Sharpe Ratio Volatility: Is It a Puzzle?*

Alberto Romero
Sauder School of Business
University of British Columbia

JOB MARKET PAPER

This Version: November 23, 2012

Abstract

Recent literature empirically documents large time-series variation in the market Sharpe ratio, spurring theoretical explanations of this phenomenon. I revisit the empirical evidence and ask whether estimates of Sharpe ratio volatility may be biased due to limitations of the standard OLS methods used in estimation. Based on simulated data from a standard calibration of the long-run risks model, I find that OLS methods used in prior literature produce Sharpe ratio volatility five times larger than its true variability. The difference arises due to measurement error. To address this issue, I propose the use of filtering techniques that account for the Sharpe ratio's time variation. I find that these techniques produce Sharpe ratio volatility estimates of less than 15% on a quarterly basis, which matches more closely the predictions of standard asset pricing models. Additionally, my results have practical implications for portfolio allocation, where upward-biased estimates of Sharpe ratio volatility imply excessive portfolio rebalancing.

JEL classification: G12; C32.

Keywords: Sharpe ratio; Filtering; OLS methods; Time-varying moments of returns; Volatility.

*I am grateful to Adlai Fisher and Lorenzo Garlappi for their guidance and supervision. I also thank Harjoat Bhamra, Murray Carlson, Jason Chen, Alan Kraus, Howard Kung, Tan Wang, Milka Dimitrova, Vincent Gregoire and Gonzalo Morales for their comments and suggestions. All errors are mine. Contact information: 2053 Main Mall, Vancouver BC, Canada V6T 1Z2; Email: alberto.romero@sauder.ubc.ca.

1 Introduction

The Sharpe ratio measures the excess return of an investment relative to its standard deviation. Most leading consumption-based asset pricing theories imply a relatively stable market Sharpe ratio. However, empirical evidence suggests that there is more variability in the Sharpe ratio than standard models account for. Recently, Lettau and Ludvigson (2010) suggest that the finance literature should address this "Sharpe ratio variability puzzle." They document that the empirical standard deviation of the estimated Sharpe ratio is about 47% per quarter. In contrast, a quarterly calibration of the standard Campbell and Cochrane (1999) model produces a substantially lower volatility of 9%. In turn, Chien, Cole, and Lustig (2012) suggest that passive investors' infrequent rebalancing explains the high variability of market Sharpe ratios.

In this paper, I examine whether estimates of the variability of the Sharpe ratio might be biased due to limitations of the empirical methodology used in estimation. In particular, I show that measurement error in estimated Sharpe ratios may help to explain the Sharpe ratio volatility puzzle. To do this, I simulate data from a standard calibration of the Bansal and Yaron (2004) long-run risks (LRR) model. Following practice common in the literature, I then estimate Sharpe ratios using ordinary least squares (OLS) methods to infer the variability of the model-generated Sharpe ratios. OLS methods lead to estimates of Sharpe ratio volatility of approximately 18%, even though the true variability of the model-implied Sharpe ratio is only 3%. The difference in estimates is due to measurement error induced by the standard Sharpe ratio estimation methodology.

Once I have documented the difference between the Sharpe ratio's estimated and true volatility, I study whether improved empirical methodologies might better account for the true variability of the Sharpe ratio. In particular, I implement filtering methods, which are statistical tools that recover unobservable state variables using measurements that are observed with noise. These techniques are flexible enough to allow the econometrician to perform statistical inference based on time-varying information observed with measurement error. Moreover, the modeling representation is general enough to include time varying information as well as flexible correlation structures and errors in variables.[1]

I use two different exercises to show the limitations of OLS methods. First, I run a controlled experiment in which I have full information of the data generating process of stock returns, state variable dynamics and parameter values. I use simulated data from the LRR model as calibrated

[1] See Hamilton (1994), Kim and Nelson (1999) and Doucet, de Freitas, Gordon, and Smith (2001) for an introduction to Filtering methods. Crisan and Rozovskii (2011) provide a more recent literature review in nonlinear filtering methods.

by Bansal and Yaron (2004) to estimate conditional means, variances and Sharpe ratios. The use of artificial data from a fully specified economy is important because it allows the economic reasons that drive the variation in Sharpe ratios to be isolated. Moreover, information about model specification and state-variable dynamics is incorporated within the filtering estimation procedure. Furthermore, the tractability of the LRR model allows data to be simulated with relative ease.

