidence for a procyclical component, \( \text{Cov}_t (\Delta d_{t+1}, \Delta c_{t+1}) \), and a countercyclical component, \( \text{Cov}_t (r_{t+1}^m - \Delta d_{t+1}, \Delta c_{t+1}) \). Because the procyclical component dominates, the puzzle is resolved. In contrast, most of the literature assumes that the amount of risk is acyclical or strictly countercyclical. I establish 10 stylized facts (of which 7 are new) related to the Duffee Puzzle that serve as testable hypotheses for a proposed theoretical model.

Second, I formulate a new DGP for the consumption-dividend joint dynamics with a minimum number of state variables to be used in consumption-based asset pricing models. The new DGP features two new state variables: countercyclical macroeconomic uncertainty and procyclical consumption-dividend comovement. Both the analytical solutions and estimation results of this parsimonious DGP demonstrate the ability to accommodate all stylized facts related to consumption and dividend growth, which is an improvement to the existing DGPs in the literature.

Then, I solve a variant of the Campbell and Cochrane model that uses the new DGP and accommodates the Duffee Puzzle. The three state variables are the procyclical risk appetite and two new ones from the new DGP. Numerical solutions demonstrate that the revised model fits all cyclical moments related to the Duffee Puzzle, which is a contribution to the literature. I also confront the model with a broad array of unconditional fundamental moments, asset price
statistics and price dividend ratio variance decompositions to over-identify the model. In particular, the comovement state variable positively predicts the log price dividend ratio through the cash flow channel, which is quantitatively confirmed using actual data. Furthermore, a multivariate contemporaneous regression of the log price dividend ratio on the three state variables reveals that the procyclical comovement state variable explains 13% of the fitted log price dividend ratio variability. Other notable asset pricing implications are (1) a lower equity premium and (2) a more volatile price dividend ratio. The procyclical amount of dividend risk “hedges” the countercyclical price of risk and the countercyclical volatility in the total amount of risk, rendering equity less risky.

To substantiate the plausibility of the procyclical comovement as the new state variable, I also examine its ability to help price the cross section of expected returns. I find a significant and positive price of comovement risk: investors demand higher compensation from stocks that comove more positively with the consumption-dividend comovement. Value stocks have significantly higher comovement loadings than growth stocks, which explains 75% of the value premium in this sample. Second-moment state variables (countercyclical macroeconomic uncertainty and procyclical consumption-dividend comovement) account for 95% of the total explanatory power of the three state variables in explaining the cross section of asset returns, whereas risk aversion accounts for 72% of price dividend ratio variability and thus is a dominant factor in explaining the time variation in market returns.
Appendices

A  Estimation procedure for the new DGP in Section 3

The consumption-dividend dynamics in Section 3 accommodates the Duffee Puzzle and the time-varying macroeconomic uncertainty, and thus introduces several new parameters. I provide parameter choices based on a two-step estimation procedure. Consumption growth and dividend growth have the following joint dynamics:

\[
\begin{align*}
\Delta c_{t+1} &= \bar{c} + \sigma_c \tilde{\omega}_{c,t+1} + \sigma_n \tilde{\omega}_{n,t+1}, \\
n_{t+1} &= (1 - \phi_n) \pi + \phi_n n_t + \sigma_n \tilde{\omega}_{n,t+1}, \\
\Delta d_{t+1} &= \tilde{d} + \phi_d (V_{c,t} - \bar{V}_c) + b_t \sigma_c \tilde{\omega}_{c,t+1} + \sigma_d \tilde{\omega}_{d,t+1}, \\
b_{t+1} &= (1 - \phi_b) \tilde{b} + \phi_b b_t + \lambda b_t \sigma_c \tilde{\omega}_{c,t+1}, \\
V_{c,t} &= \sigma_c^2 + \sigma_n^2 n_t, \\
\bar{V}_c &= \sigma_c^2 + \sigma_n^2 \pi,
\end{align*}
\]

where the consumption fundamental shock, \(\tilde{\omega}_{c,t+1}\), is a centered Gaussian shock with standard deviation equal to 1, the consumption event shock, \(\tilde{\omega}_{n,t+1}\), follows a centered heteroskedastic gamma distribution with a strictly positive shape parameter equal to \(\sigma_n\) and a scale parameter equal to 1, and the dividend-specific shock, \(\tilde{\omega}_{d,t+1}\), follows a centered homoskedastic gamma distribution with a strictly positive shape parameter equal to \(\sigma_d\) and a scale parameter equal to 1. Or,

\[
\tilde{\omega}_{c,t+1} \sim i.i.d. N(0, 1); \quad \tilde{\omega}_{n,t+1} \sim \Gamma(n_t, 1) - n_t; \quad \tilde{\omega}_{d,t+1} \sim \Gamma(V_d, 1) - V_d.
\]

The three fundamental shocks are mutually independent. Consumption growth and dividend growths are observable; the two latent processes in the system are the macroeconomic uncertainty state variable \(n_t\) and the conditional sensitivity of dividend growth to consumption growth \(b_t\). \(V_{c,t}\) denotes the total consumption conditional variance, which is a linear function of \(n_t\), \(\pi\) denotes the unconditional mean of process \(x\), and \(\phi_d\) denotes the conditional mean feedback of process \(x\) to itself or another variable. Here is the full set of parameters in the joint dynamics: \(\{\bar{c}, \sigma_c, \sigma_n, \pi, \phi_n, \sigma_n\}\); cash flow, \(\{\tilde{d}, \phi_d, \sigma_c, \tilde{b}, \phi_b, \lambda\}\).

By design, there is no feedback from the cash flow growth process to the consumption growth process. The estimation procedure is described with two steps. The first step estimates the consumption growth system. I use a filtration-based maximum likelihood methodology of Bates (2006) to estimate the latent time series \(n_t\) and extract the two consumption shock realizations, the fundamental shock \(\tilde{\omega}_{c,t+1}\) and the event shock \(\tilde{\omega}_{n,t+1}\). The conditional consumption growth variance and its long-run average is then implied, \(\hat{V}_c\) and \(\bar{V}_c\).

The second step takes the dividend growth data, \(\Delta d_{t+1}\), and the predetermined state variable levels and shocks from the previous step, \(\hat{V}_{c,t}, \bar{V}_{c,t}, \tilde{\omega}_{c,t+1}\), and finds the cash flow parameters such that the sum of the log likelihoods of the implied cash flow-specific shock \(\tilde{\omega}_{d,t+1}\) is maximized. To provide estimation convenience without loss of statistical power, I first project dividend growth onto a vector of ones and \(\hat{V}_{c,t} - \bar{V}_{c,t}\), and obtain the estimates for \(\hat{d}, \phi_d\). Then, I use the residuals to estimate the rest of the cash flow parameters. The MLE estimation does not impose constraints on the non-negativity of \(b_t\) estimates, but imposes one constraint to ensure a valid gamma density function for \(\tilde{\omega}_{d,t+1}\) at any time stamp \(t\):

\[
-\sigma_d V_d \geq \max_{\forall t \in 1, \ldots, T} \left( \Delta d_{t+1} - \tilde{d} - \phi_d (\hat{V}_{c,t} - \bar{V}_{c,t}) - b_t \tilde{\omega}_{d,t+1} \right).
\]

A gamma distribution is right-tailed and is bounded below; \(\sigma_d\) is expected to be negative given the strongly negative dividend growth skewness; therefore, the zero-mean dividend-specific shock \(\sigma_d \tilde{\omega}_{d,t+1}\) is expected to be left-tailed with an upper bound at \(-\sigma_d V_d > 0\). The constraint above states that the maximum dividend-specific shock is within the upper bound, and thus a gamma density function at any time stamp \(t\) is defined.
B  Recession identification criteria in simulated monthly series

In the moment matching exercises, data moments and simulation moments during recession and non-recession months are calculated in Table 6. The empirical recession dummy variable uses the NBER recession indicator (from the NBER website) which is based on patterns in quarterly GDP growth. However, theoretical models in this paper do not generate output. Instead, I use patterns in consumption growth (the only macroeconomic variable in the model) to identify recessions in the simulated months. Here is the algorithm:

1. **Quarterly Growth**: aggregate the monthly consumption growth data into a quarterly frequency;
2. **Standardization**: de-mean the quarterly consumption growth by a 49-quarter moving average (24+1+24), and divide it with its long-term/unconditional standard deviation;
3. **Fundamental Cyclical Events**: identify the quarters as recessions if there are at least two consecutive standardized consumption growth drops that are \(< -0.9;\)
4. **Extreme Events**: for an extreme event (when standardized consumption growth values are \(< = -2),\) if its adjacent quarter(s) before and/or after exhibit(s) negative standardized growths, then the extreme event and its adjacent quarter(s) are considered as recession periods. If the second adjacent quarters before and/or after also have standardized growth \(< -0.9,\) then these quarters + adjacent one(s) + extreme event are identified as recession periods. Evaluate backward and forward until the “ripple” effect is considered diminished, or \(> = -0.9.\)
5. **Trough Points with Positive Growths**: given the recessions identified in (3) and (4), if there is a recession period lasting for at least three quarters and the following quarter has a positive growth rate (usually large), then this quarter is also considered as a recession period.

I test the algorithm using the actual monthly consumption data, and compare the identified recessions with the NBER recession periods. This methodology is able to identify 7 out of the 8 NBER recessions between January 1959 and June 2014, with a significant correlation 80%. The quarterly regression coefficient is 0.9038 (SE=0.0507), which fails to reject the null hypothesis of 1 at the 5% significance level.

C  Intermediate models in Section 4

In this section, I introduce the two intermediate models and the final model as analyzed in Section 4. The first intermediate model “M(1)” is an adapted Campbell and Cochrane model that features one state variable (the surplus consumption ratio), homoskedastic shocks, constant expected dividend growth, and a constant consumption-dividend comovement. The second intermediate model “M(2)” is an adapted Bekaert and Engstrom model that builds on M(1) to incorporate a countercyclical macroeconomic uncertainty as a salient state variable. Lastly, the final model in this paper “M(3)” can be viewed as a generalization of M(2) where the consumption-dividend comovement is now procyclical. All three models price dividend claims.

