The Long-run Risks Model: What Differences Can an Extra Volatility Factor Make?

Guofu Zhou
Washington University[1]

and

Yingzi Zhu
Tsinghua University

Current Version: March, 2013

[1]This paper supersedes an early version of the paper circulated under title "A Long-run Risks Model with Long- and Short-run Volatilities: Explaining Predictability and Variance Risk Premium." We are grateful to Doron Avramov, Ravi Bansal, Jason Beeler, Geert Bekaert, Tim Bollerslev, Henry Cao, Long Chen, Peter Christoffersen, Bjørn Eraker, Satadru Hore, Raymond Kan (the CICF discussant), Dana Kiku (the EFA discussant), Hong Liu, Zhongjin Lu, Ivan Shaliastovich, Jack Strauss, George Tauchen, Liping Wang, Yaping Wang (the Five Star Finance Forum discussant), Hao Zhou, and seminar/conference participants at East China University of Science and Technology, National University of Singapore, Tsinghua University, CKGSB 2009 summer workshop, 2010 China International Conference in Finance, 2010 European Finance Association annual meetings, and 2011 Beijing Five Star Finance Forum.

Correspondence: Guofu Zhou, Olin School of Business, Washington University, St. Louis, MO 63130; e-mail: zhou@wustl.edu, phone: 314-935-6384; and Yingzi Zhu, School of Economics and Management, Tsinghua University, Beijing, China 100084; e-mail: zhuyz@sem.tsinghua.edu.cn, phone: 86-10-62786041.

The Long-run Risks Model:
What Differences Can an Extra Volatility Factor Make?

In this paper, we extend the long-run risks model of Bansal and Yaron (2004) by allowing both a long- and a short-run volatility components in the evolution of economic fundamentals. While the Bansal and Yaron model has received increasing attention recently, it has major difficulties in explaining a large negative market variance risk premium, the predictability of consumption growth by price-dividend ratio, the predictability of market volatility by price-dividend ratio, and the importance of discount rate risk relative to cash flow news. By adding one more volatility factor, our extension not only makes the new model consistent with the volatility literature that the stock market is driven by two, rather than one, volatility factors, but also resolves all the aforementioned challenging issues facing the original Bansal and Yaron model.

1. Introduction

The long-run risks model of Bansal and Yaron (henceforth BY, 2004) explains simultaneously the equity risk premium puzzle, the risk-free rate puzzle, the high level of market volatility, and many other stylized facts about the stock market, consumption and price-dividend ratio. Consequently, it is not surprising that it has attracted a lot of attention, with important subsequent studies by Bansal, Khatchatrian, and Yaron (2005), Bansal, Kiku, and Yaron (2007), Bansal, Kiku, and Yaron (2012), Chen, Collin-Dufresne, and Goldstein (2009), Eraker (2008), Hansen, Heaton, and Li (2008), Pakos (2008), Avramov and Hore (2009), Bansal, Dittmar, and Kiku (2009), Beeler and Campbell (2012), Constantinides and Ghosh (2011), Drechsler and Yaron (2011), and Ferson, Nallareddy and Xie (2012), among others.

However, there remain three major problems with the model. First, the BY model implies a much stronger predictability of consumption growth by price-dividend ratio than observed in the data, as pointed out by Bansal, Kiku and Yaron (henceforth BKY, 2007a). Second, the model implies a variance risk premium that is too small to match that of the options market, as shown by Drechsler and Yaron (2011). Third, the model asserts a much stronger predictive power of price-dividend ratio on future stock return volatility than is found in the data, as demonstrated by Beeler and Campbell (2012) in their extensive analysis of the BY model. The existing solutions to the first two problems remain piecemeal and incomplete, and the solution to the third problem is unknown. For example, BKY provides an improved version of the model that deals with the first problem, but their approach does not address the second. While Drechsler and Yaron (2011) provide a jump process solution to the second problem, their approach does not solve the first. A simple and important question is: Can the BY model be extended in such a way to overcome all the three problems simultaneously?

Another problem with the BY and BKY models, which has not received attention in the literature, is that the underlying volatility process is inconsistent with data. Recent studies on stock return volatility by Alizadeh, Brandt, and Diebold (2002), Chernov, Gallant, Ghysels and Tauchen (2003), Chacko and Viceira (2005), Adrian and Rosenberg (2008), Christoffersen, Jacobs, Ornthanalai and Wang (2008), and Lu and Zhu (2010), among others, show that there are two volatility factors in the stock market. The two factors are important not only in understanding the returns on options, but also in explaining the cross section returns on stocks. Hence, given the empirical facts from the large literature on volatility, it is

important to resolve the inconsistency of the one-factor model with the volatility dynamics of the stock market. At present, Bansal and Shaliastovich (2010) seems the only other study that uses a two-factor volatility model in the context of long-run risks models. They show that by separating short- and long-run consumption volatility, a discrete-time version of the two-factor volatility model helps to match the expectation hypothesis violations in bond and foreign exchange markets. In contrast to their study, we focus here on extending the more basic and fundamental BY and BKY models in a more general way, and to resolve all the aforementioned problems faced by the BY and BKY models.

Another motivation of our paper comes from the long-standing academic debate on the importance of discount rate news relative to the cash flow news in driving the variations of stock returns. Theoretically, the Campbell and Cochrane (1999) habit formation model focuses on the time varying discount rate, while the BY model emphasizes on long-run cash flow risk. Empirically, Campbell (1993) and Campbell and Vuolteenaho (2004) find that the large portion of stock market return variance is driven by changing discount rates, while Chen and Zhao (2009) argue that the results are far from conclusive due to the uncertainty in modeling stock return predictability. Empirical studies on the BY/BKY model, e.g. Beeler and Campbell (2012) and Avramov and Cederburg (2012), point out that the BY model implies small discount rate risk relative to the cash-flow risk. By introducing the two volatility factors, we show that the extended model has enough flexibility to allow the discount rate risk to play a greater role in driving stock return variance than the cash flow news.

Economically, how to we understand the two volatility components in the new model? First, economically, the long-run volatility factor is usually considered to be associated with business cycles (e.g., Stock and Watson 2002), and the short-run volatility factor is related to short-term risks. Adrian and Rosenberg (2008) find that a short-run volatility factor is important to capture both cross-sectional and time series properties of asset prices, and is closely related to the tightness of financial constraints. When market participants are constrained by capital requirement, their ability to absorb temporal imbalances in order flows is hampered and short-term volatility runs up. Li (2012), incorporating financial intermediary into the Lucas tree model, demonstrates that the short-run volatility is related to the leverage ratio of overall financing sector.

Second, as pointed out by Beeler and Campbell (2012), BKY's calibration differs from

BY's in that it increases the persistence of consumption volatility. In doing so, BKY increases the relative importance of the volatility state variable to the long-run risks, improving hence the fit of the data by reducing the high predictability of consumption and dividend growth rate from stock price-dividend ratio. In the mean time, however, the BKY model inevitably implies much greater predictive power of price-dividend ratio for future return volatility than found in the data. The one-factor volatility model, as we show analytically in the paper, implies that the high predictive power of stock price-dividend ratio for stock return volatility must co-exist with its high predictive power for both consumption and dividend volatilities. In other words, one-factor volatility models, such as those of the BY and BKY, cannot capture different degrees of predictability on stock return, consumption and dividend volatilities due to the fact that they all are governed by one single volatility factor. Hence, to explain empirically the relatively high predictability of stock price-dividend ratio for consumption volatility, and low predictability for stock return volatility, two volatility components are needed.

Thirdly, the BY model implies that the risk premium associated with volatility risk is small when the volatility component is highly persistent. Hence, to match the large and negative volatility risk premium, one has to introduce a volatility component with high frequency. Drechsler and Yaron (2011) show that by adding a jump component they can increase the magnitude of the variance risk premium. Intuitively, a jump process is a volatility component with least persistence. However, an addition of jump does not solve the volatility predictability issue mentioned above, because the jumps do not increase the state variable as an additional volatility component does. As a result, with the addition of one more volatility component, our new model can resolve the predictability problem raised by Beeler and Campbell (2012).

Methodologically, we cast our extension of the BY and BKY models in continuous-time. This allows us not only to avoid the occasional negative volatility problem in the discrete-time counterparts, but also to obtain approximate analytical solutions for many functions of economic interest, such as derivative prices, measures of volatility and the slope coefficients of various predictive regressions. These are useful not only for empirical analysis, but also for qualitative insights in understanding long-run risks models. In particular, we show analytically that the two volatility components are necessary to match the variance risk premium and to explain the volatility predictability. Empirically, using the generalized

method of moments (GMM) test, we reject the one-factor volatility BY/BKY model strongly in favor of the proposed two-factor model. In addition, our model yields reasonable estimates of the parameters and matches empirical moments well.

The rest of the paper is organized as follows. Section 2 provides a short review of the BY model. Section 3 provides the extension with both the long- and short-run volatility components, and solves various functions of interest approximate analytically. Section 4 presents an empirical analysis of the new model and compares it with the BY and BKY models. Section 5 concludes.

2. A short review of BY

In this section, we provide a short review of the BY model, which will be useful for understanding our extension and its comparison with other models.

The BY model assumes a representative investor who has Epstein-Zin-Weil recursive preferences (Epstein and Zin 1989, Weil 1989) and maximizes the life-time utility,

$$V_t = \left[C_t^{\frac{1-\gamma}{\theta}} + \beta (E_t[V_{t+1}^{1-\gamma}])^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}},$$

where C_t is consumption at time t, γ the coefficient of risk aversion, ψ the intertemporal elasticity of substitution (IES), and $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$. When $\theta = 1$, i.e., $\gamma = \frac{1}{\psi}$, the EZ utility becomes additive power utility; when $\theta < 1$, i.e., $\gamma > \frac{1}{\psi}$, as demonstrated by Epstein and Zin (1989), the investor prefers early resolution of uncertainty, which, intuitively, is due to the time aggregator being a convex function with power $\frac{1}{\theta} > 1$. As illustrated by Bansal and Yaron (2004), it is necessary to impose $\gamma > 1$ and $\psi > 1$ to account for the high stock market risk premium in the long-run risks model. For discussions on the magnitude of IES, refer to Hansen and Singleton (1982), Attanasio and Weber (1989) and Vissing-Jorgensen (2002), who argue that the IES is well over 1. However, Hall (1988) and Campbell (1999) estimate the IES to be well below 1. Although the debate on the magnitude of IES is far from settled, due to space constraint, we do not focus on the estimation of IES in this paper but rather follow Basal and Yaron (2004) on their estimation.

The investor makes his optimal portfolio decision under the discrete-time processes for consumption and dividends, the economic fundamentals, as follows:

$$
\begin{aligned}
\log(C_{t+1}/C_t) &= \mu + X_t + \sigma_t \eta_{t+1}, \\
X_{t+1} &= \alpha X_t + \varphi_x \sigma_t e_{t+1}, \\
\sigma_{t+1}^2 &= \sigma^2 + \kappa(\sigma_t^2 - \sigma^2) + \sigma_w w_{t+1}, \\
\log(D_{t+1}/D_t) &= \mu_d + \phi X_t + \varphi_d \sigma_t u_{t+1},
\end{aligned} \tag{1}
$$

where X_t is the long-run risk that affects both consumption and dividend growth, $\log(C_{t+1}/C_t)$ and $\log(D_{t+1}/D_t)$, and is persistent with autoregression (AR) coefficient α and volatility $\varphi_x \sigma_t$; the variance process σ_t^2 is also a time-varying AR process with coefficient κ; $\eta_{t+1}, e_{t+1}, w_{t+1}$ and u_{t+1} are independent shocks drawn from the standard normal distribution. ϕ is the dividend growth leverage ratio, φ_x and φ_d are volatility leverage ratios for long-run risks and dividend growth, *resp.*

The intuition behind the model is that X_t captures the long-run growth prospects of the economy. Shocks in both long-run X_t and short-run η_{t+1} drive the consumption growth and asset prices. With $\theta < 1$, the fear for adverse long-run growth requires high risk premium to compensate for. Since the long- and short-run shocks in dividend growth and asset returns can be very volatile, the BY model can successfully explain, among other stylized facts, the equity risk premium, the risk-free rate, and the volatility of the market return.

However, there are some strong implications from the original BY model that are inconsistent with the data. Empirically, the price-dividend ratio has little power in predicting consumption growth, but the model implies the opposite. BKY address this issue by increasing the persistence of the volatility to make it a more important factor, so the relative importance of long-run consumption risk factor in the price-dividend ratio can be reduced, so is its predictability on consumption growth. However, this persistent volatility factor implies much greater predictive power of stock price-dividend ratio for future stock return volatility than is found in the data, as pointed out by Beeler and Campbell's (2012) extensive study. In addition, the increase in the persistence of the volatility entails a small (virtually zero) variance risk premium that is inconsistent with the data. Drechsler and Yaron (2011) allow jumps in the volatility process that is capable of explaining the large negative variance risk premium, but their extension does not resolve the predictability problem. Since multi-factor stochastic volatility models are necessary in capturing volatility risk premium and the

volatility term structure dynamics (see, e.g., Christoffersen, Jacobs, Ornthanalai and Wang, 2008, Lu and Zhu, 2010, and Zhou and Zhu, 2011), we adopt this framework in this paper to allow two volatility factors in the long-run risks model.

3. The new long-run risks model

In this section, we first motivate our dynamic processes for the state variables in the new long-run risks model, and then solve the model in terms of the state variables. Subsequently, we provide approximate analytical solutions to functions of interest: the consumption-wealth ratio, market prices of risks, price-dividend ratio, and the market return volatility. Then we derive analytical results for the predictability regression coefficients of the excess return, consumption and dividend growth rate and their volatilities, as well as variance risk premium.