I then implement two econometric techniques: I run standard OLS regressions and then I apply filtering techniques. I compare both sets of results with the closed-form expressions implied by the LRR model as a benchmark. My results show that the Sharpe ratios based on standard OLS methods are more volatile than the estimates obtained with filtering techniques. Moreover, the volatility estimates obtained via filtering differ from the true value by less than 1%, which is a significant improvement over OLS estimates. The main driver of this result is the use of conditioning information within the estimation process.

There are a number of reasons why a filtering approach can improve upon predictive regressions to estimate expected returns and conditional volatilities. First, filtering explicitly acknowledges that both expected returns and volatilities are time varying. Filtering techniques aggregate the entire history of realized returns parsimoniously; in contrast, predictive regressions use lagged predictors to form estimates of expected returns and volatilities. Instead of adding lags to a vector autorregressive (VAR) model, which would increase the number of parameters to be estimated, a latent variable approach such as filtering incorporates the information contained in the history of observed returns. Moreover, filtering techniques are flexible enough to be used with large information sets without relying on additional instruments that may be misspecified (Ferson, Sarkissian, and Simin (2003)). Finally, filtering is more robust to structural breaks than are OLS techniques (Rytchkov (2012)), since it is insensitive to robust shifts in relations over the long run. For example, in the predictability literature, a substantial shift in the dividend-price ratio destroys its forecasting power.[2] Also, robustness to structural breaks makes the filtering approach more valuable from an ex-ante point of view, when it is unclear whether structural breaks will occur.

The standard method used in the literature to estimate Sharpe ratios is to use fitted moments from first-stage predictive regressions as proxies for the unobserved conditional mean and volatility. Such a technique has some important drawbacks. First, the dynamics of the conditional mean and volatility are determined by the joint conditional distribution of the first-stage predictors. Thus, with any model misspecification, such as omitted variables, the dynamics of the fitted moments would not necessarily correspond to the dynamics of the true moments. In

[2]See Lettau and Van Nieuwerburgh (2008) for a detailed explanation.

addition, even if the predictive models for the conditional mean and volatility are well specified, the effect of errors in variables, which are induced by the first-stage regressions, is not trivial to quantify in a VAR model.

Simulating data of stock returns by means of theoretical models is a powerful tool because the economic reasons that drive the simulated time-series variation are fully identified. However, theoretical models are abstractions, and by definition misspecified. An alternative form of analyzing stock returns is via reduced form models, which are statistical representations that do not impose any economic structure and thus aim to better describe historical data. To infer Sharpe ratios and their variability from the data, I conduct a second exercise based on the reduced form model by Brandt and Kang (2004). In this model, expected returns and volatilities are estimated as latent variables and identified from the history of returns. The main advantage of this approach is that it does not rely on prespecified predictors and is not subject to errors in variables or model misspecification. I apply filtering techniques to estimate the parameters of the model and to extract estimates of conditional moments of returns as well as conditional Sharpe ratios. As a result, my estimate for quarterly Sharpe ratio volatility using the reduced form model is in the order of 5% to 10%, whereas my estimate for the quarterly Sharpe ratio volatility using the OLS methods is 42%.

Consistent with the results of the simulation exercise, I find that conditioning information drives the results above. Reduced form models do not rely on predetermined conditioning variables to estimate conditional moments: The state variables are identified from the history of returns. Standard OLS techniques generate fitted moments from a set of predictive regressions as proxies for the unobservable conditional mean and volatility. The fitted moments depend on the joint distribution of these predictors. Consequently, any model misspecification would generate fitted moments that do not correspond to the true dynamics of the conditional mean and volatility, and thus, the dynamics of the Sharpe ratio.

My findings have important implications in an asset management context since the Sharpe ratio is a commonly used measure of performance evaluation. For investors willing to allocate their wealth between the market portfolio and the risk-free instrument, the market Sharpe ratio becomes a natural benchmark of their investments. If this ratio is highly volatile, the variation needs to be taken into account for hedging and rebalancing purposes. Indeed, Lustig and Verdelhan (2012) report that accounting for time variation in Sharpe ratios may lead to optimal trading strategies that differ markedly from buy-and-hold strategies.