- **Fundamentals:**

  \[ M(1): \quad \Delta c_{t+1} = \epsilon + \sigma_c \tilde{\omega}_{c,t+1} + \sigma_n \tilde{\omega}_{n,t+1}, \]
  \[ \tilde{\omega}_{c,t+1} \sim i.i.d. N(0,1), \tilde{\omega}_{n,t+1} \sim \Gamma(\pi,1) - \pi, \]
  \[ \Delta d_{t+1} = \bar{d} + \bar{\sigma}_d \tilde{\omega}_{c,t+1} + \sigma_d \tilde{\omega}_{d,t+1}, \]
  \[ \tilde{\omega}_{d,t+1} \sim \Gamma(V_d,1) - V_d, \]
  \[ \epsilon = 0.0025, \sigma_c = 0.0029, \sigma_n = -0.0023, \pi = 0.3742, \]
  \[ \bar{d} = 0.0025, \sigma_d = -0.0008, V_d = 89.9322, \bar{\delta} = 0.4447; \]

  \[ M(2): \quad \Delta c_{t+1} = \epsilon + \sigma_c \tilde{\omega}_{c,t+1} + \sigma_n \tilde{\omega}_{n,t+1}, \]
  \[ n_{t+1} = (1 - \phi_n)\pi + \phi_n n_t + \sigma_n \tilde{\omega}_{n,t+1}, \]
  \[ \tilde{\omega}_{c,t+1} \sim i.i.d. N(0,1), \tilde{\omega}_{n,t+1} \sim \Gamma(n_t,1) - n_t, \]
\[ \Delta d_{t+1} = \bar{d} + \phi_d(V_{c,t} - V_c) + \sigma_d \bar{w}_{d,t+1} + \sigma_d \tilde{w}_{d,t+1}, \]  
\[ V_{c,t} = Var_t(\Delta c_{t+1}) = \sigma_c^2 + \sigma_n^2 n_t, V_c = E(V_{c,t}), \]  
\[ \bar{w}_{d,t+1} \sim \Gamma(V_d, 1) - V_d, \]  
\[ \pi = 0.0025, \sigma_c = 0.0029, \sigma_n = -0.0023, \pi = 0.3742, \phi_n = 0.9500, \sigma_{nn} = 0.2772, \]  
\[ \bar{d} = 0.0025, \phi_d = -568.0871, \sigma_d = -0.0008, V_d = 89.9322, \bar{r} = 0.4447; \]

M(3): The new DGP in this paper (as shown in Table 5).

• The surplus consumption ratio, M(1)~M(3):

\[ s_{t+1} = (1 - \phi_s) \bar{\pi} + \phi_s s_t + \lambda_t(\Delta c_{t+1} - \bar{\pi}). \]  

• The sensitivity functions:

\[ \lambda_t \left\{ \begin{array}{ll} \frac{1}{\gamma} \sqrt{1 - 2(s_t - \bar{\pi})} - 1, & s_t \leq s_{\text{max},t}, \\ 0, & s_t > s_{\text{max},t}. \end{array} \right. \]  

\[ \bar{\pi}_t = \log(\bar{S}_t), \]  

\[ s_{\text{max},t} = \bar{\pi}_t + \frac{1}{2}(1 - \bar{\pi}_t)^2, \]  

\[ M(1): \bar{S}_t = \sqrt{(\sigma_c^2 + \sigma_n^2 \bar{\pi}) \frac{\gamma}{1 - \phi_s}}. \]  

\[ M(2)/3: \bar{S}_t = \sqrt{(\sigma_c^2 + \sigma_n^2 n_t) \frac{\gamma}{1 - \phi_s}}. \]

• The real risk free rates (approximated at the third order for demonstration purpose):

\[ M(1): r_{f1} = -\ln(\beta + \gamma \bar{\pi} + \gamma(1 - \phi_s)(\bar{\pi}_t - s_t) - \frac{1}{2} \gamma^2 (1 + \lambda_t)^2 \sigma_c^2 - |\gamma(1 + \lambda_t)\sigma_n - \ln(1 + \gamma(1 + \lambda_t)\sigma_n)| \bar{\pi}} \]

\[ \approx -\ln(\beta + \gamma \bar{\pi} + \gamma(1 - \phi_s)(\bar{\pi}_t - s_t) - \frac{1}{2} \gamma^2 (1 + \lambda_t)^2 \sigma_c^2 - \frac{1}{2} \gamma^2 (1 + \lambda_t)^2 \sigma_n^2 \bar{\pi} + \frac{1}{3} \gamma^3 (1 + \lambda_t)^3 \sigma_n^2 \bar{\pi}} \]

\[ \approx rf_{cc} + \frac{1}{3} \gamma^3 (1 + \lambda_t)^3 \sigma_n^2 \bar{\pi} \]  

\[ M(2)/3: r_{f2} = -\ln(\beta + \gamma \bar{\pi} + \gamma(1 - \phi_s)(\bar{\pi}_t - s_t) - \frac{1}{2} \gamma^2 (1 + \lambda_t)^2 \sigma_c^2 - |\gamma(1 + \lambda_t)\sigma_n - \ln(1 + \gamma(1 + \lambda_t)\sigma_n)| n_t} \]

\[ \approx -\ln(\beta + \gamma \bar{\pi} + \gamma(1 - \phi_s)(\bar{\pi}_t - s_t) - \frac{1}{2} \gamma^2 (1 + \lambda_t)^2 \sigma_c^2 - \frac{1}{2} \gamma^2 (1 + \lambda_t)^2 \sigma_n^2 n_t + \frac{1}{3} \gamma^3 (1 + \lambda_t)^3 \sigma_n^2 n_t} \]

\[ \approx rf_{cc} + \frac{1}{3} \gamma^3 (1 + \lambda_t)^3 \sigma_n^2 n_t, \]

\[ rf_{cc} = -\ln(\beta + \gamma \bar{\pi} - \frac{(1 - \phi_s) \gamma}{2}). \]
D Quadratic approximation to the moment generating function of a random variable that is a linear combination of Gaussian, $\chi^2$, and gamma shocks

Suppose a random variable $x$ receives three shocks,

$$x = \mu + x_1 \omega + x_2 (\omega^2 - 1) + x_3 (\varepsilon - \alpha),$$
$$\omega \sim \text{i.i.d.} \mathcal{N}(0,1),$$
$$\omega^2 \sim \text{i.i.d.} \chi^2(1),$$
$$\varepsilon \sim \Gamma(\alpha,1),$$

where $\mu$ is the unconditional mean of variable $x$, and $\{x_1, x_2, x_3\}$ are constant coefficients. Recall the moment generating function (mgf) for a standard Gaussian shock is $mgf_{\omega}(\nu) = \exp(\nu^2/2)$, the mgf for a $\chi^2$ shock is $mgf_{\chi^2}(\nu) = (1 - 2\nu)^{-1/2}$, and the mgf for a gamma shock with a unit scale parameter is $mgf_{\varepsilon}(\nu) = (1 - \nu)^{-\alpha}$ where $\alpha$ is the shape parameter. The three shocks are uncorrelated. Therefore,

$$mgf_x(\nu) = E[\exp(\nu x)]$$
$$= \exp(\nu \mu) E[\exp(\nu x_1 \omega + \nu x_2 (\omega^2 - 1) + \nu x_3 (\varepsilon - \alpha))]$$
$$= \exp(\nu \mu - \nu x_2 - \nu x_3 \alpha) mgf_{\omega}(\nu x_1) mgf_{\chi^2}(\nu x_2) + mgf_{\varepsilon}(\nu x_3)$$
$$= \exp(\nu \mu - \nu x_2 - \nu x_3 \alpha) \exp \left\{ \frac{1}{2} (\nu x_1)^2 \right\} (1 - 2\nu x_2)^{-1/2} (1 - \nu x_3)^{-\alpha}$$
$$= \exp(\nu \mu - \nu x_2 - \nu x_3 \alpha) \exp \left\{ \frac{1}{2} (\nu x_1)^2 - \frac{1}{2} \ln (1 - 2\nu x_2) - \alpha \ln (1 - \nu x_3) \right\}.$$  (D.2)

It can be easily shown that the quadratic approximation to $\ln(1-x)$ is $-x - \frac{1}{2} x^2$. Applying the quadratic approximation to $\ln(1 - 2\nu x_2)$ and $\ln(1 - \nu x_3)$, the mgf becomes,

$$mgf_x(\nu) \approx \exp(\nu \mu - \nu x_2 - \nu x_3 \alpha) \exp \left\{ \frac{1}{2} (\nu x_1)^2 + \nu x_2 + (\nu x_2)^2 + \nu x_3 \alpha + \frac{1}{2} (\nu x_3)^2 \alpha \right\}$$
$$= \exp(\nu \mu) \exp \left\{ \frac{1}{2} (\nu x_1)^2 + (\nu x_2)^2 + \frac{1}{2} (\nu x_3)^2 \alpha \right\}$$
$$= \exp(\nu E(x)) \exp \left\{ \frac{1}{2} \nu^2 V(x) \right\}.$$  (D.3)

Define $X = \exp(x)$ and set $\nu = 1$,

$$E(X) \approx \exp \left\{ E(x) + \frac{1}{2} V(x) \right\}.$$  (D.4)

E Solving the theoretical model from Section 4 (approximate analytical solution)

In this section, I solve the theoretical model in Section 4 or M(3), with an approximate analytical solution using a similar procedure in the spirit of Bansal and Yaron (2004). The log linearization of the Euler equation becomes complex with gamma shocks and $\chi^2(1)$ shocks. Thus, to derive an “approximate” analytical solution, I use a quadratic approximation to the moment generating functions of the random variable $m_{t+1} + r_{t+1}$ (Appendix D), and impose a linear approximation to obtain the log market return. In this economy, the log valuation ratio $pd_t$ has the conjecture as follows:

$$pd_t = A_0 + A_1 s_t + A_2 b_t + A_3 b_t^2 + A_4 m_t.$$  (E.1)
The Campbell and Shiller linearization to market return shows,

\[ r_{t+1}^m = \Delta d_{t+1} + a_1 p d_{t+1} - p d_t + a_0. \]  

(E.2)