3.1. The model and solution

Our model extends the BY and BKY models in the continuous-time framework with two volatility factors. Parallel to the discrete-time model (1), we consider the following model for the consumption and dividends processes and their related variables,

$$\begin{aligned}
\frac{dC_t}{C_t} &= (\mu + X_t)dt + \sqrt{V_{1t}\delta_c + V_{2t}(1-\delta_c)}dZ_{1t}, \\
dX_t &= -\alpha X_t dt + \varphi_x \sqrt{V_{1t}\delta_x + V_{2t}(1-\delta_x)}dZ_{2t}, \\
\frac{dD_t}{D_t} &= (\mu_d + \phi X_t)dt + \varphi_d \sqrt{V_{1t}\delta_d + V_{2t}(1-\delta_d)}dB_t + \sigma_{dc}\sqrt{V_{1t}\delta_c + V_{2t}(1-\delta_c)}dZ_{1t}, \\
&\quad + \sigma_{dx}\sqrt{V_{1t}\delta_x + V_{2t}(1-\delta_x)}dZ_{2t} + \sigma_{dv}\sqrt{V_{1t}}dw_{1t} + \sigma_{dv2}\sqrt{V_{2t}}dw_{2t}, \\
dV_{1t} &= \kappa_1(\bar{V}_1 - V_{1t})dt + \sigma_1\sqrt{V_{1t}}dw_{1t}, \\
dV_{2t} &= \kappa_2(\bar{V}_2 - V_{2t})dt + \sigma_2\sqrt{V_{2t}}dw_{2t}, \qquad 0 < \kappa_1 < \kappa_2,
\end{aligned} \tag{2}$$

where $dZ_{1t}, dZ_{2t}, dB_t, dw_{1t}$ and dw_{2t} are independent Brownian motions. This model specification, which sets a convenient framework for our GMM estimation and model comparison study in later sections, is the most general one for two-factor models that allow for analytical solutions, and nests both continuous-time BY and BKY models[1]. When

[1]In the original BY/BKY models, variance is modelled as Gaussian process as opposed to square-root process in this paper. Hence we refer to BY/BKY models as long-run risks models with one-factor square-root variance process.

$\delta_x = 1, \delta_c = 1, \delta_d = 1$, and $\sigma_{dc} = \sigma_{dx} = \sigma_{dv} = \sigma_{dv2} = 0$, the model reduces to the continuous-time BY model. When $\delta_x = 1, \delta_c = 1, \delta_d = 1$, and $\sigma_{dc} = \sigma_{dv2} = 0$, it becomes the continuous-time BKY model. Further, if we set $\delta_x = 1, \delta_c = 1, \delta_d = 1$ and $\sigma_{dv2} = 0$, we get the most general one-factor model. Note that in the new model, we keep the dividend leverage ratio ϕ and long-run risks volatility leverage ratio φ_x as in BY, while the dividend leverage ratio becomes $\sqrt{\varphi_d^2 + \sigma_{dc}^2 + \sigma_{dx}^2 + \sigma_{dv}^2 + \sigma_{dv2}^2}$. As shown in BY, the numerical calibration of these parameters is important for the model performance. Hence, in our empirical study we keep these parameters as close to BY as possible.

The key feature of the new model is that the consumption growth has a variance level of $V_{1t}\delta_c + V_{2t}(1 - \delta_c)$, a convex combination of the long- and short-run variances V_{1t} and V_{2t}. This convex combination decomposes the total variance into two plausible components. The same variance decomposition, adjusted by the volatility leverage factor φ_x, is also applied to the long-run risk X_t. The dividend growth process is treated similarly, except that it allows for various covariations with C_t and X_t, as in the BKY model. The two components of volatility, as derived from model, will enter into the stock price volatility, which helps to match the market variance premium and the volatility predictability with the data. Further, the long- and short-run variances follow two independent standard square-root Heston (1993) processes, which avoid the negative variance problem of Gaussian specification by BY/BKY models.

There are several reasons for selecting the two volatility factors as opposed to jumps. First, although a jump-diffusion process can potentially solve the market variance risk premium problem (see, e.g., Drechsler and Yaron, 2011), Christoffersen, Jacobs, Ornthanalai and Wang (2008) show that valuation models with two volatility factors fit option prices better than a jump-diffusion model. The superior performance is partly attributable to its improved ability to model the smirk and the path of spot volatility, but its most distinctive feature is its ability to model the volatility term structure. Second, to resolve various predictability issues raised in Beeler and Campbell (2012), as shown later, we need to introduce at least one more volatility factor to account for the observed volatility predictability pattern of the market return, consumption, and dividend by the stock price-dividend ratio, while adding jumps do not help on this. Third, the jumps do not impact on the return predictability hence no impact on the variance decomposition. As mentioned in introduction, we target at a more flexible model allowing for discount rate risk driving the unexpected

return variance. Therefore, we do not introduce jumps into our model for simplicity.

In their study of bond and foreign exchange markets, Bansal and Shaliastovich (2010) use a two-factor volatility model in the long-run risks set-up. Their model is a special case of the model here with $\delta_x = 0, \delta_c = 1$. From an economic standpoint, Bansal and Shaliastovich (2010) provide an intriguing motivation for two-factor volatility structure by introducing uncertainty about measures of future consumption growth, and showing that the additional volatility component can arise through learning from signals with time-varying precision. However, as shown below, we need a more general form of the two-factor volatility model of Equation (2) to resolve the empirical difficulties facing the BY and BKY models.

To solve the equilibrium prices and other quantities of interest, following the BY model, we use the Epstein-Zin-Weil preference, but in continuous-time. Based on Duffie and Epstein (1992), we can define the intertemporal value function recursively by

$$J_t = E_t[\int_t^T f(C_s, J_s)ds]. \tag{3}$$

Thus the representative investor's objective is to choose consumption to optimize the value function, that is,

$$J_t = \max_{\{C_s\}} E_t[\int_t^T f(C_s, J_s)ds], \tag{4}$$

where $f(C, J)$ is a *normalized aggregator* related to current consumption C_t and continuation value function J_t, and is given by

$$f(C_t, J_t) = \frac{\beta}{1 - \frac{1}{\psi}}(1-\gamma)J[(\frac{C}{((1-\gamma)J)^{\frac{1}{1-\gamma}}})^{1-\frac{1}{\psi}} - 1], \tag{5}$$

with β the rate of time preference, $\gamma > 0$ the relative risk aversion, and $\psi > 0$ the intertemporal elasticity of substitution (IES). If we set $\psi = 1/\gamma$ in (5), as shown by Duffie and Epstein (1992), we obtain the standard additive expected utility of constant relative risk aversion (CRRA).

In our model setting, the value function $J = J(W_t, X_t, V_{1t}, V_{2t})$ is a function of wealth and three state variables. The HJB equation for J is

$$\max_{\{C\}}\{f(C, J) + \mathcal{A}^c J\} = 0 \tag{6}$$

where $\mathcal{A}^c$ is the infinitesimal generator[2] associated with vector process $(C_t, X_t, V_{1t}, V_{2t})$ de-

[2]See for derivation details are in Appendix A.1.

fined in Equation (2). Conjecturing a solution for J of the following form,

$$J(W_t, X_t, V_{1t}, V_{2t}) = \exp(A_0 + A_1 X_t + A_2 V_{1t} + A_3 V_{2t}) \frac{W_t^{1-\gamma}}{1-\gamma}, \tag{7}$$

we use standard log-linear approximation, which Campbell (1993) develops in discrete-time, and Chacko and Viceira (2005) use first in continuous-time. Specifically, let g_1 be the long-term mean of the consumption-wealth ratio,[3]

$$g_1 = \exp(E[c_t - \omega_t]), \tag{8}$$

where the lowercase variables are the log variables. With the standard log-linear approximation we have

$$\frac{C_t}{W_t} = \exp(c_t - \omega_t) \approx g_1 - g_1 \log g_1 + g_1 \log(C_t/W_t). \tag{9}$$

A key step is to derive consumption-wealth ratio explicitly in terms of state variables to substitute out the wealth in J of Equation (7), because the HJB equation is most conveniently expressed with respect to consumption process as in Equation (6) due to the consumption-based formulation in our model. In contrast, Campbell (1993) uses the consumption-wealth ratio to substitute out the consumption and formulates an ICAPM type of pricing model. Theoretically these two formulations are equivalent, although the former emphasizes on the time series property while the later on the cross-sectional implication of asset prices.

We show in Appendix A.1 that, with the log-linear approximation, the first order condition

$$f_C = J_W$$

leads to the consumption-wealth ratio as

$$\frac{C_t}{W_t} = \beta^{\psi} \exp\{(A_{0a} + A_{1a} X_t + A_{2a} V_{1t} + A_{3a} V_{2t})\}, \tag{10}$$

where $A_{ia} = A_i \frac{1-\psi}{1-\gamma}$ for $i = 0, 1, 2, 3$. Substituting out the wealth in J conjectured in Equation (7), we can solve the HJB equation to obtain the value function. The solution is approximate in general and exact when $\psi = 1$. Armed with this solution, we are ready to solve for various functions of interest, which are also approximate in general and exact when $\psi = 1$.[4] In particular, we have

$$A_1 = \frac{1-\gamma}{(g_1 + \alpha)\psi}$$

[3] It can be solved endogenously once the model parameters are known. Appendix A.3 provides the details.

[4] As found in similar approximations elsewhere as well as our own studies, the approximate solution is accurate around commonly calibrated parameters.

which is negative when $\gamma > 1$. In addition, A_2 and A_3 are both positive, which means that a rise in long-run consumption growth expectation increases the value function, while a rise in consumption volatility lowers the value function.

3.2. Consumption-wealth ratio

We examine the model implied consumption-wealth ratio, where wealth is defined as the present value of future consumption stream. Based on Equation (10), the consumption-wealth ratio is loglinear in the state variables, and has similar functional form as in the BY model. In particular,

$$A_{1a} = -\frac{1 - \frac{1}{\psi}}{g_1 + \alpha}, \tag{11}$$

which is exactly the same as the continuous analogue of BY's $-A_1$.[5] Hence, the same interpretation applies that, when $\psi > 1$, $A_{1a} < 0$, which means that a rise in expected consumption growth lowers the consumption-wealth ratio, and the substitution effect dominates; when $\psi > 1$, a rise in expected consumption growth increases the consumption-wealth ratio, and the income effect dominates. In addition, the consumption-wealth ratio is more sensitive to the expected growth rate when the persistence of expected growth shocks, measured by $1/\alpha$, increases.

However, there are now two volatilities in Equation (10). This is expected since they are in the basic dynamics equations. Due to their symmetric formulations entered into the consumption and dividends dynamics, the two volatility components impact on the ratio in the same way, with proportional coefficients A_{2a} and A_{3a} depending on the volatility parameters via similar functional forms. When $\psi > 1$, both A_{2a} and A_{3a} are positive. The same intuition of the original BY model about volatility also holds here. For example, a rise in either of the volatilities will make consumption more volatile, which lowers asset valuations and increases the risk premia on all assets. In addition, an increase in the persistence of volatility shocks, that is, a decrease in either κ_1 or κ_2, will magnify the effects of volatility shocks on valuation ratios, since the investor would perceive changes in economic uncertainty as being long lasting.

[5]Note that Bansal and Yaron (2004) use the ratio of wealth to consumption, but we use the ratio of consumption to wealth. Hence our A_{ia}'s have the opposite sign of theirs. The same applies to the price-dividend ratio below.

Empirically, the aggregate wealth, which consists of both asset wealth and human capital, is not directly observable. However, Jagannathan and Wang (1996) and Lettau and Ludvigson (2001), among others, argue that it may be expressed in observable variables of consumption, asset wealth and labor income under various assumptions. In this paper, we focus on the stock price-dividend ratio and its predictability which are empirically observable.

3.3. Risk-free rate and market prices of risks

Define a state price process or pricing kernel π_t for any security with dividend process D_t and price process P_t as

$$P_t = \frac{1}{\pi_t} E_t \left[\int_t^\infty \pi_s D_s ds \right]. \tag{12}$$

In particular, for a risk-free asset with risk-free rate r_f, we have

$$-r_f dt = E_t \left[\frac{d\pi_t}{\pi_t} \right].$$

The Euler equation can be expressed in a differential form

$$E_t \left(\frac{dP_t}{P_t} \right) + \frac{D_t}{P_t} dt = r_f dt - E_t \left[\frac{d\pi_t}{\pi_t} \frac{dP_t}{P_t} \right]. \tag{13}$$

Duffie and Epstein (1992) identify a state price process for the above defined recursive utility as[6]

$$\pi_t = \exp \left[\int_0^t f_J(C_s, J_s) ds \right] f_C(C_t, J_t). \tag{14}$$

The wealth process is defined as the present value of consumption stream as

$$W_t = \frac{1}{\pi_t} E_t \left[\int_t^\infty \pi_s C_s ds \right]. \tag{15}$$

In our long-run risks model with two volatility factors, the pricing kernel is projected onto the risk space spanned by $\{Z_{1t}, Z_{2t}, w_{1t}, w_{2t}\}$ in Equation (2), and can be expressed as

$$\frac{d\pi_t}{\pi_t} = -(r_f dt + \lambda_1 dZ_{1t} + \lambda_2 dZ_{2t} + \lambda_3 dw_{1t} + \lambda_4 dw_{2t}), \tag{16}$$

[6]The intuition of the equation is apparent when one notes that it is equivalent to the standard Euler equation in discrete-time,

$$1 = E_t \left\{ \left[\beta \left(\frac{C_{t+1}}{C_t} \right)^{-1/\psi} \right]^\theta \left[\frac{1}{(1 + R_{w,t+1})} \right]^{1-\theta} (1 + R_{i,t+1}) \right\},$$

where $R_{w,t}$ the wealth portfolio return.

where r_f is the risk-free rate, and λ_i's (i=1, 2, 3, 4) are the market prices of risks corresponding to Z_{1t}, Z_{2t}, w_{1t} and w_{2t}, and can be derived by applying Ito's Lemma to Equation (14). The derivation details are in Appendix A.1. The risk-free rate is an affine function of state variables

$$r_f = r_0 + r_1 X_t + r_2 V_{1t} + r_3 V_{2t}, \tag{17}$$

with the parameters given in Equation (A16). The market prices of risks are in Equation (A17).

First, we note that $r_1 = \frac{1}{\psi} > 0$, which implies that the risk-free rate rises with higher expectation of consumption growth. Alternatively,

$$\frac{\partial E_t[(C_{t+1} - C_t)/C_t]}{\partial r_f} = \psi,$$

which says that a one percent increase in risk-free rate induces ψ percent increase in consumption growth. Further, when $\gamma = \frac{1}{\psi}$, only λ_1 is non-zero and risk premium is not correlated with expected consumption growth, hence risky asset return is related to expected consumption growth only through r_f. The above equation can also be written as

$$\psi = \frac{dE_t[(C_{t+1} - C_t)/C_t]}{dE_t[R_{t+1}]},$$

which can be used to empirically measure IES (e.g., Hansen and Singleton 1982, Hall 1988, Vissing-Jorgensen 2002).

Second, the maximal Sharpe ratio is the conditional volatility of the pricing kernel innovation, expressed as

$$\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_4^2}.$$

The market prices of risks for compensating the volatility risks are λ_3 and λ_4. When $\gamma = \frac{1}{\psi}$, the case for power additive utility, only λ_1 is non-zero. In this case, all the risk premia other than that of the instantaneous consumption growth become zero, and hence it will not be possible for standard models, such as the Breeden (1979) CCAPM model, to match the market risk premium. However, the long-run risks models with $\gamma > 1/\psi$, including BY/BKY and our new two volatility factor model, can produce the positive market prices for the long-run risks, and negative market prices for variance risks, both of which, combined with the stock market's positive beta to long-run risks and negative beta to volatility risks, contribute positively to solve equity risk premium puzzle. Hence, the relative magnitudes of

λ_i's quantify the different contributions to the maximal Sharpe ratio by various risk factors, and will be examined empirically later with comparison to BY model.