Given equations above, the log market return can be approximately expressed as a linear function of the state variables and four independent shocks to the economy:

\[ r_{t+1}^m = d - \phi_d \sigma_d^2 \pi + a_1 (A_0 + A_1 (1 - \phi_d) \pi + A_2 (1 - \phi_d) b + A_3 (1 - \phi_d) b)^2 + A_4 (1 - \phi_d) b) - A_0 + a_0 + A_1 (a_1 \phi_s - 1)s_1 + (a_1 A_2 \phi_b + 2a_1 A_3 (1 - \phi) b) b_1 + A_3 (a_1 \phi_s^2 - 1) b_2^2 + (a_1 A_2 \phi_b + a_4 \sigma_b^2) n_1 + (a_1 A_2 \lambda + a_1 A_2 \lambda + 2a_1 A_2 (1 - \phi) b_1) (1 + a_1 A_2 \lambda b_1) \sigma_c \tilde{\omega}_{c,t+1} + a_3 \sigma_c^2 \lambda b_1^2 (\tilde{\omega}_{c,t+1})^2 + a_1 (A_1 \lambda \sigma_n + A_4 \sigma_{nn}) \tilde{\omega}_{n,t+1} + \sigma_d \tilde{\omega}_{d,t+1}. \]  

(E.3)

With the approximate logarithm of the Euler equation and by equating the terms for the state variables, the coefficients in the price-dividend ratio equation are solved:

\[ A_1 = \frac{\gamma (1 - \phi_d)}{1 - a_1 \phi_s}, \]
\[ 0 < A_1 < \gamma. \]
\[ A_2 = \frac{(1 + 2a_1 A_3 \phi_b \lambda_b) \left[ \gamma (1 + \lambda) - (a_1 A_1 \lambda + 2a_1 A_3 (1 - \phi) b) \lambda_b \right] \sigma_b^2 - 2a_1 A_3 (1 - \phi) \phi_b b}{a_1 \phi_b - 1 + a_1 \lambda_b (1 + 2a_1 A_3 \phi_b \lambda_b) \sigma_b^2} > 0. \]
\[ A_3 = \frac{-2a_1 \phi_b \lambda_b \sigma_b^2 + 1 - a_1 \phi_b^2 + \sqrt{(2a_1 \phi_b \lambda_b \sigma_b^2 - 1 + a_1 \phi_b^2)^2 - 4a_1^2 \phi_b^2 \lambda_b^2 \sigma_b^2}}{4a_1^2 \phi_b^2 \lambda_b^2 \sigma_b^2} > 0. \]
\[ A_4 = \frac{\xi_t \pm \sqrt{\xi_t^2 - 2a_1^2 \sigma_n^2 \left( \phi_a \sigma_a^2 + \frac{1}{2} (A_1 \lambda a_1 - \gamma (1 + \lambda)) \sigma_a^2 \right)}}{2a_1 \sigma_n^2 \sigma_a^4}, \]
\[ \xi_t = 1 - \phi_a a_1 - a_1^2 A_1 \lambda \sigma_n \sigma_{nn} + \gamma (1 + \lambda) a_1 \sigma_n \sigma_{nn}, \]
\[ A_4 < 0 \iff \phi_a \sigma_a^2 < -\frac{1}{2} (A_1 \lambda a_1 - \gamma (1 + \lambda)) \sigma_n^2. \]

The equity premium equation is well-approximated with \(-Cov_t (r_{t+1}^m, m_{t+1})\) given the quadratic approximation,

\[ E_t (r_{t+1}^m) - rf_t + \frac{1}{2} Var_t (r_{t+1}^m) \approx -Cov_t (r_{t+1}^m, m_{t+1}) = \frac{\gamma (1 + \lambda)}{\text{price of consumption risk}} \times \frac{(a_1 A_1 \lambda \sigma_a^2)}{\text{1. approximate amount of consumption risk in CC}} + \frac{(a_1 A_2 \lambda + 2a_1 A_3 (1 - \phi) b) \lambda b_1) \sigma_b^2}{\text{2. additional amount of consumption risk induced by comovement}} + \frac{a_1^2 [A_1 \lambda \sigma_a^2 + A_4 \sigma_{nn} \sigma_n] n_1}{\text{3. additional amount of consumption risk induced by uncertainty}}. \]  

(E.9)
References


Table 1: Models of the Univariate Conditional Variances.

This table presents the estimation results of eight univariate conditional variance models with a constant long-run mean or with a time-varying long-run mean that is a linear function of the standardized NBER recession indicator for four variables: consumption growth, log market return, dividend growth, and the non-dividend part of the market return. 

Data “Consumption Growth” is the AR(3)-de-meaned change in log real consumption (non-durable goods and services) per capita (source: BEA). “Market Return” is the change in log real market index (including dividends) (source: CRSP). “Dividend Growth” is the change in log real dividend level per capita (source: Shiller’s website). “Market Return–Dividend Growth” is the linear difference between log market return and log dividend growth. The NBER recession indicator ($I_{NBER,t}$) is obtained from the NBER website (1=recession; 0=otherwise), the standardized NBER indicator ($SNBER_t$) is predetermined as $(I_{NBER,t} - E(I_{NBER,t}))/SD(I_{NBER,t})$. To obtain the residuals, the four variables are regressed on the NBER recession indicator to account for the the cyclical conditional mean (if any).


Models with $q_t$ see Section 2 for modeling details. Model estimation uses the maximum likelihood estimation (MLE) methodology given the specified shock distributions. The estimations for “GARCH” and “GED-GARCH” use variance targeting. The robust standard errors are shown in parentheses. Values in bold (italics) are statistically significance at a significant level of 5% (10%). Underlined models are the best models among the eight models, given the Bayesian Information Criteria (BIC). N=665 months (1959/02∼2014/06).

---

### Panel A. Consumption Growth

<table>
<thead>
<tr>
<th></th>
<th>Loglikelihood</th>
<th>Nparams</th>
<th>AIC</th>
<th>BIC</th>
<th>$SNBER_t$ coefficient (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconditional</td>
<td>2884.42</td>
<td>1</td>
<td>-5766.84</td>
<td>-5762.34</td>
<td></td>
</tr>
<tr>
<td>GARCH</td>
<td>2911.44</td>
<td>2</td>
<td>-5816.87</td>
<td>-5803.37</td>
<td></td>
</tr>
<tr>
<td>GED-GARCH</td>
<td>2913.30</td>
<td>3</td>
<td>-5820.59</td>
<td>-5807.10</td>
<td></td>
</tr>
<tr>
<td>BEGE-$n_t$-GARCH</td>
<td>2919.62</td>
<td>6</td>
<td>-5827.23</td>
<td>-5800.24</td>
<td></td>
</tr>
<tr>
<td>Unconditional, $q_t$</td>
<td>2885.92</td>
<td>1</td>
<td>-5767.84</td>
<td>-5758.84</td>
<td>0.1014 (0.0066)</td>
</tr>
<tr>
<td>GARCH, $q_t$</td>
<td>2913.14</td>
<td>3</td>
<td>-5820.28</td>
<td>-5806.78</td>
<td>0.0270 (0.0060)</td>
</tr>
<tr>
<td>GED-GARCH, $q_t$</td>
<td>2917.30</td>
<td>4</td>
<td>-5826.59</td>
<td>-5808.60</td>
<td>0.0428 (0.0099)</td>
</tr>
<tr>
<td>BEGE-$n_t$-GARCH, $q_t$</td>
<td>2924.92</td>
<td>7</td>
<td>-5835.85</td>
<td>-5804.35</td>
<td>0.1941 (0.0173)</td>
</tr>
</tbody>
</table>

---

### Panel B. Market Return

<table>
<thead>
<tr>
<th></th>
<th>Loglikelihood</th>
<th>Nparams</th>
<th>AIC</th>
<th>BIC</th>
<th>$SNBER_t$ coefficient (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconditional</td>
<td>1241.50</td>
<td>1</td>
<td>-2481.01</td>
<td>-2476.51</td>
<td></td>
</tr>
<tr>
<td>GARCH</td>
<td>1275.84</td>
<td>2</td>
<td>-2547.68</td>
<td>-2538.68</td>
<td></td>
</tr>
<tr>
<td>GED-GARCH</td>
<td>1293.13</td>
<td>3</td>
<td>-2580.25</td>
<td>-2566.75</td>
<td></td>
</tr>
<tr>
<td>BEGE-$n_t$-GARCH</td>
<td>1307.83</td>
<td>6</td>
<td>-2603.65</td>
<td>-2576.65</td>
<td></td>
</tr>
<tr>
<td>Unconditional, $q_t$</td>
<td>1284.04</td>
<td>1</td>
<td>-2566.09</td>
<td>-2561.59</td>
<td>0.6541 (0.0888)</td>
</tr>
<tr>
<td>GARCH, $q_t$</td>
<td>1298.05</td>
<td>3</td>
<td>-2590.10</td>
<td>-2576.60</td>
<td>0.6944 (0.0938)</td>
</tr>
<tr>
<td>GED-GARCH, $q_t$</td>
<td>1307.41</td>
<td>4</td>
<td>-2606.82</td>
<td>-2588.82</td>
<td>0.6935 (0.1084)</td>
</tr>
<tr>
<td>BEGE-$n_t$-GARCH, $q_t$</td>
<td>1309.54</td>
<td>7</td>
<td>-2605.09</td>
<td>-2573.59</td>
<td>0.6077 (0.4930)</td>
</tr>
</tbody>
</table>

---

### Panel C. Dividend Growth

<table>
<thead>
<tr>
<th></th>
<th>Loglikelihood</th>
<th>Nparams</th>
<th>AIC</th>
<th>BIC</th>
<th>$SNBER_t$ coefficient (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconditional</td>
<td>2374.99</td>
<td>1</td>
<td>-4747.98</td>
<td>-4743.48</td>
<td></td>
</tr>
<tr>
<td>GARCH</td>
<td>2441.71</td>
<td>2</td>
<td>-4879.43</td>
<td>-4870.43</td>
<td></td>
</tr>
<tr>
<td>GED-GARCH</td>
<td>2485.82</td>
<td>3</td>
<td>-4965.63</td>
<td>-4952.13</td>
<td></td>
</tr>
<tr>
<td>BEGE-$n_t$-GARCH</td>
<td>2495.86</td>
<td>6</td>
<td>-4979.73</td>
<td>-4952.73</td>
<td></td>
</tr>
<tr>
<td>Unconditional, $q_t$</td>
<td>2376.52</td>
<td>1</td>
<td>-4751.05</td>
<td>-4746.55</td>
<td>-0.0897 (0.0441)</td>
</tr>
<tr>
<td>GARCH, $q_t$</td>
<td>2444.05</td>
<td>3</td>
<td>-4882.11</td>
<td>-4868.61</td>
<td>-0.1316 (0.0413)</td>
</tr>
<tr>
<td>GED-GARCH, $q_t$</td>
<td>2486.82</td>
<td>4</td>
<td>-4965.63</td>
<td>-4947.63</td>
<td>-0.0714 (0.0620)</td>
</tr>
<tr>
<td>BEGE-$n_t$-GARCH, $q_t$</td>
<td>2499.63</td>
<td>7</td>
<td>-4985.26</td>
<td>-4953.76</td>
<td>-0.1114 (0.0592)</td>
</tr>
</tbody>
</table>

continue next page
Panel D. Market Return–Dividend Growth (Non-Dividend Part of the Market Return)