Intuitively, consistent with the empirical evidence, the negative sign of the variance risk premium indicates that investors regard increases in market volatility as unfavorable shocks to the investment opportunity. In comparison with the BY and BKY models, the market variance risk premium is determined by both the long- and short-run volatilities. As will be clear later, because of the rich dynamics of these two components, the associated parameters can be chosen in such a way that the model can explain the market variance risk premium along with other facts, while the previous models fail to do so.

3.4. Price-dividend ratio

Now we derive price-dividend ratio from the market portfolio return defined as

$$r_{m,t}dt \equiv E_t\left(\frac{dP_t}{P_t}\right) + \frac{D_t}{P_t}dt = r_f dt - E_t\left[\frac{d\pi_t}{\pi_t}\frac{dP_t}{P_t}\right] \tag{18}$$

where $r_{m,t}$ is the continuous compound market portfolio return. Campbell and Shiller (1988) show that, if the price-dividend ratio is stationary, the budget constraint (in discrete time)

$$P_{t+1} = (P_t - D_t)(1 + R_{m,t+1})$$

can be approximated by

$$d_t - p_t = \sum_{j=1}^{\infty} \rho^j (r_{m,t+j} - \Delta d_{t+j}) + const \tag{19}$$

where d_t and p_t are logarithms of D_t and P_t, and *const* is a constant related to the linearization parameter. This relation provides clear intuitive interpretation for the price-dividend predictability. The price-dividend ratio may be obtained by summation of expected future market returns and dividend growth, as demonstrated in the right hand side of Equation (19). However, the direct computation can be cumbersome. In our model setting, we derive the price-dividend ratio in a more convenient way as shown in Appendix A2, which shows that

$$\frac{D_t}{P_t} = \exp\{(A_{0m} + A_{1m}X_t + A_{2m}V_{1t} + A_{3m}V_{2t})\}, \tag{20}$$

where A_{im}'s are constants given in Equation (A22), (A23) and (A24) in Appendix A.2. In particular,

$$A_{1m} = -\frac{\phi - \frac{1}{\psi}}{g_{1m} + \alpha} \tag{21}$$

with g_{1m}, similar to g_1, as the long-term mean of D_t/P_t. Note that when the dividend leverage ratio $\phi > 1$, we have $|A_{1m}| > |A_{1a}|$, which says that news about expected growth rate leads to a larger reaction in the price of the dividend claim than in the price of the consumption claim. In addition, when ϕ is big enough, A_{2m} and A_{3m} are positive, resulting in the well-known volatility leverage effect, i.e., shocks to returns are negatively correlated with shocks to variance process.

The price process is obtained by applying Ito's Lemma to Equation (20)

$$\frac{dP_t}{P_t} = [c_3 + c_4 X_t + c_5 V_{1t} + c_6 V_{2t}]dt + \sqrt{c_1 V_{1t} + c_2 V_{2t}} dZ_t,$$

where c_i's ($i = 1$ to 6) are constants given in (A25) and (A26). dZ_t is a new Brownian motion defined accordingly, and hence the variance of the price process is

$$V_t = c_1 V_{1t} + c_2 V_{2t}, \tag{22}$$

with c_1 and c_2 defined in Equation (A25) of Appendix A.2. Equation (22) shows that the market volatility is a simple linear function of the long- and short-run volatilities of the economy. This provides an economic interpretation of the volatility components that are identified econometrically in the option literature.

3.5. Predictability of excess returns, consumption and dividends

To examine predictability of the variables, we, following Beeler and Campbell (2012), consider the following three K-period regressions

$$(r_{t+j} - r_{f,t+j}) + \cdots + (r_{t+j+K} - r_{f,t+j+K}) = \alpha_{jK} + \beta(p_t - d_t) + \epsilon_{jKt}, \tag{23}$$

$$\Delta c_{t+j} + \cdots + \Delta c_{t+j+K} = \alpha_{jK} + \beta(p_t - d_t) + \epsilon_{jKt}, \tag{24}$$

and

$$\Delta d_{t+j} + \cdots + \Delta d_{t+j+K} = \alpha_{jK} + \beta(p_t - d_t) + \epsilon_{jKt}, \tag{25}$$

where r and r_f are the stock market rate of return and risk-free rate, respectively; and c_t and d_t are logarithms of consumption and dividends. To explain the observed regression patterns, our idea is to derive the regression slope coefficients as functions of the model parameters that can be chosen in such a way to make the model implied regression slope coefficients match closely with those of the data.

We provide the formulas for $K = 1$ only for notational simplicity, while the general case is a straightforward extension. As $K = 1$, the regressors of the above three regressions all have the generic functional form of

$$dY_t = [a_0 + a_1 X_t + a_2 V_{1t} + a_3 V_{2t}]dt + \sqrt{b_1 V_{1t} + b_2 V_{2t}} dZ_t, \tag{26}$$

where dY_t corresponds to excess return $d \ln P_t + \frac{D_t}{P_t} - r_f dt$, consumption growth $d \ln C_t$ and dividend growth $d \ln D_t$, respectively. The parameter a_i's are given in Equations (A29) to (A31) in Appendix A.4 for stock market excess return, consumption growth, dividend growth.

The regression slope coefficient is then analytically obtained (see Appendix A.4) as

$$\beta = \frac{\mathrm{Cov}(\Delta_\tau y, p - d)}{\mathrm{Var}(p - d)}, \tag{27}$$

where

$$\begin{aligned}
\mathrm{Cov}(\Delta_\tau y, p - d) &= -\Big[a_1 A_{1m} \frac{\sigma_x^2}{2\alpha^2}(1 - e^{-\alpha\tau}) + a_2 A_{2m} \frac{\sigma_1^2 \bar{V}_1}{2\kappa_1^2}(1 - e^{-\kappa_1 \tau}), \\
&\qquad + a_3 A_{3m} \frac{\sigma_2^2 \bar{V}_2}{2\kappa_2^2}(1 - e^{-\kappa_2 \tau})\Big] \\
\mathrm{Var}(p - d) &= A_{1m}^2 \frac{\sigma_x^2}{2\alpha} + A_{2m}^2 \frac{\sigma_1^2 \bar{V}_1}{2\kappa_1} + A_{3m}^2 \frac{\sigma_2^2 \bar{V}_2}{2\kappa_2},
\end{aligned}$$

Equation (27) shows that with two volatility components rather than one, the covariance function in the numerator of β depends on three rather than two factors, precisely how the volatility components and the model parameters contribute to the degree of predictability.

3.6. Predictability of volatility: excess returns, consumption and dividends

In this subsection, we derive analytically the predictive coefficients of volatility by excess returns, consumption growth and dividend growth, which are not matched well in BY/BKY models as demonstrated in Beeler and Campbell (2012). In this paper, we use the same volatility measure as in Beeler and Campbell (2012). There are two steps in computing this measure. First, we run an AR(1) regression of each variable of interest y_{t+1} as

$$y_{t+1} = b_0^{vol} + b_1^{vol} y_t + u_{t+1}, \tag{28}$$

where y_{t+1} is the excess return or consumption growth or dividend growth. Second, the K-period realized volatility is defined as the sum of the absolute values of the residuals,

$$\text{Vol}_{t,t+K-1} = \sum_{k=0}^{K-1} |u_{t+k}| \tag{29}$$

over K periods with K, as before, the horizon of interest.

Then, the predictability of volatility is examined from the regression of the log of K-period realized volatility on the logarithm of price-dividend ratio,

$$\ln[\text{Vol}_{t+1,t+K}] = \alpha + \beta_{vol}(p_t - d_t) + \xi_t. \tag{30}$$

To match the volatility predictability implied by the data, we need to solve the continuous-time version of the regression slope.

First, we note that the innovation process u_{t+1} in the AR(1) process (28) is a discrete-time version of dZ_t in equation (26). Thus, in the continuous limit, the discrete realized volatility defined in Equation (29) can be written as

$$\text{Vol}_{t,t+\tau} = \int_0^\tau \sqrt{V_t}|dZ_t| = \frac{2}{\sqrt{2\pi}} \int_0^\tau \sqrt{V_t}dt, \tag{31}$$

where we have used the relation

$$|dZ_t| = \frac{2}{\sqrt{2\pi}}dt$$

for a standard Brownian Motion Z_t. Then, a key step in approximating the log of τ period realized volatility in Equation (31) is to use the following approximate equality,

$$\ln \int_0^\tau \sqrt{V_t}|dZ_t| \approx Const + \frac{1}{2\tau} \int_0^\tau \frac{V_t}{\bar{V}}dt, \tag{32}$$

where $\bar{V}$ is the unconditional mean of V_t and τ the horizon of interest similar to K (see Appendix A.5 for derivation). Then, we can obtain approximately the volatility slope,

$$\beta_{vol} = \frac{\text{Cov}(\ln[\text{Vol}_{t+1,t+K}], p-d)}{\text{Var}(p-d)}, \tag{33}$$

where

$$\text{Cov}(\Delta_\tau y, p-d) = -\frac{1}{2\tau\bar{V}} \left[b_1 A_{2m} \frac{\sigma_1^2 \bar{V}_1}{2\kappa_1^2}(1 - e^{-\kappa_1 \tau}) + b_2 A_{3m} \frac{\sigma_2^2 \bar{V}_2}{2\kappa_2^2}(1 - e^{-\kappa_2 \tau}) \right],$$

with $\text{Var}(p-d)$ as given by Equation (A33), and b_1 and b_2 given by $b_1 = c_1$ and $b_2 = c_2$ in the case of excess return volatility; $b_1 = \delta_c$, $b_2 = 1 - \delta_c$ in the case of consumption growth volatility; and

$$b_1 = \varphi_d^2 \delta_d + \sigma_{dc}^2 \delta_c + \sigma_{dx}^2 \delta_x + \sigma_{dv}^2, \quad b_2 = \varphi_d^2(1-\delta_d) + \sigma_{dc}^2(1-\delta_c) + \sigma_{dx}^2(1-\delta_x) + \sigma_{dv2}^2$$

in the final case of dividend growth volatility. Equation (33) shows nicely how the volatilities are predicted by the price-dividend ratio.

3.7. Variance risk premium

In this subsection, we follow Drechsler and Yaron (2011), among others, to define the variance risk premium (VRP) as the difference between the the objective or physical expectation and the risk-neutral expectation of the variance of aggregate stock market return over a period τ_0. The risk-neutral expectation of variance is also known as variance swap rate (VS), which can be replicated by model-free method utilizing options on all strike prices. Variance swap contracts are popular OTC volatility derivatives betting on realized variance (RV) against pre-determined VS rate (see Egloff, Leppold and Wu 2009). The VIX index, which has been published by CBOE (Chicago Board of Exchange) since 2003, is the square root of the VS rate on S&P 500 index with 30-day maturity (see CBOE 2009, and Zhang and Zhu 2006 for details). In fact, as Lu and Zhu (2010) and many others show, the VIX is almost always bigger than the realized volatility because investors with more risk aversion are willing to pay a positive risk premium to long stock market volatility. For that reason, VIX is referred to as the market's fear gauge by financial commentators.

Following Zhang and Zhu (2006), we first define the time t expected future realized variance over period τ_0 under the physical measure, which is denoted as RV_t. Using the market variance process defined in Equation (22), we have[7]

$$\begin{aligned} RV_t &= \tfrac{1}{\tau_0} \quad E_t^P[\int_t^{t+\tau_0} V_s ds] = E_t^P[\int_t^{t+\tau_0} \frac{1}{\tau_0} \sum_{i=1}^{2} c_i V_{is} ds] \\ &= \sum_{i=1}^{2} c_i \int_t^{t+\tau_0} ds \frac{1}{\tau_0} E_t^P[V_{is}]. \end{aligned} \tag{34}$$

where $E^P[\cdot]$ denotes the expectation under physical probability. Using the volatility factor dynamics defined in Equation (2), we have

$$E_t^P[V_{is}] = \bar{V}_i + (V_{it} - \bar{V}_i)e^{-\kappa_i(s-t)}, \tag{35}$$

[7]The RV is annualized, so is the variance swap rate VS as well as the VRP. The VRP data provided in table is annualized with a scaling factor 10000/12.

for $i = 1, 2$. Substituting Equation (35) into Equation (34), we have

$$\mathrm{RV}_t = \sum_{i=1}^{2} c_i (A_i^P + B_i^P V_{it}), \tag{36}$$

where A_i^P and B_i^P $(i = 1, 2)$ are constants given by

$$A_i^P = \bar{V}_i \left[1 - \frac{1 - e^{-\kappa_i \tau_0}}{\kappa_i \tau_0} \right], \qquad B_i^P = \frac{1 - e^{-\kappa_i \tau_0}}{\kappa_i \tau_0}.$$

Then we derive the time t expected future realized variance over time period τ_0 under the risk-neutral probability. The market prices of risk for V_{1t} and V_{2t} are λ_3 and λ_4 of Equation (A17), hence the associated risk premia are $\nu_1 V_{1t}$ and $\nu_2 V_{2t}$, which are proportional to V_{1t} and V_{2t}, and the risk-neutral processes are also square-root processes. We show in Appendix A.6 that the variance swap rate is given by

$$\mathrm{VS}_t = \sum_{i=1}^{2} c_i (A_i^Q + B_i^Q V_{it}), \tag{37}$$

with A_i^Q and B_i^Q defined in Equation (A42).

Hence the VRP, defined as the difference between RV of Equation (36) and VS rate of Equation (37), can be expressed as

$$VRP \equiv \mathrm{RV}_t - \mathrm{VS}_t = \sum_{i=1}^{2} c_i [(A_i^P - A_i^Q) + (B_i^P - B_i^Q) V_{it}], \tag{38}$$

where the coefficients are analytically determined by the model parameters. This equation is convenient for our empirical study below.

4. Empirical performance

In this section, we first describe the data and estimation procedure for our empirical study. We then present the empirical results and model comparison with BY and BKY models. Finally we examine the impact of the two volatility structure in our new model on predictability issues raised by Beeler and Campbell (2012) as well as large and negative variance risk premium as in the data.

4.1. Data

The stock market and macro economic data are from Beeler and Campbell (2012).[8] The stock index returns are monthly. The risk-free rates are the 30 day returns on the Treasury bills. The consumption data include total nominal nondurables and services consumption, which are deflated both by the rate of population growth and by inflation. The population data are the year end values from the Census Bureau. These values are used for annual data and the fourth quarter of quarterly data. Other quarterly population values are interpolated assuming a constant geometric growth rate within a year. Stock returns, dividend growth and consumption are all deflated with the CPI. For yearly inflation, the rate of inflation is the log of the ratio of the CPI in the last month of the current year to the CPI in the last month of the previous year.

We compute the model-free VRP from CBOE's VIX series from 1990 to 2007 and the realized monthly variance based on S&P 500 index. Due to lack of VIX and hence VRP data before 1990, we work out a multiple linear regression of the VRP onto consumption growth, dividend growth, excess stock return, price-dividend ratio, risk-free rate and realized variance of stock, and extrapolate the relationship to period between 1930 and 1990. There are two considerations for choosing this method of extrapolation. First, the extrapolation method preserves the mean value of VRP data. Second, we concern about the covariance matrix estimation of the data for error estimation, this method preserves the covariance between VRP and other variables under available sample. This extrapolation of data is also necessary to obtain a positive-definite covariance matrix estimation.

We also construct the annual volatility of consumption, excess equity return, and dividend growth time series from quarterly data with the method described in Equations (28) and (29). Thus, we obtain annual year end data from 1930 to 2007 including consumption growth, dividend growth, equity return, price-dividend ratio, risk-free rate, realized annual volatility of consumption, dividend and excess return, and VRP.

[8] We are grateful to Jason Beeler for providing us the data.

4.2. GMM tests

We use the GMM to estimate and test the model. The base case pricing restrictions include the mean value, standard deviation and the first order autocorrelation coefficients of the log consumption growth, log dividend growth, real risk-free rate, excess stock return and log stock price dividend ratio. These 15 base case moments, which are presented in the left panel of Table 2, have been extensively studied in the long-run risks model literature, e.g., Bansal and Yaron (2004), Bansal, Kiku and Yaron (2007a,b), and Beeler and Campbell (2012).

Besides the base case moments, we pay special attention to the restrictions on 6 predictability regression coefficients given in Equations (27) and (33). Given the unconditional moments of the base case, these predictability coefficients will be sensitive to the additional state variable of the second volatility factor. Finally, we include two additional restrictions on variance risk premium, the mean and standard deviation. These 8 additional moments are shown in the right panel of Table 2.

In our GMM test, we minimize a weighted sum of squared differences of the 23 target moments between the model derived and the market implied. Instead of using the standard optimal weighting matrix, we use a diagonal one with weights adjusting the moments to the same order of magnitude. This technique is used by Zhou (1994) to isolate parameters for easy estimation. Then the GMM test statistic has to be constructed accordingly, the details of this and the moment conditions are presented in Appendices A.7 and A.8.

4.3. Moment matching and model comparison

In this subsection, we present the GMM estimation and test results. Following BY and BKY model calibration, we fix the preference parameters at $\gamma = 10$ and $\psi = 1.5$ for comparison[9]. Since the two mean levels of the two volatility factors, $\bar{V}_1$ and $\bar{V}_2$, are not identified independently, we set $\bar{V}_1 = \bar{V}_2$. Hence, there are in our two-factor model a total of 19 parameters,

$$\theta = (\mu, \alpha, \varphi_x, \delta_c, \delta_x, \mu_d, \phi, \varphi_d, \delta_d, \sigma_{dc}, \sigma_{dx}, \sigma_{dv}, \sigma_{dv2}, \bar{V}_1 = \bar{V}_2, \sigma_1, \kappa_1, \sigma_2, \kappa_2, \beta).$$

Table 1 provides the estimation results under the heading "New". For comparison, we also provide the corresponding values of the continuous-time version of the BY and BKY

[9]Since the BY and BKY models use different β, the rates of time preference, we choose to optimize β in our model for comparison.

models based on their calibrations.[10]

There are a few notables. First, all the parameters governing the consumption growth, μ, α and φ_x, are virtually the same across the models[11]. Second, the parameters of the dividend growth process are, however, different between their models and ours. In our model, the dividend growth leverage ratio ϕ is 4.37 while the values for BY and BKY are 3 and 2.5. Intuitively, increasing ϕ will increase the A_{1m}, the price-dividend ratio sensitivity to the long-run risk component. This increase is required by the need to match the predictability regression parameters. The dividend volatility leverage ratio for BY is $\varphi_d = 4.5$, for BKY is $\sqrt{\varphi_d^2 + \sigma_{dc}^2} = 6.5$, and for the new model is $\sqrt{\varphi_d^2 + \sigma_{dc}^2 + \sigma_{dx}^2 + \sigma_{dv}^2 + \sigma_{dv2}^2} = 4.95$, which is between BY and BKY. Third, the substantial difference occurs for the volatility process, as expected. While the volatility of the BKY model appears too persistent, with half life around 50 years, and too much so than that of the BY model of 4.5 years. our new model has a long-run component with half life around 20 years, and a short-run one around one month, which are consistent with aforementioned literature on stochastic volatility.

To understand the models further, it will be useful and informative to assess the moments and their sample estimates, as did so by Beeler and Campbell (2012). Table 2 presents the 23 moments that are used in our GMM estimates. We provide the results given by BY, BKY, our new model, and the data (with yearly time interval from 1930 to 2006). The left panel provides the 15 basic asset pricing moments discussed. As with the BY and BKY models, these moments are well matched by our model. In particular, our model produces all the stylized facts on the market, with an equity risk premium 6.25% and a risk-free rate 1.18%, close to the market data of 6.20% and 0.99%, *resp.* Further, Table 3 presents the comparison between our new model and BY model on the conditional volatility of pricing kernel (the maximal Sharpe ratio) and the relative contributions of different economic shocks. The result of the BY model is from Bansal and Yaron (2004). The maximal Sharpe ratio in our model is 0.60 comparing with 0.73 of BY model, the difference of which is mainly due to the modelling differences, such as the continuous-time vs discrete-time model, and the square-root process vs Gaussian process for variance. Further, the contribution from the fluctuating economic

[10]Since the BY and BKY parameters are not calibrated but obtained from BY and BKY papers, we postpone to present the standard errors of our GMM estimates of the new model in Table 4.

[11]Note that the long-run risk volatility multiple φ_x is 0.044 and 0.038 for BY and BKY models, we have multiplied by 12 because in our model X_t is annualized expected consumption growth while in their discrete time model, it is monthly growth rate.

uncertainty in our new model comes from two volatility factors, in contrast to BY's one volatility factor. Despite the differences, the relative contributions from the different risks to pricing kernel are roughly the same. For example, the contribution of the long-run risks is 49% in BY model, and 47% in the new model. In addition, the combined contribution from the two volatility factors in the new model is 36%, in comparison to the 39% of the one volatility factor in BY model. Similar to BY model, the independent consumption shocks contribute only 15% to the total variance of the pricing kernel, and the maximal Sharpe ratio with only i.i.d. consumption growth rate is only 0.23. In conclusion, our model matches well the basic asset pricing moments studied in BY model, and the main difference between the models comes from the different structure of volatility factors.

The major differences among the models occur in the right panel of Table 2, where $\beta(r_e)$, $\beta(\Delta c)$ and $\beta(\Delta d)$ are the regression coefficients of the market excess return, consumption and dividend growth on the price-dividend ratio in a one-year horizon. For the market excess return, the data implies a $\beta(r_e) = -0.059$. While our model matches well with a beta of $\beta(r_e) = -0.073$, the BY model, with a beta of $\beta(r_e) = -0.007$, does not match this important predictability. In contrast, the BKY improves over the BY model substantially in this aspect, matching well with a beta of $\beta(r_e) = -0.078$. Overall, the BKY model performs very well in explaining the predictability of excess returns, consumption and dividends, so does our new model.

However, the good performance of the BKY model comes at a high cost of matching the predictive regression coefficients of the volatility regressions of excess return, consumption and dividend growth on price-dividend ratio, $\beta_{vol}(r_e)$, $\beta_{vol}(\Delta c)$ and $\beta_{vol}(\Delta d)$. For $\beta_{vol}(r_e)$, the BKY value is -1.315, which is more than 10 times greater in magnitude than that of the data, -0.081. For both $\beta_{vol}(\Delta c)$ and $\beta_{vol}(\Delta d)$, the values of the BKY model are about 3 times greater in magnitude than those of the data too. On the other hand, the BY model provides the volatility predictability 10 times smaller than that of BKY model, and fails to match the data as well. In contrast, our new model matches well with the data for all the three volatility regression coefficients. As explained later in Section 4.5, the key reason is that both BY and BKY models rely on a one-factor volatility model which makes the matching difficult for $\beta_{vol}(r_e)$, $\beta_{vol}(\Delta c)$ and $\beta_{vol}(\Delta d)$.

Finally, we examine the last two target moments in Table 2, i.e., the mean and standard

deviation of VRP. While the data[12] show a large negative VRP at -8.89 with standard deviation of 26.98, both the BY and the BKY models imply a VRP of only -0.005 and -0.010, which are too small in magnitude to explain the observed variance risk premium. In contrast, our new model implies a VRP of -5.57, which is very close to the empirical mean level. Moreover, the standard deviation of the VRP in our model is 29.71, which also matches well with the data.

The GMM test and comparison results are presented in the Table 4. The upper panel shows the parameters with estimation errors and the model test for our new two-factor model. At the bottom of this panel, we present the J test of the model, which is the standardized sum of squared moment matching errors. As derived in Appendix A.7, J should be χ^2-distributed with degree of 4. Our model has $J = 7.53$ corresponding to a p-value of $p = 0.11$, which means we cannot reject the model at 10% confidence level.

Finally, we compare our model with the BY and BKY models, both of which can be considered as special or "restricted" cases of our model. Specifically, the BKY model can be obtained through restrictions, i.e.,

$$\bar{V}_2 = \sigma_2 = \kappa_2 = \sigma_{dv2} = 0,$$

and

$$\delta_c = \delta_x = \delta_d = 1.$$

BY model can be obtained by additional restrictions of

$$\sigma_{dc} = \sigma_{dx} = \sigma_{dv} = 0.$$

We use the same covariance matrix for the matching moments as the unrestricted two-factor model, and compute the $J(\text{BKY}) - \text{J}(\text{New}) \sim \chi^2(7)$ and $J_T(\text{BY}) - \text{J}_\text{T}(\text{New}) \sim \chi^2(10)$ using the moments given by BKY and BY. The results are reported in the lower panel of Table 4. It shows that $J_T(\text{BKY}) - \text{J}_\text{T}(\text{New}) = 376$ and $J_T(\text{BY}) - \text{J}_\text{T}(\text{New}) = 293$, a very large increase in the errors for matching moments, corresponding to p-value very close to zero. This means that our new model outperforms the BY and the BKY models substantially from a statistical point of view as well.

[12] According to market convention, the VRP data presented here is converted to monthly VRP through a factor of $\frac{1}{12}$, and multiplied by 10000.

4.4. Return variance decomposition

To obtain further intuition of the model, we examine its implication on return variance decomposition, which was first proposed by Campbell and Shiller (1988) and inspired a large volume of literature across many disciplines including finance, accounting, and macroeconomics (see Chen and Zhao, 2009 for a review). In particular, using aggregate stock market data from 1927 to 1988, Campbell (1991) shows that about 1/3 of the variance of unexpected returns is attributed to the cash flow risk, 1/3 to the discount rate risk, and 1/3 to the covariance between these two. Debate on which one is more important in driving stock market return variation, discount rate risk or cash flow risk, is highly inconclusive due to model uncertainty in return predictability. Using data from 1953 to 2001 with different predictive variables, Chen and Zhao (2009) find that discount-rate risk accounts for anywhere from 10% to more than 80% of return variance.

The long-run risks model, successful in explaining the level and volatility of stock market return, attributes almost all unexpected variation of the stock return to changing cash flow by emphasizing on cash flow risk. In contrast, the habit formation model by Campbell and Cochrane (1999) emphasize on discount rate risk by changing investors' risk attitude. We show that by extending the one volatility BY model to two volatility model, we provide enough flexibility for the model to generate more variation due to discount rate risk.

We apply the variance decomposition of Campbell (1991) with our model estimation. Table 5 demonstrates the results of our new model in comparison with BKY model[13]. Note that in our estimation, BKY attributes 99.5%, 8.2% and −7.6% of total variation to cash flow risk, discount rate risk, and covariance between these two, while Avramov and Cederburg's (2012) estimation is 106%, 4.4% and −11.2%, *resp.* The difference is due to the fact that we cast BY/BKY model in continuous-time with square-root volatility process in contrast to their discrete-time model with Gaussian volatility process. Both cases show that the BKY model attributes too small a portion of stock return variance to discount rate than aforementioned studies. On the other hand, as shown in the lower panel of Table 5, our new two-factor model attributes 23.3% to discount rate news, which is in line with empirical results of Campbell (1991). The ability of our two volatility model to generate more discount rate variation is attributed to the additional short-run volatility factor, which has much

[13]Here we only compare with BKY calibration because BKY matches predictability better than BY calibration.

higher volatility of volatility (vol of vol) and hence higher volatility of the discount rate. It turns out that this short-run volatility factor is also very important to explain volatility predictability and variance risk premium as demonstrated next.

4.5. Volatility predictability

In this subsection, we examine further the statistical properties of the volatility regression coefficients. Table 2 shows that for one-year horizon, the regression slope coefficients β_{vol} for volatilities of excess return, consumption and dividends are $-0.123, -0.128$, and -0.146 in BY model, and $-1.315, -1.420$, and -1.483 in BKY model, both of which show little difference across the three regressions. The results are consistent with our theoretical analysis. Equation (32) shows that the volatility regression coefficients should be the same for all three variables based on a one-factor model. Indeed, the three regressors in regression equation (30) can all be written as

$$\ln \int_0^\tau \sqrt{V_t}|dZ_t| \approx Const + \frac{1}{2\tau}\int_0^\tau \frac{V_t}{\bar{V}}dt, \tag{39}$$

where V_t corresponds generically to volatilities of excess return, consumption and dividend growth. Note that in one-factor volatility model, the three volatilities will be the same up to a scaling factor to the single volatility component, hence the regressors are all the same because $\frac{V_t}{\bar{V}}$ will be invariant to a scaling factor. Consequently β_{vol}'s will be the same for the three variables, and this holds for different horizons. Hence, it is critical to use a two-factor volatility model to explain the varying degrees of volatility predictability.

Note that the volatility predictability in the BKY model is ten-fold larger than that in the BY model, due to a much more persistent volatility component in BKY. In our proposed two-factor model, the volatility regression slope coefficients match remarkably well with those from the data. While the data implies significantly varying beta values of $-0.081, -0.481$ and -0.530 for excess return, consumption growth and dividend growth, so does the model with $-0.167, -0.317$ and -0.546.