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Loglikelihood</th>
<th>Nparams</th>
<th>AIC</th>
<th>BIC</th>
<th>(SNBER_t) coefficient (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconditional</td>
<td>1241.08</td>
<td>1</td>
<td>-2480.17</td>
<td>-2475.67</td>
<td></td>
</tr>
<tr>
<td>GARCH</td>
<td>1276.97</td>
<td>2</td>
<td>-2549.95</td>
<td>-2540.95</td>
<td></td>
</tr>
<tr>
<td>GED-GARCH</td>
<td>1294.55</td>
<td>3</td>
<td>-2583.10</td>
<td>-2569.60</td>
<td></td>
</tr>
<tr>
<td>BEGE-(n_t)-GARCH</td>
<td>1303.12</td>
<td>6</td>
<td>-2594.24</td>
<td>-2567.24</td>
<td></td>
</tr>
<tr>
<td>Unconditional, (q_t)</td>
<td>1284.03</td>
<td>1</td>
<td>-2566.05</td>
<td>-2561.55</td>
<td>0.6577 (0.0888)</td>
</tr>
<tr>
<td>GARCH, (q_t)</td>
<td>1300.38</td>
<td>3</td>
<td>-2594.76</td>
<td>-2581.26</td>
<td>0.6822 (0.0873)</td>
</tr>
<tr>
<td>GED-GARCH, (q_t)</td>
<td>1308.63</td>
<td>4</td>
<td>-2609.27</td>
<td>-2591.27</td>
<td>0.6689 (0.0991)</td>
</tr>
<tr>
<td>BEGE-(m_t)-GARCH, (q_t)</td>
<td>1313.71</td>
<td>7</td>
<td>-2613.43</td>
<td>-2581.93</td>
<td>0.7571 (0.1036)</td>
</tr>
</tbody>
</table>
Table 2: Parameter Estimates for the Best Univariate Conditional Variance Models.

This table presents the best univariate conditional variance models (denoted with underlines in Table 1) of consumption growth, log market return, dividend growth, and the non-dividend part of the market return. Denote $\tilde{\epsilon}_{t+1}$ as the zero-mean innovation at time $t+1$, and $h_t$ as the conditional variance of $\tilde{\epsilon}_{t+1}$. The best conditional variance model for consumption growth (Panel A of Table 1), log market return (Panel B of Table 1) and the non-dividend part of the market return (Panel D of Table 1) is “GED-GARCH-$q_t$”:

$$\tilde{\epsilon}_{t+1} \sim GED(0, h_t, \tau);$$

$$V_t(\tilde{\epsilon}_{t+1}) \equiv h_t = \bar{h}(1 + q_t + \alpha [\tilde{\epsilon}_{t-1}^2 - \bar{h}(1 + q_{t-1})] + \beta [h_{t-1} - \bar{h}(1 + q_{t-1})],$$

where $\alpha + \beta < 1$, $\alpha > 0$, $\beta > 0$; $\tau > 0$ is the shape parameter of the Generalized Error Distribution. The best conditional variance model for dividend growth (Panel C of Table 1) is “BEGE-$n_t$-GARCH-$q_t$”:

$$\tilde{\epsilon}_{t+1} = \sigma_p \tilde{\omega}_{p,t+1} - \sigma_n \tilde{\omega}_{n,t+1} \sim \Gamma(p, 1) - \bar{p}, \tilde{\omega}_{n,t+1} \sim \Gamma(n_t, 1) - n_t;$$

$$V_t(\tilde{\epsilon}_{t+1}) \equiv h_t = \sigma_p^2 p + \sigma_n^2 n_t; n_t = \pi(1 + q_t + \alpha_n [\tilde{\epsilon}_{t-1}^2 \sigma_p^2 - \pi(1 + q_{t-1})] + \beta_n [n_{t-1} - \pi(1 + q_{t-1})]),$$

where $\alpha_n + \beta_n < 1$, $\alpha_n > 0$, $\beta_n > 0$; $p$ ($\sigma_p$) is the shape (scale) parameter of the zero-mean homoskedastic good-environment gamma shock $\{\tilde{\omega}_{p,t+1}\}$, and $\{n_t\}$ ($\sigma_n$) the shape parameter of the zero-mean heteroskedastic bad-environment gamma shock $\{\tilde{\omega}_{n,t+1}\}$; $\sigma_p$, $\sigma_n$, $p$ and $n_t$ are strictly positive to satisfy statistical properties of gamma distributions. The time-varying long-run conditional mean of the conditional variances is a multiple of the standardized NBER recession indicator:

$$q_t = \nu S NBER_t.$$

More model descriptions and estimation methodologies are discussed in Section 2. Robust standard errors are shown in parentheses. Values in bold (italics) are statistically significant at a significant level of 5% (10%). N=665 months (1959/02~2014/06).

<table>
<thead>
<tr>
<th>Panel A. Consumption Growth, GED-GARCH-$q_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sqrt{\bar{h}}$</td>
</tr>
<tr>
<td>Est.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B. Market Return, GED-GARCH-$q_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sqrt{\bar{h}}$</td>
</tr>
<tr>
<td>Est.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C. Dividend Growth, BEGE-$n_t$-GARCH-$q_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sqrt{\bar{p}}$</td>
</tr>
<tr>
<td>Est.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel D. Market Return–Dividend Growth, GED-GARCH-$q_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sqrt{\bar{h}}$</td>
</tr>
<tr>
<td>Est.</td>
</tr>
</tbody>
</table>
Given the best univariate conditional variance estimates from Table 2 this table formally identifies the cyclicality of the conditional comovements (correlation in Panel A, covariance in Panel B, and beta in Panel C) between the innovations of consumption growth and the innovations of different components of the market return using the DCC-$q_t$ framework. In particular, the amount of risk ($\text{Cov}(\Delta c_{t+1}, r_{m,t+1})$) is decomposed into two components: the “exogenous” component ($\text{Cov}(\Delta c_{t+1}, \Delta d_{t+1})$) and the “endogenous” component ($\text{Cov}(\Delta c_{t+1}, r_{m,t+1} - \Delta d_{t+1})$). The DCC-$q_t$ Model The bivariate dependence framework is designed to jointly estimate and test the cyclicality of the conditional correlation. The conditional correlation matrix $\text{Corr}_t$ is modeled with a quadratic form, $(Q_t^{-1} - Q_t (Q_t^{-1})^{-1})$, where $Q_t$ is the diagonal matrix with the square root of the diagonal element of $Q_t$ on the diagonal (so the diagonal entries of $\text{Corr}_t$ are strictly equal to 1). The off-diagonal element of $\text{Corr}_t$ is the conditional correlation (or equivalently, the conditional covariance) of the standardized disturbance obtained from Table 2 $z_{t+1} \equiv [z_{1,t+1}, z_{2,t+1}]'$. The DCC-$q_t$ model is expressed as follows,

$$Q_t = \frac{\theta}{2} \sum_{i=1}^{T} z_i z_i' + \alpha_{12} [z_t z_t' - \bar{Q}_{12} \frac{1}{1 + q_t} 1 + q_t^{-1} 1 + q_t^{-1}] + \beta_{12} [Q_{t-1} - \bar{Q}_{12} \frac{1}{1 + q_{t-1}} 1 + q_{t-1}^{-1}],$$

where parameter $\bar{Q}_{12}$ is the off-diagonal term of the predetermined constant conditional correlation matrix $\bar{Q} = \frac{\theta}{2} \sum_{i=1}^{T} z_i z_i'$, and $q_t$ is a multiple of the standardized NBER recession indicator, $q_t = \nu S_{\text{NBER}} t$.

The DCC Model With $\nu = 0$ (or a constant long-run mean), the DCC model (Engle, 2002) is a natural null hypothesis of the DCC-$q_t$ model. Implied Conditional Covariance and Conditional Beta Given the conditional correlation estimates from Panel A of this table and the conditional variance estimates from Table 2 the conditional covariance and the conditional beta of dividend growth to consumption growth innovations are derived and regressed on the NBER recession indicator ($I_{\text{NBER}} = 1$ during recessions; 0 otherwise) in Panel B and Panel C. Other Notations “LL”, loglikelihood; “LR”, likelihood ratio; $b(I_{\text{NBER}})$. OLS coefficient on the NBER recession indicator. Robust standard errors are shown in parentheses. Values in bold (italics) are statistically significant at a significant level of 5% (10%). N=665 months (1959/02~2014/06).