In short, the additional volatility factor is instrumental to explain the differences in volatility predictability across excess returns, consumption and dividends. As it turns out, it is also fundamentally important in explaining the market variance premium, as discussed further below.

4.6. Variance risk premium

As mentioned above, one of the difficulties of BY/BKY model is to explain the market variance risk premium (VRP) which is large and negative. In this subsection we examine in detail why the BY/BKY model fail to match the market VRP. Consider the following BY/BKY one volatility factor process which is cast in our continuous setting:

$$dV_t = \kappa(\bar{V} - V_t)dt + \sigma\sqrt{V_t}dw_t$$

The model implied VRP is $-\nu V_t$, and the coefficient ν, similar to Equation (A38) but with only one volatility factor, is given by

$$\nu = \frac{1-\gamma\psi}{1-\gamma}A_2\sigma^2, \tag{40}$$

which is proportional to σ^2. For persistent process with κ very small, given the unconditional vol of vol, $\frac{\sigma^2\bar{V}}{2\kappa}$, which is inversely proportional to κ, the parameter σ^2 and hence the VRP coefficient in Equation (40) cannot be large enough to match the market data. Another way to look at it is, similar to Equation (A41), that the VRP coefficient $\nu < \kappa$. In BKY calibration, with a value of $\kappa = 0.012$ and an annual return volatility of 20%, the VRP in absolute value will be bounded by

$$|\text{VRP}_{\text{BKY}}| < 0.012 \times 20\%^2/12 \times 10000 = 0.4,$$

which is much smaller than a value of 8.89 from our data,[14] as reported in Table 6. With an additional volatility component that is much less persistent than both BY and BKY, our new model can produce the desired VRP mostly contributed by this short-run volatility, which is consistent with evidence from volatility derivatives market (e.g., Egloff, Leippold, and Wu, 2009, Lu and Zhu 2010). In addition, our new model matches the autoregression coefficient AR1 (with one month lag) of the VRP well. While the data implies an AR1 value of 0.54, our model yields a value of 0.47. This is also consistent with that of Bollerslev, Tauchen and Zhou (2009). In contrast, both BY and BKY models imply a too large value of 0.99 for AR1, a result of their single factor structure of the volatility process that is highly persistent.

As mentioned earlier, a jump-diffusion model, Drechsler and Yaron (2011), is an alternative to explain the negative and large market VRP in the long-run risks model. Table 6,

[14]Note that the VRP estimated using Drechsler and Yaron (2011) data is −12.67. The difference between their data and ours is due to the difference of sampling period.

citing their result, shows that their model indeed explains well the magnitude of VRP. Similar models are also proposed by Bansal, Gallant and Tauchen (2007), Eraker (2008), and Eraker and Shaliastovich (2008), Bollerslev, Sizova, and Tauchen (2012), among others. However, as evident from our earlier analysis, it is very difficult for the Drechsler and Yaron (2011) model to explain the volatility predictability cross-sectionally because it has only one state variable. In addition, studies in the volatility literature (see, e.g., Christoffersen, Jacobs, Ornthanalai and Wang, 2008, and Lu and Zhu, 2010) show that the two-factor volatility model is generally preferred in explaining the large negative VRP as well as variance term structure. Hence, our extension of the BY and BKY models by adding another volatility factor seems to offer a promising route for future applications and further extension of the long-run risks models.

5. Conclusion

One of the fundamental problems in asset pricing is to explain various stylized facts about the equity market, for which the rational and general equilibrium long-run risks model of Bansal and Yaron (2004) seems to have come a long way toward this goal. Hence, it is not surprising that there are subsequently a number of important studies along this line of research, including Bansal, Khatchatrian, and Yaron (2005), Bansal, Kiku, and Yaron (2007a, b), Chen, Collin-Dufresne, and Goldstein (2009), Constantinides and Ghosh (2011), Drechsler and Yaron (2011), Eraker (2008), Hansen, Heaton, and Li (2008), Pakos (2008), Avramov and Hore (2009), Bansal, Dittmar, and Kiku (2009), Beeler and Campbell (2012) and Ferson, Nallareddy and Xie (2011). However, in the equity market, there are still three major problems confronting the model. It has difficulty in explaining the predictability of consumption, and the varying degrees of volatility predictability of excess stock returns, consumption and dividends. Moreover, the model fails completely in explaining the large negative market variance risk premium. While Bansal, Kiku, and Yaron (2007b) and Drechsler and Yaron (2011) address one of the problems each, there have been no studies that can resolve all the three problems simultaneously. In addition, the one factor stock volatility process does not match the findings in the volatility literature and it cannot explain the importance of discount rate risk relative to cash flow news.

This paper proposes a simple extension of the Bansal and Yaron (2004) and Bansal,

Kiku, and Yaron (2007b) models that solves all aforementioned problems simultaneously, while retaining many of the desired properties of the original models in explaining well other stylized facts. In our new model, there are two components of volatility: the long- and short-run volatilities, both of which enter into the dynamics of both the consumption and dividend growth rates. As a result, the equilibrium stock volatility process has two volatility factor to be consistent with findings of the volatility literature. Hence, it may not be surprising that the new model can justify the large negative variance risk premium. The flexibility permitted by the two-factor volatility model also allows for more realistic modeling of the discount rate risk in driving unexpected stock return and for explaining the aforementioned challenging issues facing the BY and BKY models. Looking forward, due to wide applications of the long-run risks model of Bansal and Yaron (2004), such as in the cross-section of the equity market and the currency market, it will be of interest to see how conclusions of these applications might be altered in light of the proposed new model. Additionally, as pointed out by Beeler and Campbell (2012), it is an open challenge to extend the long-run risks models to the fixed income markets. All of these issues are important and exciting topics of future research.

Appendix A

A.1 Derivation for the A_i's, risk-free rate and market price of risk

First, we re-write the normalized aggregator f defined in Equation (5) as

$$f(C,J) = \frac{\beta}{1-\frac{1}{\psi}}(1-\gamma)J[G-1],$$

where

$$G \equiv (\frac{C}{((1-\gamma)J)^{\frac{1}{1-\gamma}}})^{1-\frac{1}{\psi}}. \tag{A1}$$

Then, taking partial derivatives of $f(C,J)$ with respect to J and C , we have

$$f_J = (\theta-1)\beta G - \beta\theta \tag{A2}$$

and

$$f_C = \beta\frac{G}{C}(1-\gamma)J. \tag{A3}$$

Conjecturing a solution for J of the following form,

$$J(W_t, X_t, V_{1t}, V_{2t}) = \exp(A_0 + A_1X_t + A_2V_{1t} + A_3V_{2t})\frac{W_t^{1-\gamma}}{1-\gamma}, \tag{A4}$$

and using the standard envelope condition $f_C = J_W$, we have

$$C = J_W^{-\psi}[(1-\gamma)J]^{\frac{1-\gamma\psi}{1-\gamma}}\beta^{\psi}. \tag{A5}$$

Substituting (A3) and (A4) into (A5), we obtain

$$\frac{C}{W} = \beta^{\psi}\exp\left[(A_0 + A_1X_t + A_2V_{1t} + A_3V_{2t})\frac{1-\psi}{1-\gamma}\right]. \tag{A6}$$

and hence J can be re-written as

$$J(C_t, X_t, V_{1t}, V_{2t}) = \beta^{-\psi(1-\gamma)}\exp[\psi(A_0 + A_1X_t + A_2V_{1t} + A_3V_{2t})]\frac{C_t^{1-\gamma}}{1-\gamma}, \tag{A7}$$

Further substituting (A6) and (A4) into (A1), we get

$$\beta G = \frac{C_t}{W_t}.$$

Applying the log-linear approximation, we obtain

$$\beta G = \frac{C_t}{W_t} \approx g_1 - g_1\log g_1 + g_1\log(\beta G). \tag{A8}$$

This implies that

$$f = \theta J(\beta G - \beta) \approx \theta J \left[g_1 \frac{1-\psi}{1-\gamma}(A_0 + A_1 X_t + A_2 V_{1t} + A_3 V_{2t}) + \xi \right], \tag{A9}$$

where $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$ and

$$\xi = g_1 - g_1 \log g_1 + g_1 \psi \log \beta - \beta. \tag{A10}$$

Substituting (A9) into the HJB Equation (6),

$$\begin{aligned} & f(C,J) + c \cdot (\mu + X) J_c + \frac{1}{2}[\delta_c V_1 + (1-\delta_c) V_2] C^2 J_{CC} + J_X \cdot (-\alpha X) + \frac{1}{2}\varphi_x^2 [\delta_x V_1 + (1-\delta_x) V_2] J_{XX} \\ & + J_{V_1} \cdot \kappa_1 (\bar{V}_1 - V_1) + \frac{1}{2}\sigma_1^2 v_1 J_{V_1 V_1} + J_{V_2} \cdot \kappa_2 (\bar{V}_2 - V_2) + \frac{1}{2}\sigma_2^2 V_2 J_{V_2 V_2} = 0, \end{aligned} \tag{A11}$$

where $\{C_t\}$ is the optimal consumption process, and we have used the definition of

$$\mathcal{A}^c J \equiv \sum_i b(z) \frac{\partial J(z)}{\partial z} + \sum_{i,j} (\sigma \sigma^T)_{i,j}(z) \frac{\partial^2 J}{\partial z_i z_j},$$

with $z = (C, X, V_1, V_2)$ and $b(z)$ and $\sigma(z)$ the drift and diffusive terms for z defined in Equation (2). Collecting the terms containing constant, X_t, V_{1t} and V_{2t}, *resp*, we have

$$\begin{aligned} & \theta g_1 \frac{1-\psi}{1-\gamma} A_0 + \theta \xi + (1-\gamma)\mu + \kappa_1 \bar{V}_1 \psi A_2 + \kappa_2 \bar{V}_2 \psi A_3 = 0 \\ X: \quad & \theta g_1 \frac{1-\psi}{1-\gamma} A_1 + (1-\gamma) - \alpha \psi A_1 = 0 \\ V_1: \quad & \theta g_1 \frac{1-\psi}{1-\gamma} A_2 - \frac{1}{2}\gamma(1-\gamma)\delta_c + \frac{1}{2}\varphi_x^2 \delta_x \psi^2 A_1^2 - \kappa_1 \psi A_2 + \frac{1}{2}\sigma_1^2 \psi^2 A_2^2 = 0 \\ V_2: \quad & \theta g_1 \frac{1-\psi}{1-\gamma} A_3 - \frac{1}{2}\gamma(1-\gamma)(1-\delta_c) + \frac{1}{2}\varphi_x^2 (1-\delta_x) \psi^2 A_1^2 - \kappa_2 \psi A_3 + \frac{1}{2}\sigma_2^2 \psi^2 A_3^2 = 0. \end{aligned}$$

Solving the above algebraic equations, we obtain

$$\begin{aligned} A_0 &= \frac{1}{g_1 \psi}[\theta \xi + (1-\gamma)\mu + \kappa_1 \bar{V}_1 \psi A_2 + \kappa_2 \bar{V}_2 \psi A_3], \\ A_1 &= \frac{1-\gamma}{(g_1+\alpha)\psi}, \\ A_2 &= \frac{-b_1 - \sqrt{b_1^2 - 4a_1 c_1}}{2a_1}, \\ A_3 &= \frac{-b_2 - \sqrt{b_2^2 - 4a_2 c_2}}{2a_2}, \end{aligned} \tag{A12}$$

with

$$\begin{aligned} a_1 = \tfrac{1}{2}\sigma_1^2 \psi^2, \quad b_1 = -(g_1 + \kappa_1)\psi, \quad c_1 = -\tfrac{1}{2}\gamma(1-\gamma)\delta_c + \tfrac{1}{2}\varphi_x^2 \delta_x \tfrac{(1-\gamma)^2}{(g_1+\alpha)^2}, \\ a_2 = \tfrac{1}{2}\sigma_2^2 \psi^2, \quad b_2 = -(g_1 + \kappa_2)\psi, \quad c_2 = -\tfrac{1}{2}\gamma(1-\gamma)\delta_c(1-\delta_c) + \tfrac{1}{2}\varphi_x^2 (1-\delta_x) \tfrac{(1-\gamma)^2}{(g_1+\alpha)^2}. \end{aligned}$$

We then derive the risk-free rate and market prices of risks. Recall that the pricing kernel is given by Equation (14). Based on the definition for f in Equation (5), we have

$$\begin{aligned} f_J &= \xi_1 - g_1(A_1X_t + A_2V_{1t} + A_3V_{2t})\frac{1-\gamma\psi}{1-\gamma}, \\ f_C &= \beta^{\psi\gamma}\exp\left[(B + A_1X_t + A_2V_{1t} + A_3V_{2t})\frac{1-\gamma\psi}{1-\gamma}\right]C_t^{-\gamma}, \end{aligned}$$

where

$$\xi_1 = (\theta-1)\xi - \beta - g_1\frac{1-\gamma\psi}{1-\gamma}A_0. \tag{A13}$$

Applying Ito's Lemma to π_t in Equation (14), we have

$$\frac{d\pi_t}{\pi_t} = -(r_f dt + \lambda_1 dZ_{1t} + \lambda_2 dZ_{2t} + \lambda_3 dw_{1t} + \lambda_4 dw_{2t}), \tag{A14}$$

where the risk-free rate r_f and the market prices of risks, $\lambda_i, i = 1,2,3,4$, are given below.

First, the risk-free rate is

$$r_f = r_0 + r_1X_t + r_2V_{1t} + r_3V_{2t}, \tag{A15}$$

where

$$\begin{aligned} r_0 &= -(\xi_1 + (\kappa_1A_2\bar{V}_1 + \kappa_2A_3\bar{V}_2)\frac{1-\gamma\psi}{1-\gamma} - \gamma\mu), \\ r_1 &= \frac{1}{\psi}, \\ r_2 &= (g_1+\kappa_1)A_2\frac{1-\gamma\psi}{1-\gamma} - \frac{1}{2}\left(\frac{1-\gamma\psi}{1-\gamma}\right)^2(A_1^2\varphi_x^2\delta_x + A_2^2\sigma_1^2) - \frac{1}{2}\gamma(\gamma+1)\delta_c, \\ r_3 &= (g_1+\kappa_2)A_3\frac{1-\gamma\psi}{1-\gamma} - \frac{1}{2}\left(\frac{1-\gamma\psi}{1-\gamma}\right)^2[A_1^2\varphi_x^2(1-\delta_x) + A_3^2\sigma_2^2] - \frac{1}{2}\gamma(\gamma+1)(1-\delta_c). \end{aligned} \tag{A16}$$

Second, the market prices of risks are

$$\begin{aligned} \lambda_1 &= \gamma\sqrt{V_{1t}\delta_c + V_{2t}(1-\delta_c)}, \\ \lambda_2 &= -\frac{1-\gamma\psi}{1-\gamma}A_1\varphi_x\sqrt{V_{1t}\delta_x + V_{2t}(1-\delta_x)}, \\ \lambda_3 &= -\frac{1-\gamma\psi}{1-\gamma}A_2\sigma_1\sqrt{V_{1t}}, \\ \lambda_4 &= -\frac{1-\gamma\psi}{1-\gamma}A_3\sigma_2\sqrt{V_{2t}}. \end{aligned} \tag{A17}$$

Q.E.D.

A.2 Derivation for the A_{im}'s

Let

$$\frac{D_t}{P_t} = \exp\{(A_{0m} + A_{1m}X_t + A_{2m}V_{1t} + A_{3m}V_{2t})\}. \tag{A18}$$

A key step in the derivation is to use the following pricing relation given in Equation (13). With similar loglinear approximation as Equation (A8), we can approximate the ratio as

$$\frac{D_t}{P_t} \approx g_{0m} + g_{1m}\log\frac{D_t}{P_t} = g_{0m} + g_{1m}((A_{0m} + A_{1m}X_t + A_{2m}V_{1t} + A_{3m}V_{2t})), \tag{A19}$$

where

$$g_{0m} = g_{1m} - g_{1m}\log g_{1m}.$$

Applying Ito's lemma to (A18), we have

$$\frac{dP_t}{P_t} = \frac{dD_t}{D_t} - (A_{1m}dX_t + A_{2m}dV_{1t} + A_{3m}dV_{2t}) + \frac{1}{2}A_{1m}^2(dX_t)^2 + \frac{1}{2}A_{2m}^2(dV_{1t})^2 + \frac{1}{2}A_{3m}^2(dV_{2t})^2.$$

Hence,

$$\begin{aligned} E_t(\frac{dP_t}{P_t})/dt &= \mu_d + \phi X_t + \alpha A_{1m}X_t - \kappa_1 A_{2m}(\bar{V}_1 - V_{1t}) - \kappa_2 A_{3m}(\bar{V}_2 - V_{2t}) \\ &\quad + \frac{1}{2}A_{1m}^2\varphi_x^2[V_{1t}\delta_x + V_{2t}(1-\delta_x)] + \frac{1}{2}A_{2m}^2\sigma_1^2 V_{1t} + \frac{1}{2}A_{3m}^2\sigma_2^2 V_{2t}. \end{aligned} \tag{A20}$$

The risk premium term in Equation (13) can thus be written as

$$\begin{aligned} -E_t\left[\frac{d\pi_t}{\pi_t}\frac{dP_t}{P_t}\right]/dt &= \sigma_{dc}\lambda_1\sqrt{V_{1t}\delta_c + V_{2t}(1-\delta_c)} - (A_{1m}\varphi_x - \sigma_{dx})\lambda_2\sqrt{V_{1t}\delta_x + V_{2t}(1-\delta_x)} \\ &\quad -(A_{2m}\sigma_1 - \sigma_{dv})\lambda_3\sqrt{V_{1t}} - (A_{3m}\sigma_2 - \sigma_{dv2})\lambda_4\sqrt{V_{2t}}, \end{aligned} \tag{A21}$$

where λ_1, λ_2, λ_3 and λ_4 are market prices of risks as defined in Equation (A17), and $d\pi_t/\pi_t$ is the pricing kernel as defined in Equation(16).

Now, substituting (A19), (A20), (A21), and risk-free rate as defined in (17) into Equation (13), and collecting terms containing X_t, we obtain

$$A_{1m} = -\frac{\phi - \frac{1}{\psi}}{g_{1m} + \alpha}. \tag{A22}$$

Collecting terms containing V_{1t} and V_{2t}, *resp*, we obtain an equation for A_{2m},

$$a_{2m}A_{2m}^2 + b_{2m}A_{2m} + c_{2m} = 0$$

with

$$a_{2m} = \frac{1}{2}\sigma_1^2, \quad b_{2m} = g_{1m} + \kappa_1 - \frac{1-\gamma\psi}{1-\gamma}A_2\sigma_1^2, \quad c_{2m} = (\frac{1}{2}A_{1m}^2 - \frac{1-\gamma\psi}{1-\gamma}A_1A_{1m})\varphi_x^2\delta_x + r_2.$$

Solving it, we have

$$A_{2m} = \frac{-b_{2m} \pm \sqrt{b_{2m}^2 - 4a_{2m}c_{2m}}}{2a_{2m}}. \tag{A23}$$

We choose the root that goes to zero when σ_1 goes to zero. This is because when σ_1, or a_{2m} goes to zero, the price sensitivity to V_1 should be zero.

Similarly, we obtain an equation for A_{3m},

$$a_{3m}A_{3m}^2 + b_{3m}A_{3m} + c_{3m} = 0$$

with

$$a_{3m} = \frac{1}{2}\sigma_2^2,\ b_{2m} = g_{1m} + \kappa_2 - \frac{1-\gamma\psi}{1-\gamma}A_3\sigma_2^2,\ c_{3m} = (\frac{1}{2}A_{1m}^2 - \frac{1-\gamma\psi}{1-\gamma}A_1A_{1m})\varphi_x^2(1-\delta_x) + r_3.$$

The solution is

$$A_{3m} = \frac{-b_{3m} \pm \sqrt{b_{3m}^2 - 4a_{3m}c_{3m}}}{2a_{3m}}, \tag{A24}$$

where we choose the root in a similar fashion as for A_{2m} above.

Finally, collecting the constant terms in Equation (13), we obtain

$$\mu_d - \kappa_1 A_{2m}\bar{V}_1 - \kappa_2 A_{3m}\bar{V}_2 + g_{0m} + g_{1m}A_{0m} + r_0 = 0,$$

and re-arrange terms to get

$$A_{0m} = -\frac{1}{g_{1m}}\left[\mu_d - \kappa_1 A_{2m}\bar{V}_1 - \kappa_2 A_{3m}\bar{V}_2 + g_{1m} - g_{1m}\log g_{1m} + r_0\right].$$

So far, we obtain all the A_{im} coefficients.

To obtain the market return volatility, we apply Ito's Lemma to Equation (20) and obtain

$$\begin{aligned}
\frac{dP_t}{P_t} &= [\mu_d - (A_{2m}\kappa_1\bar{V}_1 + A_{3m}\kappa_2\bar{V}_2) + (\phi + \alpha A_{1m})X_t \\
&\quad +(\frac{1}{2}A_{1m}^2\varphi_x^2\delta_x + \frac{1}{2}A_{2m}\sigma_1^2 + A_{2m}\kappa_1 - A_{1m}\sigma_{dx}\varphi_x\delta_x - A_{2m}\sigma_1\sigma_{dv})V_{1t} \\
&\quad +(\frac{1}{2}A_{1m}^2\varphi_x^2(1-\delta_x) + \frac{1}{2}A_{3m}\sigma_2^2 + A_{3m}\kappa_2 - A_{1m}\sigma_{dx}\varphi_x(1-\delta_x) - A_{3m}\sigma_2\sigma_{dv2})V_{2t}]dt \\
&\quad +\varphi_d\sqrt{V_{1t}\delta_d + V_{2t}(1-\delta_d)}dB_t + \sigma_{dc}\sqrt{V_{1t}\delta_c + V_{2t}(1-\delta_c)}dZ_{1t} \\
&\quad +(\sigma_{dx} - A_{1m}\varphi_x)\sqrt{V_{1t}\delta_x + V_{2t}(1-\delta_x)}dZ_{2t} \\
&\quad +(\sigma_{dv} - A_{2m}\sigma_1)\sqrt{V_{1t}}dw_{1t} + (\sigma_{dv2} - A_{3m}\sigma_2)\sqrt{V_{2t}}dw_{2t} \\
&= [c_3 + c_4X_t + c_5V_{1t} + c_6V_{2t}]dt + \sqrt{c_1V_{1t} + c_2V_{2t}}dZ_t,
\end{aligned}$$

where c_i ($i = 1$ to 6) are constants, dZ_t is a new Brownian motion defined accordingly, and hence the variance of the price process is

$$V_t = c_1 V_{1t} + c_2 V_{2t},$$

which is the same given in Equation (22) with

$$\begin{aligned} c_1 &= \varphi_d^2 \delta_d + \sigma_{dc}^2 \delta_c + (\sigma_{dx} - A_{1m}\varphi_x)^2 \delta_x + (\sigma_{dv} - A_{2m}\sigma_1)^2, \\ c_2 &= \varphi_d^2 (1-\delta_d) + \sigma_{dc}^2 (1-\delta_c) + (\sigma_{dx} - A_{1m}\varphi_x)^2 (1-\delta_x) + (\sigma_{dv2} - A_{3m}\sigma_2)^2, \end{aligned} \tag{A25}$$

and the parameters for the drift term are

$$\begin{aligned} c_4 &= \phi + \alpha A_{1m}, \\ c_5 &= (\frac{1}{2} A_{1m}^2 \varphi_x^2 \delta_x + \frac{1}{2} A_{2m}\sigma_1^2 + A_{2m}\kappa_1 - A_{1m}\sigma_{dx}\varphi_x\delta_x - A_{2m}\sigma_1\sigma_{dv}), \\ c_6 &= (\frac{1}{2} A_{1m}^2 \varphi_x^2 (1-\delta_x) + \frac{1}{2} A_{3m}\sigma_2^2 + A_{3m}\kappa_2 - A_{1m}\sigma_{dx}\varphi_x(1-\delta_x) - A_{3m}\sigma_2\sigma_{dv2}). \end{aligned} \tag{A26}$$

Q.E.D.

A.3 Solutions to g_1 and g_{1m}

Note that the derived solutions depend on the approximation constant g_1, which can be solved endogenously. Given the model parameters, we can compute the unconditional mean of consumption-wealth ratio as a function of the parameters,

$$\begin{aligned} g_1 &= E\left(\frac{C}{W}\right) = \beta^{\psi} \exp\{A_{0a}\} \exp\left\{\frac{1}{4} A_{1a}^2 \varphi_x^2 \frac{(\bar{V}_1\delta_x + \bar{V}_2(1-\delta_x))}{\alpha}\right\} \\ &\quad \cdot \exp\left\{-\frac{2\kappa_1 \bar{V}_1}{\sigma_1^2} \log(1 - \frac{A_{2a}}{2\kappa_1/\sigma_1^2})\right\} \cdot \exp\left\{-\frac{2\kappa_2 \bar{V}_2}{\sigma_2^2} \log(1 - \frac{A_{3a}}{2\kappa_2/\sigma_2^2})\right\}. \end{aligned} \tag{A27}$$

Note that the A_{ia}'s on the right hand side are also functions of g_1. Substituting A_{ia} as function of g_1 into Equation (A27), we obtain a nonlinear function in terms of g_1 only, and hence g_1 can be solved in terms of the fundamental parameters of the model, and can be computed numerically with many available algorithms.

Similarly, we can solve g_{1m} endogenously based on dividend-price ratio given as

$$\begin{aligned} g_{1m} &= E\left(\frac{D}{P}\right) = \exp\{A_{0m}\} \exp\left\{\frac{1}{4} A_{1m}^2 \varphi_x^2 \frac{(\bar{V}_1\delta_x + \bar{V}_2(1-\delta_x))}{\alpha}\right\} \\ &\quad \cdot \exp\left\{-\frac{2\kappa_1 \bar{V}_1}{\sigma_1^2} \log(1 - \frac{A_{2m}}{2\kappa_1/\sigma_1^2})\right\} \cdot \exp\left\{-\frac{2\kappa_2 \bar{V}_2}{\sigma_2^2} \log(1 - \frac{A_{3m}}{2\kappa_2/\sigma_2^2})\right\}. \end{aligned} \tag{A28}$$

This can be solved numerically as above. Q.E.D.

A.4 Predictability of variables

The regressors of the three regressions given in Equation (24) all have the generic functional form of

$$dY_t = [a_0 + a_1 X_t + a_2 V_{1t} + a_3 V_{2t}]dt + \sqrt{b_1 V_{1t} + b_2 V_{2t}} dZ_t,$$

given in Equation (26) where dY_t corresponds to excess return $d\ln P_t + \frac{D_t}{P_t} - r_f dt$, consumption growth $d\ln C_t$ and dividend growth $d\ln D_t$, respectively. For stock market excess return, we have

$$\begin{aligned} a_1 &= c_4 + r_1 + g_{1m} A_{1m}, \\ a_2 &= c_5 - \frac{c_1}{2} + r_2 + g_{2m} A_{2m}, \\ a_3 &= c_6 - \frac{c_2}{2} + r_3 + g_{3m} A_{3m}, \end{aligned} \tag{A29}$$

where c_1, c_2, c_4, c_5 and c_6 are defined in Equations (A25) and (A26).

For consumption growth, we have

$$a_1 = 1,\ a_2 = -\frac{\delta_c}{2},\ a_3 = -\frac{1-\delta_c}{2}. \tag{A30}$$

For dividend growth, we have

$$a_1 = \varphi,\ a_2 = -\frac{\varphi_d^2 \delta_d + \sigma_{dc}^2 \delta_c + \sigma_{dv}^2 + \sigma_{dx}^2 \delta_x}{2}, a_3 = -\frac{\varphi_d^2 (1-\delta_d) + \sigma_{dc}^2 (1-\delta_c) + \sigma_{dv2}^2 + \sigma_{dx}^2 (1-\delta_x)}{2}. \tag{A31}$$

We want to show Equations (27). Given Equations (26) and (20), and denoting $\mathrm{Cov}(x,y) \equiv < x, y >$, and $pd_t \equiv p_t - d_t$, we have

$$\begin{aligned} & < \int_t^{t+\tau} dy_s, p_t - d_t > \\ = & \int_t^{t+\tau} ds < a_0 + a_1 X_s + a_2 V_{1s} + a_3 V_{2s}, pd_t > \\ = & \int_t^{t+\tau} ds [a_1 < x_s, pd_t > + a_2 < V_{1s}, pd_t > + a_3 < V_{2s}, pd_t >] \\ = & -\int_0^{\tau} ds \big[a_1 A_{1m} \frac{\sigma_x^2}{2\alpha} e^{-\alpha s} + a_2 A_{2m} \frac{\sigma_1^2 \bar{V}_1}{2\kappa_1} e^{-\kappa_1 s} + a_3 A_{3m} \frac{\sigma_2^2 \bar{V}_2}{2\kappa_2} e^{-\kappa_2 s} \big] \end{aligned}$$

where

$$\sigma_x^2 = \varphi_x^2[\bar{V}_1\delta_x + \bar{V}_2(1-\delta_x)]$$

Integrating the above equation, we obtain Equation (27), where

$$\begin{aligned}
\mathrm{Cov}(\Delta_\tau y, p-d) &= -\Big[a_1 A_{1m}\frac{\sigma_x^2}{2\alpha^2}(1-e^{-\alpha\tau}) + a_2 A_{2m}\frac{\sigma_1^2\bar{V}_1}{2\kappa_1^2}(1-e^{-\kappa_1\tau}), \\
&\qquad + a_3 A_{3m}\frac{\sigma_2^2\bar{V}_2}{2\kappa_2^2}(1-e^{-\kappa_2\tau})\Big] \qquad \text{(A32)} \\
\mathrm{Var}(p-d) &= A_{1m}^2\frac{\sigma_x^2}{2\alpha} + A_{2m}^2\frac{\sigma_1^2\bar{V}_1}{2\kappa_1} + A_{3m}^2\frac{\sigma_2^2\bar{V}_2}{2\kappa_2}, \qquad \text{(A33)}
\end{aligned}$$

with

$$\sigma_x^2 = \varphi_x^2[\bar{V}_1\delta_x + \bar{V}_2(1-\delta_x)]$$

which are the same given in the text. We have used the unconditional covariance:

$$\begin{aligned}
< X_t, X_s > &= \frac{\sigma_x^2}{2\alpha}e^{-\alpha|t-s|} \\
< V_{it}, V_{is} > &= \frac{\sigma_i^2\bar{V}_i}{2\kappa_i}e^{-\kappa_i|t-s|}
\end{aligned}$$

for $i = 1, 2$.