<table>
<thead>
<tr>
<th>Series 1: Market Return ($r_m$)</th>
<th>Consumption Growth ($\Delta c$)</th>
<th>Dividend Growth ($\Delta d$)</th>
<th>$r_m - \Delta d$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Conditional Correlation</strong></td>
<td><strong>DCC</strong></td>
<td><strong>DCC-$q_t$</strong></td>
<td><strong>DCC</strong></td>
</tr>
<tr>
<td>$\bar{Q}_{12}$</td>
<td>0.2022</td>
<td>0.2022</td>
<td>0.2298</td>
</tr>
<tr>
<td>($\text{fix}$)</td>
<td>($\text{fix}$)</td>
<td>($\text{fix}$)</td>
<td>($\text{fix}$)</td>
</tr>
<tr>
<td>$\alpha_{12}$</td>
<td>0.0223</td>
<td>0.0227</td>
<td>0.1431</td>
</tr>
<tr>
<td>(0.1463)</td>
<td>(0.1636)</td>
<td>(0.0637)</td>
<td>(0.0641)</td>
</tr>
<tr>
<td>$\beta_{12}$</td>
<td>0.9657</td>
<td>0.9340</td>
<td>0.8300</td>
</tr>
<tr>
<td>(0.0740)</td>
<td>(0.0830)</td>
<td>(0.0255)</td>
<td>(0.0262)</td>
</tr>
<tr>
<td>$\nu$</td>
<td>$-1.539$</td>
<td>$-0.7999$</td>
<td>$0.4445$</td>
</tr>
<tr>
<td>$\text{LL}(\text{param})$</td>
<td>112.30</td>
<td>116.62</td>
<td>57.62</td>
</tr>
<tr>
<td>Wald test stats. ($H_0=\text{DCC}$)</td>
<td>-</td>
<td>5.69</td>
<td>-</td>
</tr>
<tr>
<td>P-value</td>
<td>-</td>
<td>1.702%</td>
<td>-</td>
</tr>
<tr>
<td>LR test stats. ($H_0=\text{DCC}$)</td>
<td>-</td>
<td>8.64</td>
<td>-</td>
</tr>
<tr>
<td>P-value</td>
<td>-</td>
<td>0.329%</td>
<td>-</td>
</tr>
</tbody>
</table>

| **Panel B. Conditional Covariance** | **DCC** | **DCC-$q_t$** | **DCC** | **DCC-$q_t$** | **DCC** | **DCC-$q_t$** |
|----------------------------------|----------------------------------|-----------------------------|------------------|
| $b(I_{\text{NBER}})$ ($\times 10^{-5}$) | $-0.2849$ | $-0.5030$ | $-0.5358$ | $-0.8919$ | $0.2389$ | $0.3776$ |
| $\text{SE}(b(I_{\text{NBER}}))$ ($\times 10^{-5}$) | (0.1505) | (0.2607) | (0.2562) | (0.2564) | (0.0987) | (0.0993) |

| **Panel C. Conditional Beta** | **DCC** | **DCC-$q_t$** | **DCC** | **DCC-$q_t$** | **DCC** | **DCC-$q_t$** |
|----------------------------------|----------------------------------|-----------------------------|------------------|
| $b(I_{\text{NBER}})$ | $-0.1418$ | $-0.3501$ | $-0.3268$ | $-0.7131$ | $0.1850$ | $0.3630$ |
| $\text{SE}(b(I_{\text{NBER}}))$ | (0.0719) | (0.0713) | (0.0482) | (0.0493) | (0.0719) | (0.0719) |
In addition to Campbell and Cochrane (1999)’s habit-formation model and Bansal and Yaron (2004)’s long-run risk model, five variants of the two models that focus on modeling more realistic shocks are evaluated at their abilities to meet the stylized facts established in Section 2. To provide intuitions about the non-dividend part of the market return (\(r_t\)), I follow the convention in the long-run risk literature to approximate \(r_t\) with changes in log dividend \((\Delta d_t)\). The Campbell and Shiller linearization of the price dividend ratio according to BE2004 gives \(\sigma_{(r_t)}^{(\Delta d_t)}\) procyclical when the scale parameter of bad uncertainty shock \(\sigma_{(c_t)}^{(\Delta d_t)}\) is greater than the scale parameter of bad uncertainty in dividend \((\sigma_{(d_t)}^{(\Delta d_t)})\), with changes in log dividend \(\Delta d_t\) being countercyclical in BE2004.

<table>
<thead>
<tr>
<th>Data</th>
<th>Habit</th>
<th>BE2009</th>
<th>BY2004</th>
<th>BTZ2006</th>
<th>LRR</th>
<th>BKY2012</th>
<th>SSY2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Var (\Delta c_t)</td>
<td>Pro-</td>
<td>Pro-</td>
<td>Const.</td>
<td>Const.</td>
<td>(5)</td>
<td>Const.</td>
<td></td>
</tr>
<tr>
<td>Cov (\Delta d_t, \Delta c_t)</td>
<td>Pro-</td>
<td>Pro-</td>
<td>Const.</td>
<td>Const.</td>
<td>(6)</td>
<td>Const.</td>
<td></td>
</tr>
<tr>
<td>Var (\Delta d_t)</td>
<td>Pro-</td>
<td>Pro-</td>
<td>Const.</td>
<td>Const.</td>
<td>(7)</td>
<td>Const.</td>
<td></td>
</tr>
<tr>
<td>Cov (\Delta d_t, \Delta c_t)</td>
<td>Pro-</td>
<td>Pro-</td>
<td>Const.</td>
<td>Const.</td>
<td>(8)</td>
<td>Const.</td>
<td></td>
</tr>
<tr>
<td>Var (\Delta c_t)</td>
<td>Pro-</td>
<td>Pro-</td>
<td>Const.</td>
<td>Const.</td>
<td>(9)</td>
<td>Const.</td>
<td></td>
</tr>
<tr>
<td>Cov (\Delta d_t, \Delta c_t)</td>
<td>Pro-</td>
<td>Pro-</td>
<td>Const.</td>
<td>Const.</td>
<td>(10)</td>
<td>Const.</td>
<td></td>
</tr>
</tbody>
</table>

Boxed moments are mismatched by all representative models.
Table 5: The New DGP for the Joint Consumption-Dividend Dynamics.

In the new DGP, consumption and dividend growth have the following joint dynamics:

\[
\Delta c_{t+1} = \tau + \sigma_c \tilde{\omega}_c,_{t+1} + \sigma_n \tilde{\omega}_n,_{t+1},
\]

\[
n_{t+1} = (1 - \phi_n) \pi + \phi_n n_t + \sigma_n \tilde{\omega}_n,_{t+1},
\]

\[
\Delta d_{t+1} = \bar{d} + \phi_d (V_{c,t} - \nabla c) + b_t \sigma_c \tilde{\omega}_c,_{t+1} + \sigma_d \tilde{\omega}_d,_{t+1},
\]

\[
b_{t+1} = (1 - \phi_b) \delta + \phi_b b_t + \lambda_b \sigma_c \tilde{\omega}_c,_{t+1},
\]

\[V_{c,t} = \sigma_c^2 + \sigma_n^2 n_t,
\]

\[\nabla c = \sigma_c^2 + \sigma_n^2 \sigma_n^2,
\]

where the consumption fundamental shock, \(\tilde{\omega}_c,_{t+1}\), is a centered Gaussian shock with unit standard deviation, the consumption event shock, \(\tilde{\omega}_n,_{t+1}\), follows a centered heteroskedastic gamma distribution with a strictly positive shape parameter equal to \(n_t\) and a unit scale parameter, and the dividend-specific shock, \(\tilde{\omega}_d,_{t+1}\), follows a centered homoskedastic gamma distribution with a strictly positive shape parameter equal to \(V_d\) and a unit scale parameter. Or,

\[\tilde{\omega}_c,_{t+1} \sim i.i.d. N(0, 1); \tilde{\omega}_n,_{t+1} \sim \Gamma(n_t, 1) - n_t; \tilde{\omega}_d,_{t+1} \sim \Gamma(V_d, 1) - V_d.\]

Data As in the empirical part of the paper, I use the AR(3)-de-meaned consumption growth and the original dividend growth as \(\Delta c_{t+1}\) and \(\Delta d_{t+1}\) in the DGP estimation. Panels Panels A and B present the estimation results, where “ADF Test” denotes the augmented Dickey-Fuller tests with the null that the two latent processes, \(n_{t+1}\) and \(b_{t+1}\), follow unit root processes. Panel C presents the monthly unconditional moments of the three filtered shocks. Panel D shows the correlations between the business cycle indicators (the NBER indicator and the detrended consumption-wealth ratio, \(\hat{cay}_Q\), introduced in Lettau and Ludvigson (2001); source: NBER website and Martin Lettau’s website) and the three filtered shocks aggregated to the quarterly frequency. Robust standard errors are shown in parentheses in Panels A, B, and D; bootstrapped standard errors are shown in parentheses in Panel C. Values in bold are statistically significant at a significant level of 5%. N=665 months (1959/02~2014/06).

<table>
<thead>
<tr>
<th></th>
<th>Panel A. Estimation Results, Consumption</th>
<th>Panel B. Estimation Results, Dividend</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\tau)</td>
<td>0.0025</td>
<td>0.0025</td>
</tr>
<tr>
<td>(0.0001)</td>
<td>(0.0003)</td>
<td></td>
</tr>
<tr>
<td>(\pi)</td>
<td>0.3742</td>
<td>0.4447</td>
</tr>
<tr>
<td>(1.6099)</td>
<td>(1.0181)</td>
<td></td>
</tr>
<tr>
<td>(\sigma_c)</td>
<td>0.0029</td>
<td>-568.0871</td>
</tr>
<tr>
<td>(0.0001)</td>
<td>(139.0511)</td>
<td></td>
</tr>
<tr>
<td>(\phi_n)</td>
<td>0.9500</td>
<td>0.8824</td>
</tr>
<tr>
<td>(0.0264)</td>
<td>(0.0468)</td>
<td></td>
</tr>
<tr>
<td>(\sigma_n)</td>
<td>-0.0023</td>
<td>0.0008</td>
</tr>
<tr>
<td>(0.0005)</td>
<td>(0.0000)</td>
<td></td>
</tr>
<tr>
<td>(\sigma_{nn})</td>
<td>0.2772</td>
<td>14.0978</td>
</tr>
<tr>
<td>(0.1027)</td>
<td>(1.3764)</td>
<td></td>
</tr>
<tr>
<td>(\phi_d)</td>
<td>-0.0025</td>
<td></td>
</tr>
<tr>
<td>(0.0374)</td>
<td>(0.0649)</td>
<td></td>
</tr>
<tr>
<td>(\phi_b)</td>
<td>0.3842</td>
<td></td>
</tr>
<tr>
<td>(0.0387)</td>
<td>(0.0667)</td>
<td></td>
</tr>
<tr>
<td>(\sigma_d)</td>
<td>89.9322</td>
<td></td>
</tr>
<tr>
<td>(3.4724)</td>
<td>(3.4724)</td>
<td></td>
</tr>
<tr>
<td>(V_d)</td>
<td>89.9322</td>
<td></td>
</tr>
<tr>
<td>(3.4724)</td>
<td>(3.4724)</td>
<td></td>
</tr>
<tr>
<td>ADF Test</td>
<td>-4.298***</td>
<td>-4.764***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\bar{d})</td>
<td>0.0025</td>
<td></td>
</tr>
<tr>
<td>(0.0003)</td>
<td>(0.0003)</td>
<td></td>
</tr>
<tr>
<td>(\bar{b})</td>
<td>0.4447</td>
<td></td>
</tr>
<tr>
<td>(1.0181)</td>
<td>(1.0181)</td>
<td></td>
</tr>
<tr>
<td>(\lambda_b)</td>
<td>14.0978</td>
<td></td>
</tr>
<tr>
<td>(1.3764)</td>
<td>(1.3764)</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.0017</td>
<td>-0.2703</td>
</tr>
<tr>
<td>(0.0374)</td>
<td>(0.0649)</td>
<td></td>
</tr>
<tr>
<td>S.D.</td>
<td>0.9680</td>
<td>0.2483</td>
</tr>
<tr>
<td>(0.0387)</td>
<td>(0.0667)</td>
<td></td>
</tr>
<tr>
<td>Skew</td>
<td>0.1875</td>
<td>5.4390</td>
</tr>
<tr>
<td>(0.1326)</td>
<td>(0.0653)</td>
<td></td>
</tr>
<tr>
<td>xKurt</td>
<td>0.4491</td>
<td>40.5216</td>
</tr>
<tr>
<td>(0.5023)</td>
<td>(0.0673)</td>
<td></td>
</tr>
<tr>
<td>NBER</td>
<td>-0.2703</td>
<td>0.2483</td>
</tr>
<tr>
<td>(0.0649)</td>
<td>(0.0653)</td>
<td></td>
</tr>
<tr>
<td>(\hat{cay}_Q)</td>
<td>-0.2154</td>
<td>0.0561</td>
</tr>
<tr>
<td>(0.0658)</td>
<td>(0.0673)</td>
<td></td>
</tr>
<tr>
<td>(\hat{cay}_Q)</td>
<td>-0.2154</td>
<td>0.0350</td>
</tr>
<tr>
<td>(0.0674)</td>
<td>(0.0674)</td>
<td></td>
</tr>
</tbody>
</table>