Similar computation applies to obtain Equation (33). Q.E.D.

A.5 Predictability of volatilities

First, we prove Equation (32). To do so, we apply the following approximation:

$$\frac{1}{\tau}\int_0^\tau \exp(x_s)ds \approx \exp(\frac{1}{\tau}\int_0^\tau x_s ds) \qquad \text{(A34)}$$

for any process x_s. This is equivalent to an approximation of arithmetic mean by geometric mean. It is a good approximation when the variation of x_t is small in magnitude both across time and probability space. This is true for our variance processes because the magnitude of the variance is generally in the order of $10^{-3} \sim 10^{-4}$, and the variation of $\log V_t$ is within 1. Applying the above approximation to $\log V_t$, we have

$$\frac{1}{\tau}\int_0^\tau \sqrt{V_t}dt = \frac{1}{\tau}\int_0^\tau \exp(\frac{1}{2}\ln V_t)dt \approx \exp(\frac{1}{2\tau}\int_0^\tau \ln V_t dt). \qquad \text{(A35)}$$

Hence,

$$\begin{aligned}
\ln \frac{1}{\tau}\int_0^\tau \sqrt{V_t}dt &\approx \frac{1}{2\tau}\int_0^\tau \ln V_t dt \\
&= \frac{1}{2\tau}\left[\int_0^\tau \ln \bar{V} + \int_0^\tau \ln(1+\frac{V_t-\bar{V}}{\bar{V}})dt\right] \\
&\approx \frac{1}{2}\ln \bar{V} + \frac{1}{2\tau}\int_0^\tau \frac{V_t-\bar{V}}{\bar{V}}dt \\
&= Const + \frac{1}{2\tau\bar{V}}\int_0^\tau V_t dt, \qquad \text{(A36)}
\end{aligned}$$

which is Equation (32).

Because of the approximation above, we can express the volatilities as an integral of $b_1V_{1s} + b_2V_{2s}$ over $(t, t+\tau)$. Plugging these terms into the definition of the covariance, we then obtain Equation (33).

Then we provide the derivation of the AR(1) coefficient. Consider a stochastic process of the form

$$b_1V_{1t} + b_2V_{2t},$$

where b_1 and b_2 are constants. Due to independence between V_1 and V_2, the unconditional auto-covariance can be evaluated as

$$b_1^2\frac{\sigma_1^2\bar{V}_1}{2\kappa_1}\exp(-\kappa_1\tau) + b_2^2\frac{\sigma_2^2\bar{V}_2}{2\kappa_2}\exp(-\kappa_2\tau)$$

and the unconditional variance can be evaluated as

$$b_1^2\frac{\sigma_1^2\bar{V}_1}{2\kappa_1} + b_2^2\frac{\sigma_2^2\bar{V}_2}{2\kappa_2}.$$

Hence, the AR(1) coefficient can be computed easily based on above. Q.E.D.

A.6 Derivation of VRP

We derive the time t expected future realized variance over time period τ_0 under the risk-neutral probability. The market prices of risk for V_{1t} and V_{2t} are λ_3 and λ_4 of Equation (A17), hence the risk premia associated with V_{1t} and V_{2t} are

$$\lambda_3\sigma_1\sqrt{V_{1t}} = -\nu_1V_{1t}, \quad \text{and} \quad \lambda_4\sigma_2\sqrt{V_{2t}} = -\nu_2V_{2t}, \qquad \text{(A37)}$$

where

$$\nu_1 = \frac{1-\gamma\psi}{1-\gamma}A_2\sigma_1^2, \quad \text{and} \quad \nu_2 = \frac{1-\gamma\psi}{1-\gamma}A_3\sigma_2^2. \qquad \text{(A38)}$$

Hence, the risk-neutral processes for V_{1t} and V_{2t} are

$$\begin{aligned} dV_{1t} &= \kappa_1^Q(\frac{\kappa_1}{\kappa_1^Q}\bar{V}_1 - V_{1t})dt + \sigma_1\sqrt{V_{1t}}dw_{1t}^Q, \\ dV_{2t} &= \kappa_2^Q(\frac{\kappa_2}{\kappa_2^Q}\bar{V}_2 - V_{2t})dt + \sigma_2\sqrt{V_{2t}}dw_{2t}^Q, \end{aligned} \tag{A39}$$

where the risk-neutral mean-reversion coefficients for V_{it} are defined as

$$\kappa_i^Q = \kappa_i - \nu_i \tag{A40}$$

for $i = 1, 2$. In order for well-defined risk-neutral processes in Equation (A39), we need to have κ_i^Q's to be positive such that

$$\nu_1 < \kappa_1 \quad \text{and} \quad \nu_2 < \kappa_2. \tag{A41}$$

Now we compute the squared VIX, or more generally, variance swap rate VS_t with maturity τ_0, defined as the risk neutral expectation of the variance. Because the risk-neutral process of Equation (A39) and the physical process of Equation (22) are both Heston (1993) processes, we, following the derivation for (36), obtain Equation (37)

$$\text{VS}_t = \sum_{i=1}^{2} c_i(A_i^Q + B_i^Q V_{it}),$$

where the constants A_i^Q and B_i^Q $(i = 1, 2)$ are given by

$$A_i^Q = \frac{\kappa_i \bar{V}_i}{\kappa_i^Q}\left[1 - \frac{1 - e^{-\kappa_i^Q \tau_0}}{\kappa_i^Q \tau_0}\right], \qquad B_i^Q = \frac{1 - e^{-\kappa_i^Q \tau_0}}{\kappa_i^Q \tau_0}. \tag{A42}$$

Q.E.D.

A.7 GMM test

Denote $h(\theta)$ as the vector of target moments implied by the model given parameter set θ. We choose 23 target moments as described in the text. Let h_T be sample vector from data with size T corresponding to the target moments, and expressed as

$$h_T = \phi(g_T) \tag{A43}$$

with

$$g_T \equiv \frac{1}{T}\sum_{t=1}^{T} x_t \tag{A44}$$

where x_t is a vector representing market data, the details are given below. The GMM estimator $\{\theta_T : T \geq 1\}$ is defined as

$$\min_{\theta_T}[h(\theta_T) - h_T]'W[h(\theta_T) - h_T] \tag{A45}$$

for some positive definite weighting matrix W. If the model is true and data is stationary, then the GMM estimator must be consistent (Hansen 1982).

By optimizing the quadratic form of Equation (A45), and substituting Equation (A43) into the first order condition, we obtain

$$A_T[h(\theta_T) - h_T] = A_T[\phi(g(\theta_T)) - \phi(g_T)] = 0, \tag{A46}$$

with

$$A_T = \frac{\partial h'(\theta_T)}{\partial \theta_T} W, \tag{A47}$$

and

$$D_T = \frac{\partial h(\theta_T)}{\partial \theta_T'} \tag{A48}$$

For a consistent estimator θ_T, asymptotically we have Taylor expansion

$$\text{plim}[\phi(g(\theta_T)) - \phi(g_T)] = \frac{d\phi(\theta_0)}{d\mu} \times \text{plim}[g(\theta_T) - g_T] \tag{A49}$$

Let $A \equiv \text{plim} A_T$ and $D \equiv \text{plim} D_T$, following Zhou (1994), the covariance matrix for the target moments is

$$\Lambda_T = \frac{1}{T}(I - D(AD)^{-1}A) \left[\frac{d\phi}{d\mu}\right] S \left[\frac{d\phi}{d\mu}\right]' (I - D(AD)^{-1}A)'$$

where S is the spectral matrix defined as

$$S \equiv \sum_{j=-\infty}^{\infty} E x_t x_{t-j}.$$

Denote

$$J = (h(\theta_T) - h_T)\Lambda_T(h(\theta_T - h_T))'$$

which measures the sum of squared errors of target moments,

$$J \sim \chi^2(\# \text{ ofmoments} - \# \text{ of parameters}),$$

In addition, if J_r is J-statistics with the same covariance matrix for a restricted version of the model, then

$$J_r - J \sim \chi^2(\#\text{ of restrictions})$$

where the number of restrictions is the number of parameters that is restricted in one-factor model. Q.E.D.

A.8 Moment conditions for GMM test

In this appendix, we present the moment conditions for GMM estimation. The 23-dimensional vector $h_T(\theta)$ in the quadratic form $h_T'(\theta)W_T(\theta)h_T(\theta)$ that we choose to minimize are the differences between the model functions and their sample values. The first 15 moment functions are

$$\begin{array}{ccc} E(\Delta c) & \sigma(\Delta c) & \text{AC1}(\Delta c) \\ E(\Delta d) & \sigma(\Delta d) & \text{AC1}(\Delta d) \\ E(r_e) & \sigma(r_e) & \text{AC1}(r_e) \\ E(r_f) & \sigma(r_f) & \text{AC1}(r_f) \\ E(p-d) & \sigma(p-d) & \text{AC1}(p-d) \end{array}$$

Denote r_x as consumption growth Δc, dividend growth Δd, excess return r_e, risk free rate r_f, and price-dividend ratio $p-d$, the above moments are given as

$$\begin{aligned} \sigma(r_x) &= \sqrt{E[(r_x)^2] - E(r_x)^2} \\ \text{AC1}(r_x) &= \frac{E[r_{x,t+1}r_{x,t}] - E[r_x]^2}{E[r_x^2] - E[r_x]^2} \end{aligned}$$

where $E(r_x)$ are easy to compute analytically given the processes in the paper.

The 16th and 17th moments are $E[\text{VRP}]$, $\sigma(\text{VRP})$, the expectation and standard deviation of variance risk premium (VRP) defined in Equation (38).

The 18th to 20th moments are the regression coefficients β's of Equations (23), (24), and (25). All the 3 regression coefficients are of the form of

$$\begin{aligned} \beta &= \frac{\text{Cov}(r_{x,t+1}, (p_t - d_t))}{\text{Var}(p_t - d_t)} \\ &= \frac{E[r_{x,t+1}, (p_t - d_t)] - E[r_{x,t+1}] \cdot E[p_t - d_t]}{E[(p_t - d_t)^2] - E[p_t - d_t]^2} \end{aligned} \tag{A50}$$

where r_x are consumption growth, dividend growth, excess return, *resp.*

The 20th to 23rd moments are the three regression β's of Equation (30) for $\Delta t = 1$ year. Specifically, they are

$$\begin{aligned}\beta_{vol} &= \frac{\mathrm{Cov}(\ln \mathrm{Vol}_{\mathrm{t,t+\tau}}, (\mathrm{p_t} - \mathrm{d_t}))}{\mathrm{Var}(p_t - d_t)} \\ &= \frac{E[\ln \mathrm{Vol}_{\mathrm{t,t+\tau}} \cdot (\mathrm{p_t} - \mathrm{d_t})] - \mathrm{E}[\ln \mathrm{Vol}_{\mathrm{t,t+\tau}}] \cdot \mathrm{E}[\mathrm{p_t} - \mathrm{d_t}]}{E[(p_t - d_t)^2] - E[p_t - d_t]^2}\end{aligned}$$

where $\mathrm{Vol}_{\mathrm{t,t+\tau}}$ is given in Equation (31) and stands for volatility of consumption, dividend, and excess return, *resp.* With specification of all the moment conditions, and the analytical formula of the moments implied by the model that is solved in the paper, the GMM estimation and tests can be carried out as usual (see, e.g., Singleton, 2006).

We show the 26 elements of the moments in g_T as follows. The first 15 moments are:

$$E[r_x],\ E[r_x^2],\ E[r_{x,t+1}, r_{xt}]$$

where r_x stands for consumption growth, dividend growth, excess return, risk free rate, and price-dividend ratio.

Moments 16 and 17 are:

$$E[\mathrm{VRP_t}],\ \mathrm{E}[\mathrm{VRP_t^2}]$$

where VRP is the variance risk premium.

Moments 18 to 20 are:

$$E[r_{x,t+1}, (p_t - d_t)]$$

where r_x stands for consumption growth, dividend growth, and excess return.

Moments 21 to 26 are:

$$E[\log \mathrm{Vol_{xt}}],\ \mathrm{E}[(\log \mathrm{Vol_{xt}}), (\mathrm{p_t} - \mathrm{d_t})]$$

where x stands for consumption, dividend, and excess return. The data can be obtained through quarterly data regression

$$r_{x,t+1} = \alpha + \beta r_{x,t} + \epsilon_{x,t}$$

and annual expected volatility $\mathrm{Vol_{xt}}$ are obtained from

$$\mathrm{Vol_{xt}} = \sum_{\mathrm{k=1}}^{4} |\epsilon_{\mathrm{t+k}}| \tag{A51}$$

Finally, the form of function $h_T = \phi(g_T)$ that links the target functions h_T and the moments g_T, as well as its first-order derivative matrix are elementary, and can be obtained from authors upon request. Q.E.D.

References

Alizadeh, S., Brandt, M. W., Diebold, F. X., 2002. Range-based estimation of stochastic volatility Models. Journal of Finance 57, 1047–1091.

Adrian, A., Rosenberg, J., 2008. Stock returns and volatility: pricing the short-run and long-run components of market risk. Journal of Finance 63, 2997–3030.