N=665 months (1959/02~2014/06).
Table 6: Theoretical Models: Resolving the Exogenous Part of the Duffee Puzzle.

This table evaluates the abilities of three theoretical models to fit Facts (a)–(e) related to the exogenous part of the Duffee Puzzle as established in Section 2. To obtain the sample moments under Column “Data”, I calculate the unconditional moments of residuals during recession ($I_{NBER} = 1$) and non-recession ($I_{NBER} = 0$) periods; monthly empirical data cover period 1959/01-2014/06. Bootstrapped standard errors are shown in parentheses under Column “SE”. Similarly, I obtain their theoretical counterparts using the simulated datasets of the theoretical models under Columns M(1), M(2) and M(3). The algorithm for identifying recession periods in an endowment economy is described in Appendix B.

### Data Symbols
- $\sigma$, volatility
- $\rho$, correlation
- $C$, covariance
- $b$, sensitivity of $x_1$ to $x_2$ or $C(x_1, x_2)$
- $\sigma^2$, % of Amount of Risk by $C(\Delta d, \Delta c)$
- $C(\Delta d, \Delta c)$
- $C(r_m, \Delta c) \times 100\%$

### Models
- M(1) is an adapted Campbell and Cochrane (1999) model with the surplus consumption ratio ($s$) as the only state variable.
- M(2) is adapted from Bekaert and Engstrom (2017) with heteroskedastic macroeconomic uncertainty ($n$).
- M(3) is the full model in this paper with procyclical consumption-dividend comovement ($b$) as the third state variable.

<table>
<thead>
<tr>
<th>Data</th>
<th>SE</th>
<th>M(1)</th>
<th>M(2)</th>
<th>M(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma(\Delta c)$ ($I_{rece.} = 0$)</td>
<td>0.0031*** (0.0001)</td>
<td>0.0032</td>
<td>0.0031</td>
<td>0.0031</td>
</tr>
<tr>
<td>$\sigma(\Delta c)$ ($I_{rece.} = 1$)</td>
<td>0.0036 (0.0002)</td>
<td>0.0032</td>
<td>0.0036</td>
<td>0.0036</td>
</tr>
<tr>
<td>$\sigma(\Delta d)$ ($I_{rece.} = 0$)</td>
<td>0.0069* (0.0005)</td>
<td>0.00741</td>
<td>0.00741</td>
<td>0.00742</td>
</tr>
<tr>
<td>$\sigma(\Delta d)$ ($I_{rece.} = 1$)</td>
<td>0.0060 (0.0006)</td>
<td>0.00741</td>
<td>0.00741</td>
<td>0.00492</td>
</tr>
<tr>
<td>$\rho(\Delta d, \Delta c)$ ($I_{rece.} = 0$)</td>
<td>0.1823*** (0.0565)</td>
<td>0.1577</td>
<td>0.1622</td>
<td>0.1627</td>
</tr>
<tr>
<td>$\rho(\Delta d, \Delta c)$ ($I_{rece.} = 1$)</td>
<td>0.0148 (0.0591)</td>
<td>0.1571</td>
<td>0.1399</td>
<td>0.0639</td>
</tr>
<tr>
<td>$C(\Delta d, \Delta c)$ ($I_{rece.} = 0$) ($\times 10^{-5}$)</td>
<td>0.3916** (0.1523)</td>
<td>0.3703</td>
<td>0.3703</td>
<td>0.3720</td>
</tr>
<tr>
<td>$C(\Delta d, \Delta c)$ ($I_{rece.} = 1$) ($\times 10^{-5}$)</td>
<td>0.0317 (0.1609)</td>
<td>0.3703</td>
<td>0.3703</td>
<td>0.1122</td>
</tr>
<tr>
<td>$b(\Delta d, \Delta c)$ ($I_{rece.} = 0$)</td>
<td>0.4075*** (0.1419)</td>
<td>0.3686</td>
<td>0.3902</td>
<td>0.3920</td>
</tr>
<tr>
<td>$b(\Delta d, \Delta c)$ ($I_{rece.} = 1$)</td>
<td>0.0251 (0.1277)</td>
<td>0.3659</td>
<td>0.2900</td>
<td>0.0879</td>
</tr>
</tbody>
</table>

Note: the significance levels for testing the equality of non-recession ($I_{rece.} = 0$) and recession ($I_{rece.} = 1$) moments (Wald Test given the covariance-variance matrix from the GMM estimations) are denoted as follow:
- *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.  

51
Table 7: Theoretical Models: Unconditional Moments of the Duffee Puzzle Components.

This table presents 18 unconditional moments from simulated and empirical datasets. **Moment Symbols** $E$, mean; $\sigma$, volatility; $Skew$, scaled skewness; $xKurt$, excess kurtosis; $\rho$, correlation; $C$, covariance; $b(x_1, x_2)$, sensitivity of $x_1$ to $x_2$ or $\frac{C(x_1, x_2)}{\sigma(x_1)^2}$; % of Amount of Risk by $C(\Delta d, \Delta c)$, $\frac{C(\Delta d, \Delta c)}{\sigma(r_m)^2}$.

### Data

<table>
<thead>
<tr>
<th>Data</th>
<th>SE</th>
<th>M(1)</th>
<th>M(2)</th>
<th>M(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_m$</td>
<td>-</td>
<td>0.0025</td>
<td>0.0025</td>
<td>0.0025</td>
</tr>
<tr>
<td>$\sigma(r_m)^2$</td>
<td>-</td>
<td>0.0025</td>
<td>0.0025</td>
<td>0.0025</td>
</tr>
<tr>
<td>$\rho(\Delta d, \Delta c)$</td>
<td>0.1587</td>
<td>0.1558</td>
<td>0.1557</td>
<td>0.1503</td>
</tr>
<tr>
<td>$C(\Delta d, \Delta c)(\times 10^{-5})$</td>
<td>3.419 (0.1286)</td>
<td>0.3703</td>
<td>0.3703</td>
<td>0.3547</td>
</tr>
<tr>
<td>$b(\Delta d, \Delta c)$</td>
<td>0.3414 (0.1248)</td>
<td>0.3600</td>
<td>0.3598</td>
<td>0.3441</td>
</tr>
<tr>
<td>$\sigma(r_m^m - \Delta d)$</td>
<td>0.0575 (0.0017)</td>
<td>0.0259</td>
<td>0.0388</td>
<td>0.0362</td>
</tr>
<tr>
<td>$\sigma(r_m^m)$</td>
<td>0.0374 (0.0016)</td>
<td>0.0252</td>
<td>0.0377</td>
<td>0.0351</td>
</tr>
<tr>
<td>$C(r_m^m - \Delta d, \Delta c)(\times 10^{-5})$</td>
<td>2.0978 (0.2371)</td>
<td>2.3799</td>
<td>2.3583</td>
<td>2.3854</td>
</tr>
<tr>
<td>$C(\Delta d, \Delta c)(\times 10^{-5})$</td>
<td>2.4398 (0.1817)</td>
<td>2.7501</td>
<td>2.7286</td>
<td>2.8858</td>
</tr>
<tr>
<td>% of Amount of Risk by $C(\Delta d, \Delta c)$</td>
<td>14.014%</td>
<td>14.014%</td>
<td>14.014%</td>
<td>14.014%</td>
</tr>
</tbody>
</table>

Table 8: Non-DGP Model Parameter Choices (∗=annualized).

This table presents the non-DGP parameter choices and the derived parameter values. Following Campbell and Cochrane (1999), the AR(1) coefficient of the $s_t$ process ($\phi_s$) equals the AR(1) coefficient of the monthly log price dividend ratio. $r_{fCC}$ is the constant benchmark risk free rate and is chosen to match the average real 90-day Treasury bill rate, which is proxied by changes in log nominal 90-day Treasury index constructed by CRSP minus inflation rate continuously compounded; the nominal index is constructed by CRSP and the inflation rate is obtained from the Federal Reserve Bank of St. Louis. $\beta$ is the time discount parameter derived from the $r_{fCC}$ equation. Monthly data covers the period 1959/01-2014/06.

### Non-DGP parameters:

<table>
<thead>
<tr>
<th>Notation</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$</td>
<td>2</td>
</tr>
<tr>
<td>$\phi_s$</td>
<td>0.9499*</td>
</tr>
<tr>
<td>$r_{fCC}$</td>
<td>1.4854*</td>
</tr>
</tbody>
</table>

### Derived parameters:

<table>
<thead>
<tr>
<th>Notation</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.9952*</td>
</tr>
<tr>
<td>$S$</td>
<td>0.0964</td>
</tr>
<tr>
<td>$s_{max}$</td>
<td>-2.1701</td>
</tr>
</tbody>
</table>

This table presents 10 unconditional moments of financial variables from simulated and empirical datasets. **Moment Symbols** $E$, mean; $\sigma$, volatility; $ac$, first-order autocorrelation coefficient. Other details on data, models, model solutions, and simulations are described in Tables 6∼8. Bold values indicate that the simulation moment point estimates are within a 95% confidence interval of the empirical moments.