Attanasio, P. O., Weber G., 1989. Intertemporal substitution, risk aversion and the Euler equation for consumption. Economic Journal 99, 59–73.

Avramov, D., Cederburg, S., 2012. Implications of long-run risk for asset allocation decisions. Working Paper, Hebrew University.

Avramov, D., Hore, S., 2009. Momentum, information uncertainty, and leverage - an explanation based on recursive preferences. Working Paper, University of Maryland.

Bansal, R., Yaron, A., 2004. Risks for the long run: a potential resolution of asset pricing puzzles. Journal of Finance 59, 1481–1509.

Bansal, R., Dittmar, R., Kiku, D., 2009. Cointegration and consumption risks in asset returns. Review of Financial Studies 22, 1343–1375.

Bansal, R., Gallant, R., Tauchen, G., 2007. Rational pessimism, rational exuberance, and asset pricing models. Review of Economic Studies 74, 1005–1033.

Bansal, R., Khatchatrian, V., Yaron, A., 2005. Interpretable asset markets? European Economic Review 49, 531–560.

Bansal, R., Kiku, D., Yaron, A., 2007a. A note on the economics and statistics of predictability: a long run risks perspective. Working paper, Duke University and University of Pennsylvania.

Bansal, R., Kiku, D., Yaron, A., 2007b. Risks for the long run: estimation and inference. Working paper, Duke University and University of Pennsylvania.

Bansal, R., Kiku, D., Shaliastovich, I., Yaron, A., 2012. Volatility, the macro- economy and asset prices. Working paper, Duke University and University of Pennsylvania.

Bansal, R., Shaliastovich, I., 2010. A Long-run risks explanation of predictability puzzles in bond and currency markets. Working paper, Duke University.

Beeler, J., Campbell, J. Y., 2012. The long-run risks model and aggregate asset prices: an empirical assessment. Critical Finance Review 1, 141–182.

Bollerslev, T., Tauchen, G., Zhou, H., 2009. Expected stock returns and variance risk premia. Review of Financial Studies 22, 4463–4492.

Bollerslev, T., N. Sizova, N., Tauchen, G., 2012. Volatility in equilibrium: asymmetries and dynamic dependencies. Review of Finance, 16, 331–80.

Breeden, D. T., 1979. An Intertemporal asset pricing model with stochastic consumption and investment opportunities. Journal of Financial Economics 7, 265–296.

Campbell, J. Y., 1993. Intertemporal asset pricing without consumption data. American Economic Review 83, 487–512.

Campbell, J. Y., Cochrane, J. H., 1999. By forces of habit: a consumption-based explanation of aggregate stock market behavior. Journal of Political Economy 107, 205–251.

Campbell, J., Shiller, R., 1988. The dividend-price ratio and discount factors. Review of Financial Studies 1, 195–228.

Campbell, J. Y., Vuolteenaho, T., 2004. Bad beta, good beta. American Economic Review 94, 1249–1275.

CBOE, 2009. The CBOE volatility index – VIX, White Paper.

Chacko, G., Viceira, L. M., 2005. Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets. Review of Financial Studies 18, 1369–1402.

Chen, L., Collin-Dufresne, P., Goldstein, R. S., 2009. On the Relation between the Credit Spread Puzzle and the Equity Premium Puzzle. Review of Financial Studies 22, 3367–3409.

Chen, L., Zhao, X., 2009. Return decomposition. Review of Financial Studies 22, 5213–5249.

Chernov, M., Ronald, G., Ghysels, E., Tauchen, G., 2003. Alternative models for stock price dynamics. Journal of Econometrics 116, 225–257.

Christoffersen, P. F., Jacobs, K., Ornthanalai, C., Wang, Y., 2008. Option valuation with long-run and short-run volatility. Journal of Financial Economics 90, 272–297.

Constantinides, G. M., Ghosh, A., 2011. Asset pricing tests with long run risks in consumption growth. Review of Asset Pricing Studies 1, 96–136.

Drechsler, I., Yaron, A., 2011. What's vol got to do with it. Review of Financial Studies 24, 1–45.

Duffie, D., Epstein, L. G., 1992. Asset pricing with stochastic differential utility. Review of Financial Studies 5, 411–436.

Egloff, D., Leippold, M., Wu, L., 2009. The term structure of variance swap rates and optimal variance swap investments. Journal of Financial and Quantitative Analysis 45, 1279–1310.

Epstein, L. G., Zin, S., 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. Econometrica 57, 937–969.

Eraker, B., 2008. The volatility premium. Working Paper, Duke University.

Eraker, B., Shaliastovich, I., 2008. An equilibrium guide to designing affine pricing models. Mathematical Finance 4, 519–543.

Ferson, W., Nallareddy, S., Xie, B., 2012. The "out-of-sample" performance of long-run risk models. Journal of Financial Economics, forthcoming.

Hall, R. E., 1988. Intertemporal substitution in consumption. Journal of Political Economy 96, 339–357.

Hansen, L. P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.

Hansen, L. P., Heaton, J. C., Li, N., 2008. Consumption strikes back? Measuring long-run risk. Journal of Political Economy 116, 260–302.

Hansen, L. P., Singleton, K., 1982. Generalized instrumental variables estimation of non-linear rational expectations models. Econometrica 50, 1269–1286.

Heston, S., 1993. Closed-form solution for options with stochastic volatility with application to bond and currency options. Review of Financial Studies 6, 327–343.

Jagannathan, R., Wang, Z., 1996. The conditional capm and the cross-section of expected returns. Journal of Finance 51, 3–53.

Lettau, M., Ludvigson, S., 2001. Consumption, aggregate wealth, and expected stock returns. Journal of Finance 56, 815–849.

Li, K., 2012. Asset pricing with a financial sector. Working Paper, Duke University.

Lu, Z., Zhu, Y., 2010. Volatility components: evidence from VIX futures market. Journal of Futures Markets 30, 230–256.

Pakoš, M., 2008. Asset prices under doubt about fundamentals. Working Paper, Carnegie Mellon University.

Singleton, K.J., 2006. Empirical dynamic asset pricing: model specification and econometric assessment. Princeton University Press.

Stock, J., Watson, M., 2002. Has the business cycle changed and why. M. Gertler and K. Rogo (eds.) NBER Macroeconomics Annual, MIT Press.

Vissing-Jorgensen, A., 2002. Limited asset market participation and the elasticity of intertemporal substitution. Journal of Political Economy 110, 825–853.

Weil, P., 1989, The equity premium puzzle and the risk-free rate puzzle. Journal of Monetary Economics 24, 401–421.

Zhang, J. E., Zhu, Y., 2006. VIX futures. Journal of Futures Markets 26, 521–531.

Zhou, G., 1994. Analytical GMM tests: asset pricing with time-varying risk premiums. Review of Financial Studies 7, 687–709.

Zhou, G., Zhu, Y., 2011. Volatility trading: what is the role of the long-run volatility component? Journal of Financial and Quantitative Analysis 47, 273–307.

Table 1: Long-run risks parameters

The table reports the parameters for the three calibrated long-run risks models: the Bansal and Yaron (BY, 2004), and Bansal, Kiku, and Yaron (BKY, 2007a), the optimized one- and two-factor model. γ is the risk aversion parameter, ψ the IES parameter, β the discount rate. Other panels provide parameters governing the consumption, dividend and volatility dynamics:

$$
\begin{aligned}
\frac{dC_t}{C_t} &= (\mu + X_t)dt + \sqrt{V_{1t}\delta_c + V_{2t}(1-\delta_c)}dZ_{1t} \\
dX_t &= -\alpha X_t dt + \varphi_x \sqrt{V_{1t}\delta_x + V_{2t}(1-\delta_x)}dZ_{2t} \\
\frac{dD_t}{D_t} &= (\mu_d + \phi X_t)dt + \varphi_d \sqrt{V_{1t}\delta_d + V_{2t}(1-\delta_d)}dB_t + \sigma_{dc}\sqrt{V_{1t}\delta_c + V_{2t}(1-\delta_c)}dZ_{1t} \\
&\quad + \sigma_{dx}\sqrt{V_{1t}\delta_x + V_{2t}(1-\delta_x)}dZ_{2t} + \sigma_{dv}\sqrt{V_{1t}}dw_{1t} + \sigma_{dv2}\sqrt{V_{2t}}dw_{2t} \\
dV_{1t} &= \kappa_1(\bar{V}_1 - V_{1t})dt + \sigma_1\sqrt{V_{1t}}dw_{1t} \\
dV_{2t} &= \kappa_2(\bar{V}_2 - V_{2t})dt + \sigma_2\sqrt{V_{2t}}dw_{2t}, \quad \kappa_1 < \kappa_2,
\end{aligned}
$$

where the parameter values are annualized whenever applicable.

Preference parameters								
	γ	ψ	β					
BY	10	1.5	2.4%					
BKY	10	1.5	3%					
New	10	1.5	0.9%					
Consumption growth dynamics								
	μ	α	φ_x	δ_c	δ_x			
BY	0.018	0.256	0.528	1	1			
BKY	0.018	0.3	0.456	1	1			
New	0.018	0.256	0.522	1.000	0.515			
Dividend growth dynamics								
	μ_d	ϕ	φ_d	δ_d	σ_{dc}	σ_{dx}	σ_{dv}	σ_{dv2}
BY	0.018	3	4.5	1	0	0	0	0
BKY	0.018	2.5	5.96	1	2.6	0	0	0
New	0.020	4.368	1.792	1.000	-0.111	-3.062	2.001	-2.875
Volatility parameters								
	Factor 1			Factor 2				
	$\bar{V}_1$	σ_1	κ_1	$\bar{V}_2$	σ_2	κ_2		
BY	0.027^2	0.0035	0.156					
BKY	0.025^2	0.0027	0.012					
New	0.024^2	0.0026	0.035	0.024^2	0.698	9.640		

Table 2: GMM moments

This table presents the 23 target moments of the data. The fitted moments of the BY and BKY models are taken from Beeler and Campbell (2012). The fitted moments of our model are based on the GMM estimates of Table 3. The data are annual from 1930 to 2007.

	Data	BY	BKY	New		Data	BY	BKY	New
$E(\Delta c)$	1.95	1.79	1.82	1.73	$\beta(r_e)$	-0.059	-0.007	-0.078	-0.073
$\sigma(\Delta c)$	2.16	2.92	2.96	2.22	$\beta(\Delta c)$	0.012	0.114	0.022	0.039
AC1(Δc)	0.44	0.51	0.44	0.43	$\beta(\Delta d)$	0.064	0.343	0.054	0.174
$E(\Delta d)$	1.02	1.66	1.85	1.31	$\beta_{vol}(r_e)$	-0.081	-0.123	-1.315	-0.167
$\sigma(\Delta d)$	10.69	11.57	16.42	11.28	$\beta_{vol}(\Delta c)$	-0.481	-0.128	-1.420	-0.317
AC1(Δd)	0.14	0.40	0.29	0.32	$\beta_{vol}(\Delta d)$	-0.530	-0.146	-1.483	-0.546
$E(r_e)$	6.20	6.62	6.58	6.25	VRP	-8.89	-0.005	-0.010	-5.57
$\sigma(r_e)$	18.34	16.88	21.35	19.21	σ(VRP)	26.98	0.000	0.000	29.71
AC1(r_e)	0.04	0.03	0.02	0.01					
$E(r_f)$	0.99	2.56	0.99	1.18					
$\sigma(r_f)$	4.28	1.30	1.28	1.87					
AC1(r_f)	0.59	0.85	0.86	0.43					
$E(p-d)$	3.31	3.00	3.04	2.57					
$\sigma(p-d)$	0.46	0.16	0.26	0.27					
AC1$(p-d)$	0.88	0.77	0.95	0.86					

Table 3: Decomposing the variance of the pricing kernel

This table presents the volatility of pricing kernel (maximal Sharpe ratio) and the relative contribution of different shocks to the variance of the pricing kernel, for both our new model and BY model. The result of our new model is based on the GMM estimation, and that for the BY model is from Bansal and Yaron (2004). The contribution from the fluctuating economic uncertainty for our new model comes from two volatility factors, in contrast to BY's one volatility factor.

	Relative variance of shocks				
	Volatility of Pricing Kernel	Independent Consumption	Expected Growth Rate	Fluctuating Economic Uncertainty	
				Factor 1	Factor 2
New	0.60	15%	49%	25%	11%
BY	0.73	14%	47%	39%	

Table 4: Model comparison

The table reports the GMM parameter estimates and their asymptotic standard errors of the new model. We denote J is the χ^2 test statistic for the moments conditions of the new model, which is χ^2 distributed with degree of 4. We also report the GMM test results for the BY or BKY models which are restricted versions of the new model.

Parameter	Estimate	Standard Error
μ	0.018	0.032
μ_d	0.020	0.012
α	0.256	0.467
ϕ	4.368	0.113
φ_x	0.522	0.521
φ_d	1.792	0.027
δ_x	0.515	0.616
δ_c	1.000	1.126
δ_d	1.000	0.052
σ_{dc}	-0.111	0.051
σ_{dv}	2.001	0.086
σ_{dv2}	-2.875	0.084
σ_{dx}	-3.062	0.134
κ_1	0.035	0.868
$\bar{V}_1 = \bar{V}_2$	0.001	0.002
σ_1	0.003	0.078
κ_2	9.640	0.033
σ_2	0.698	0.872
β	0.009	0.115
J	7.53	
p-value	0.11	
Model Comparison with BY and BKY		
	BY	BKY
$J_r - J$	293.40	376.28
p-value	0.00	0.00

Table 5: Variance decomposition

The table reports the attribution of return variance to discount-rate shocks, cash-flow shocks, and covariance between the two shocks. The decomposition is performed on the BKY model and the new two-factor model.

Panel A: BKY model	
	Contribution to Total Variance
Discount-Rate Risk	0.082
Cash-Flow Risk	0.995
Covariance	-0.076
Total Variance	1.000
Panel B: New model	
Discount-Rate Risk	0.233
Cash-Flow Risk	0.655
Covariance	0.112
Total Variance	1.000

Table 6: Variance risk premium

The table reports variance risk premiums (VRP), the standard deviation (std) of VRP, and the monthly autoregression coefficient (AR1) for the market data, the Drechsler and Yaron (2011), the BY, the BKY and our new models. The results of the first two columns are from Drechsler and Yaron (DY, 2011). The third column is the data computed from our data set. All the values are monthly and multiplied by 10000.

	Data (DY)	DY	Data (ours)	BY	BKY	New
VRP	-12.67	-7.57	-8.89	-0.005	-0.010	-5.57
std	14.38	10.65	26.98	0.000	0.000	29.71
AR1	0.54	N/A	0.15	0.99	0.99	0.47