<table>
<thead>
<tr>
<th>Data SE M(1) M(2) M(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adapted Adapted</td>
</tr>
<tr>
<td>Campbell &amp; Bekaert &amp; Cochrane,1999 Engstrom,2017</td>
</tr>
<tr>
<td>This Paper</td>
</tr>
<tr>
<td>$s$ as State Variable</td>
</tr>
<tr>
<td>$n$ as State Variable</td>
</tr>
<tr>
<td>$b$ as State Variable</td>
</tr>
<tr>
<td>(19) * $E(r^m - rf), %$</td>
</tr>
<tr>
<td>(20) * $\sigma(r^m - rf), %$</td>
</tr>
<tr>
<td>(21)</td>
</tr>
<tr>
<td>(22)</td>
</tr>
<tr>
<td>(23) *</td>
</tr>
<tr>
<td>(24)</td>
</tr>
<tr>
<td>(25)</td>
</tr>
<tr>
<td>(26)</td>
</tr>
<tr>
<td>(27) *</td>
</tr>
<tr>
<td>(28) *</td>
</tr>
</tbody>
</table>
Table 10: Theoretical Models: Resolving the Endogenous Part of the Duffee Puzzle.

This table evaluates the abilities of three theoretical models to fit Facts (f)∼(j) related to the endogenous part of the Duffee Puzzle as established in Section 2. Other details on data, models, model solutions, and simulations are described in Tables 6∼8. Bold values indicate that the simulation moment point estimates are within a 95% confidence interval of the empirical moments.

<table>
<thead>
<tr>
<th></th>
<th>Data SE</th>
<th>M(1)</th>
<th>M(2)</th>
<th>M(3)</th>
<th>This Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Adapted</td>
<td></td>
<td>Adapted</td>
<td>Engstrom, 2017</td>
</tr>
<tr>
<td>s as State Variable</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>n as State Variable</td>
<td>-</td>
<td>-</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>b as State Variable</td>
<td>-</td>
<td>-</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

(f). \[\sigma(r^m - \Delta d) \,(I_{rece.} = 0)\] (0.0014) 0.0257 0.0288 0.0317
\[\sigma(r^m - \Delta d) \,(I_{rece.} = 1)\] (0.0056) 0.0264 0.0537 0.0502

(g). \[\sigma(r^m) \,(I_{rece.} = 0)\] (0.0013) 0.0251 0.0279 0.0308
\[\sigma(r^m) \,(I_{rece.} = 1)\] (0.0054) 0.0246 0.0530 0.0497

(h). \[C(r^m - \Delta d, \Delta c) \,(I_{rece.} = 0) \times 10^{-5}\] 2.0588* (0.1726) 2.2489 2.2284 2.3952
\[C(r^m - \Delta d, \Delta c) \,(I_{rece.} = 1) \times 10^{-5}\] 2.3637 (0.2354) 2.4519 2.5414 2.5621

(i). \[C(r^m, \Delta c) \,(I_{rece.} = 0) \times 10^{-5}\] 2.4504 (0.1687) 2.6192 2.5886 2.7672
\[C(r^m, \Delta c) \,(I_{rece.} = 1) \times 10^{-5}\] 2.3954 (0.1833) 2.8222 2.9116 2.6743

(j). \% of Amount of Risk by \[C(\Delta d, \Delta c) \,(I_{rece.} = 0)\] 15.980%*** (5.253%) 14.136% 14.248% 13.443%
\% of Amount of Risk by \[C(\Delta d, \Delta c) \,(I_{rece.} = 1)\] 1.322% (3.297%) 13.120% 12.717% 4.196%

Note: the significance levels for testing the equality of non-recession (I_{rece.} = 0) and recession (I_{rece.} = 1) moments (Wald Test given the covariance-variance matrix from the GMM estimations) are denoted as follow:

*** \(p < 0.01\), ** \(p < 0.05\), * \(p < 0.1\).
Table 11: Price Dividend Ratio Variance Decomposition.

This table presents 6 moments on the approximate variance decomposition results of the log price dividend ratio using the empirical dataset and three simulated datasets of the three theoretical models. The log price dividend ratio is estimated by univariate and multivariate regression models. **Moment Symbols** $b(x_t)$, linear regression coefficient estimate of state variable $x_t$ in a univariate or multivariate $pd$ regression, where the linear regression framework is denoted as $pd_t = f(\ldots, x_t, \ldots)$. The variance decomposition in a multivariate framework, VARC, is $\frac{\hat{b}(x)Cov(x, \hat{f})}{\text{Var}(\hat{f})}$ where $\hat{b}(x)$ is the coefficient estimate and $\hat{f}$ is the fitted dependent variable; the sum of VARCs across all variables is 100%; VARCs are shown in curly brackets. Bold (italics) values indicate that the simulation moment point estimates are within a 95% (99%) confidence interval of the empirical moments.

<table>
<thead>
<tr>
<th>Data</th>
<th>SE</th>
<th>M(1)</th>
<th>M(2)</th>
<th>M(3)</th>
<th>This Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>s as State Variable</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>n as State Variable</td>
<td>-</td>
<td>-</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>b as State Variable</td>
<td>-</td>
<td>-</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>$b(b_t)$: $pd_t = f(b_t)$</td>
<td>1.2807 (0.1292)</td>
<td>-</td>
<td>-</td>
<td>1.1146</td>
<td></td>
</tr>
<tr>
<td>$R^2$: $pd_t = f(b_t)$</td>
<td>12.900% (3.608%)</td>
<td>-</td>
<td>-</td>
<td>7.797%</td>
<td></td>
</tr>
<tr>
<td>$b(s_t)$: $pd_t = f(s_t, n_t, b_t, b_t^2)$</td>
<td>0.5815 (0.0384)</td>
<td>0.6693</td>
<td>0.5564</td>
<td>0.5368</td>
<td></td>
</tr>
<tr>
<td>VARC, %</td>
<td>{72.770%}</td>
<td>{100%}</td>
<td>{69.091%}</td>
<td>{64.317%}</td>
<td></td>
</tr>
<tr>
<td>$b(n_t)$: $pd_t = f(s_t, n_t, b_t, b_t^2)$</td>
<td>-0.1407 (0.0322)</td>
<td>-</td>
<td>0.3222</td>
<td>0.3121</td>
<td></td>
</tr>
<tr>
<td>VARC, %</td>
<td>{14.102%}</td>
<td>-</td>
<td>{30.909%}</td>
<td>{28.998%}</td>
<td></td>
</tr>
<tr>
<td>$b(b_t)$: $pd_t = f(s_t, n_t, b_t, b_t^2)$</td>
<td>0.5480 (0.1486)</td>
<td>-</td>
<td>-</td>
<td>0.8951</td>
<td></td>
</tr>
<tr>
<td>VARC, %</td>
<td>{13.386%}</td>
<td>-</td>
<td>-</td>
<td>{6.657%}</td>
<td></td>
</tr>
<tr>
<td>$b(b_t^2)$: $pd_t = f(s_t, n_t, b_t, b_t^2)$</td>
<td>2.3024 (0.9112)</td>
<td>-</td>
<td>-</td>
<td>0.9157</td>
<td></td>
</tr>
<tr>
<td>VARC, %</td>
<td>{-0.268%}</td>
<td>-</td>
<td>-</td>
<td>{0.029%}</td>
<td></td>
</tr>
<tr>
<td>$R^2$: $pd_t = f(s_t, n_t, b_t, b_t^2)$</td>
<td>49.587% (4.654%)</td>
<td>96.933%</td>
<td>62.012%</td>
<td>65.255%</td>
<td></td>
</tr>
</tbody>
</table>
Table 12: The Pricing of $b_t$ in Cross Section: Factor Loadings and Prices of Risk.


In the first panel, Panels A to D report multivariate factor loadings from regressions of each portfolio excess return on the market excess return ("Mkt"), the log surplus consumption ratio innovations ("sinnov"), the macroeconomic uncertainty innovations ("ninnov") and the procyclical consumption-dividend comovement innovations.

\[ R_{i,t} - R_{f,t} = \beta_{i,0} + \beta_{i,Mkt} (R_{Mkt,t} - R_{f,t}) + \beta_{i,sinnov} sinnov_t + \beta_{i,ninnov} ninnov_t + \beta_{i,binnov} binnov_t + \epsilon_{i,t}. \]

"5-1" denotes the difference between Portfolio 5 (highest in B/M or largest in size) and Portfolio 1 (lowest in B/M or smallest in size). "***" denotes 1% significance level, "**" 5% significance level and "*" 10% significance level. In the second stage, portfolio returns are regressed on the loadings, giving an estimate of the price of risk for each factor:

\[ E[R_i] - R_f = \lambda_0 + \lambda_{Mkt} \beta_{i,Mkt} + \lambda_{sinnov} \beta_{i,sinnov} + \lambda_{ninnov} \beta_{i,ninnov} + \lambda_{binnov} \beta_{i,binnov}. \]

"VARC" reports the variance decomposition (see Table [1] for details). Robust standard errors are shown in parentheses. Bold (italics) estimates have significance at the 5% (10%) level. N=556 months (1968/03-2014/06).

<table>
<thead>
<tr>
<th>Panel A. Multivariate Loadings on Market Excess Returns, $\beta_{i,Mkt}$</th>
<th>GROWTH</th>
<th>B/M 2</th>
<th>B/M 3</th>
<th>B/M 4</th>
<th>VALUE</th>
<th>5-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMALL</td>
<td>1.418***</td>
<td>1.2162***</td>
<td>1.0906***</td>
<td>1.0059***</td>
<td>1.0507***</td>
<td>-0.3680***</td>
</tr>
<tr>
<td>Size 2</td>
<td>1.3965***</td>
<td>1.1631***</td>
<td>1.0401***</td>
<td>0.9943***</td>
<td>1.1079***</td>
<td>-0.2886***</td>
</tr>
<tr>
<td>Size 3</td>
<td>1.3271***</td>
<td>1.1191***</td>
<td>1.0011***</td>
<td>0.9421***</td>
<td>1.0503***</td>
<td>-0.2768***</td>
</tr>
<tr>
<td>Size 4</td>
<td>1.2319***</td>
<td>1.0846***</td>
<td>1.0093***</td>
<td>0.9466***</td>
<td>1.0766***</td>
<td>-0.1553***</td>
</tr>
<tr>
<td>LARGE</td>
<td>0.9761***</td>
<td>0.9384***</td>
<td>0.8480***</td>
<td>0.8826***</td>
<td>0.9471***</td>
<td>-0.0290</td>
</tr>
<tr>
<td>5-1</td>
<td>-0.4426***</td>
<td>-0.2778***</td>
<td>-0.2425***</td>
<td>-0.1233***</td>
<td>-0.1036***</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B. Multivariate Loadings on Surplus Consumption Ratio (s) Innovations, $\beta_{i,sinnov}$</th>
<th>GROWTH</th>
<th>B/M 2</th>
<th>B/M 3</th>
<th>B/M 4</th>
<th>VALUE</th>
<th>5-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMALL</td>
<td>15.2272</td>
<td>11.875</td>
<td>12.4366*</td>
<td>10.5175</td>
<td>14.0103*</td>
<td>-1.2169</td>
</tr>
<tr>
<td>Size 2</td>
<td>10.9073</td>
<td>8.1214</td>
<td>10.9637*</td>
<td>10.2815</td>
<td>9.4299</td>
<td>-1.4774</td>
</tr>
<tr>
<td>Size 3</td>
<td>8.3100</td>
<td>6.1493</td>
<td>5.0391</td>
<td>5.0484</td>
<td>9.5382</td>
<td>1.2282</td>
</tr>
<tr>
<td>5-1</td>
<td>-5.9506</td>
<td>-2.6887</td>
<td>-2.7790</td>
<td>-4.5184</td>
<td>2.1238</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C. Multivariate Loadings on Macroeconomic Uncertainty (n) Innovations, $\beta_{i,ninnov}$</th>
<th>GROWTH</th>
<th>B/M 2</th>
<th>B/M 3</th>
<th>B/M 4</th>
<th>VALUE</th>
<th>5-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMALL</td>
<td>-0.0485*</td>
<td>-0.0489***</td>
<td>-0.0499***</td>
<td>-0.0653***</td>
<td>-0.0684***</td>
<td>-0.0199</td>
</tr>
<tr>
<td>Size 2</td>
<td>-0.0446*</td>
<td>-0.0415***</td>
<td>-0.0444***</td>
<td>-0.0666***</td>
<td>-0.0640***</td>
<td>-0.0193</td>
</tr>
<tr>
<td>Size 3</td>
<td>-0.0314</td>
<td>-0.0419***</td>
<td>-0.0458***</td>
<td>-0.0572***</td>
<td>-0.0648***</td>
<td>-0.0333* *</td>
</tr>
<tr>
<td>Size 4</td>
<td>-0.0342</td>
<td>-0.0354***</td>
<td>-0.0354*</td>
<td>-0.0405***</td>
<td>-0.0634***</td>
<td>-0.0292 *</td>
</tr>
<tr>
<td>LARGE</td>
<td>0.0038</td>
<td>-0.0306*</td>
<td>-0.0371***</td>
<td>-0.0427***</td>
<td>-0.0588***</td>
<td>-0.0546***</td>
</tr>
<tr>
<td>5-1</td>
<td>0.0523***</td>
<td>0.0183</td>
<td>0.0127</td>
<td>0.0226</td>
<td>0.0177</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel D. Multivariate Loadings on Div.-Cons. Comovement (b) Innovations, $\beta_{i,binnov}$</th>
<th>GROWTH</th>
<th>B/M 2</th>
<th>B/M 3</th>
<th>B/M 4</th>
<th>VALUE</th>
<th>5-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMALL</td>
<td>11.0184</td>
<td>7.0150</td>
<td>11.5856 *</td>
<td>7.3966</td>
<td>17.1700***</td>
<td>6.1575</td>
</tr>
<tr>
<td>Size 2</td>
<td>5.4487</td>
<td>6.8722</td>
<td>4.4826</td>
<td>1.9321</td>
<td>10.3278*</td>
<td>4.8792</td>
</tr>
<tr>
<td>Size 3</td>
<td>1.9982</td>
<td>1.4142</td>
<td>3.2624</td>
<td>5.8930</td>
<td>4.5245</td>
<td>2.5263</td>
</tr>
<tr>
<td>Size 4</td>
<td>1.3867</td>
<td>4.3867</td>
<td>5.3777</td>
<td>5.5067</td>
<td>7.1502</td>
<td>8.5369*</td>
</tr>
<tr>
<td>LARGE</td>
<td>4.9418</td>
<td>0.2228</td>
<td>-1.3690</td>
<td>5.4284</td>
<td>14.7410***</td>
<td>19.6829***</td>
</tr>
<tr>
<td>5-1</td>
<td>15.9603***</td>
<td>-6.7922</td>
<td>-12.9546***</td>
<td>-1.9682</td>
<td>-2.4349</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel E. Price of Risk (Second-Stage Fama-MacBeth)</th>
<th>$\lambda_{Mkt}$</th>
<th>$\lambda_{sinnov}$</th>
<th>$\lambda_{ninnov}$</th>
<th>$\lambda_{binnov}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Est.</td>
<td>-0.5577</td>
<td>-0.0144</td>
<td>-0.0697</td>
<td>0.6474</td>
</tr>
<tr>
<td>SE</td>
<td>(0.2682)</td>
<td>(0.0155)</td>
<td>(0.0413)</td>
<td>(0.0216)</td>
</tr>
<tr>
<td>VARC</td>
<td>38.55%</td>
<td>3.29%</td>
<td>41.75%</td>
<td>16.41%</td>
</tr>
<tr>
<td>$R^2$</td>
<td>41.40%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: The Decomposition of the Duffee Puzzle from the Empirical Analyses.

The top panel depicts the empirical estimates of the two components of amount of risk in the Duffee Puzzle. The solid black line depicts the dynamic conditional covariance of dividend growth and consumption growth (with instrument); the dashed red line depicts the dynamic conditional covariance of the non-dividend part of the market return and consumption growth (with instrument). The bottom panel depicts the time variation in the share of the consumption-dividend conditional covariance in the market return-consumption conditional covariance (namely, the amount of risk). The conditional covariance estimates in the two plots are obtained from the “DCC-qt” model estimation results as reported in Table 3. The shaded regions are the NBER recession months from the NBER website.
Figure 2: Annualized Conditional Volatility of the Two Consumption Shocks from the DGP Estimation.

The plot illustrates the magnitudes and dynamics of the two components of consumption growth volatility. The solid black line depicts the annualized conditional volatility contributed by the heteroskedastic event shock $\tilde{\omega}_{n,t+1}$, or $\sqrt{12\sigma_n^2 n_t}$. The dotted red line depicts the annualized volatility contributed by the homoskedastic fundamental shock $\tilde{\omega}_{c,t+1}$, or $\sqrt{12\sigma_c^2}$. The dashed blue line depicts the percentage of the total consumption variance explained by the heteroskedastic component, or $\frac{\sigma_n^2 n_t}{\sigma_n^2 n_t + \sigma_c^2}$. The shaded regions are the NBER recession months from the NBER website.
Figure 3: Annualized Conditional Volatility of the Two Dividend Shock Components from the DGP Estimation.

The plot illustrates the magnitudes and dynamics of the two components of the dividend growth volatility. The solid black line depicts the annualized conditional volatility contributed by the consumption fundamental shock $\tilde{\omega}_{c,t+1}$, or $\sqrt{12\sigma_{c}^{2}b_{t}^{2}}$. The dotted red line depicts the annualized volatility contributed by the dividend-specific shock $\tilde{\omega}_{d,t+1}$, or $\sqrt{12\sigma_{d}^{2}V_{d}}$. The dashed blue line depicts the percentage of the total dividend variance explained by the heteroskedastic component, or $\frac{\sigma_{d}^{2}b_{t}^{2}}{\sigma_{c}^{2}b_{t}^{2} + \sigma_{d}^{2}V_{d}}$. The shaded regions are the NBER recession months from the NBER website.
Figure 4: Economic Interpretations for the Fundamental Shock and the Event Shock.

The plot provides direct graphical evidence on the economic interpretations of the fundamental shock and the event shock in the new DGP (Equation (10)). The two shocks are estimated using a filtration-based maximum likelihood estimation methodology developed by Bates (2006); the estimation results are shown in Table 5. In this figure, the monthly filtered shock realizations—the fundamental shock $\hat{\omega}_c$ and the event shock $\hat{\omega}_n$—are summarized at the quarterly frequency and plotted against two business condition indicators. Plot A depicts the quarterly fundamental shock against the detrended consumption-wealth ratio introduced in Lettau and Ludvigson (2001). The magnitudes of the quarterly $\hat{\omega}_c$ (detrended $\hat{c}ay^{Q}$) are shown in the left (right) axis. Plot B depicts the quarterly event shock realizations. The shaded regions are the NBER recession quarters from the NBER website.
exp(s), surplus consumption ratio
P/D

P/D of the s-dimension

P/D

exp(s), surplus consumption ratio

n, macroeconomic uncertainty

Figure 5: Dependence of the Price Dividend Ratio on the State Variables s and n.

The top figure depicts the relationship between PD and exp(s), and the bottom figure depicts the relationship between PD and n. For M(1), price dividend ratio depends on s only; M(1) is depicted in solid black lines with circles. For M(2), price dividend ratio is sensitive to s and n; to explore the dependence of PD on s (n), I fix n (s) at its mean from the simulation, 0.3742 (-2.6595); M(2) is depicted in solid red lines with triangles. For M(3), price dividend ratio is sensitive to s, n and b. The dimension is reduced by fixing the other state variables at their mean and critical values (i.e., the 95% quantile value in the n_t simulation, 1.4246, the 5% quantile value in the s_t simulation, -3.5281, and the 5% quantile value in the b_t simulation, 0.3014); M(3) is depicted in blue lines with squares. Note that, as suggested by theory, the three solid lines in the top plot coincide; the two solids lines with triangles and squares in the bottom plot also coincide.
Figure 6: Dependence of the Price Dividend Ratio on the State Variable $b$.

The figure depicts the relationship between $PD$ and $b$ conditional on $(s, n)$. M(1) and M(2) are invariant of the new state variable $b$. For M(3), price dividend ratio is sensitive to $s$, $n$ and $b$. The multi-dimensional relationship is reduced by fixing the other state variables at their mean and critical values (i.e., the 95% quantile value in the $n_t$ simulation, 1.4246 and the 5% quantile value in the $s_t$ simulation, -3.5281); M(3) is depicted in blue lines with squares.