Furthermore, a mean-variance investor would have an obvious interest in understanding the volatility of Sharpe ratios. For example, in a partial equilibrium setting,[3] the Sharpe ratio

[3]Some examples are Merton (1969) and Merton (1971).

determines the fraction of wealth that an agent invests in the market portfolio. I show that if an investor uses OLS methods to determine this fraction, then the portfolio weights would exhibit extremely volatile behavior over time, which may result in high rebalancing costs. I also show that if an investor applies filtering techniques to estimate the fraction of wealth invested in the market portfolio, then these costs will be substantially lower. Even further, for a representative agent with habit formation preferences, the Sharpe ratio indicates the timing and magnitude of fluctuations of risk aversion (Constantinides (1990) and Campbell and Cochrane (1999)). Thus, the time variation in the market Sharpe ratio may provide information about the fundamental economics underlying the asset prices.

A number of studies analyze the predictable variation of the mean and volatility of stock returns from an empirical point of view.[4] However, only a few papers have investigated the time variation observed in equity Sharpe ratios. Lettau and Ludvigson (2010) measure the conditional Sharpe ratio of U.S. equities by forecasting stock market returns and realized volatility using different predictors. They obtain highly counter-cyclical and volatile Sharpe ratios and show that neither the external habit model of Campbell and Cochrane (1999) nor the LRR model Bansal and Yaron (2004) deliver Sharpe ratios volatile enough to match the data. Using a latent VAR process, Brandt and Kang (2004) also find a highly counter-cyclical Sharpe ratio. Ludvigson and Ng (2007) document the same result using a large number of predictors in a dynamic factor analysis.

Tang and Whitelaw (2011) document predictable variation in stock market Sharpe ratios. Based on a predetermined set of financial variables, the conditional mean and volatility of equity returns are constructed and combined to estimate the conditional Sharpe ratios. Tang and Whitelaw (2011) find that conditional Sharpe ratios show substantial time variation that coincides with the phases of the business cycle. Lustig and Verdelhan (2012) provide evidence that Sharpe ratios are higher in recessions than in expansions in the United States and other OECD countries. They also find that the changes in expected returns during business-cycle expansion and contractions are not explained by changes in near-term dividend growth rates. These papers focus on the counter-cyclical behavior of Sharpe ratios. My paper focuses on the conditional volatility of market Sharpe ratios and finds that the volatility estimates are substantially smaller than the evidence previously documented.

My paper is also related to Brandt and Kang (2004), Pástor and Stambaugh (2009), van Binsbergen and Koijen (2010) and Rytchkov (2012), who analyze return predictability using state-space models.[5] I contribute to the literature by focusing on the dynamic behavior of the

[4]See Lettau and Ludvigson (2010) for a comprehensive survey.

[5]In an early work in this body of literature, Conrad and Kaul (1988) use the Kalman filter to extract expected

market Sharpe ratio and by showing that standard OLS methods as applied in the literature generate measurement error which impacts estimates of Sharpe ratio volatility. Moreover, I also show that filtering techniques are a good approach for estimating the ratio's true volatility. I also find that filtering techniques are better able to capture the dynamic behavior of market Sharpe ratios.

The remainder of this paper is organized as follows. Section 2 provides a theoretical framework to interpret Sharpe ratios. Section 3 introduces the LRR model and its implications for empirical moments. Section 4 describes the simulation exercise as well as the estimation methodologies for expected returns, volatilities and Sharpe ratios. In section 5, an analysis of Sharpe ratios in reduced form models is described and the empirical results are shown. Section 6 presents asset allocation implications. Finally, conclusions are presented in section 7.

2 Sharpe Ratios in Asset Pricing

The conditional Sharpe ratio of any asset at time t, denoted by SR_t, is defined as the ratio of the conditional mean excess return to its conditional standard deviation; that is

$$SR_t = \frac{\mathbb{E}_t\left[R_{t+1} - R_{ft+1}\right]}{\sigma_t\left[R_{t+1} - R_{ft+1}\right]}, \tag{1}$$

where R_t and R_{ft} denote the gross asset return of an asset and the one-period risk-free interest rate, respectively, and the conditional expectations are based on the information available at time t.

Harrison and Kreps (1979) show that the absence of arbitrage implies the existence of a stochastic discount factor (SDF) or pricing kernel, denoted by M_t, that prices all assets in the economy.[6] An implication of no arbitrage is that the expectation of the product of the stochastic discount factor and the gross asset return of any asset must be equal to one; that is,

$$\mathbb{E}_t\left[M_{t+1}R_{t+1}\right] = 1. \tag{2}$$

An implication of (2) is that the conditional Sharpe ratio is proportional to the risk-free rate, the volatility of the pricing kernel and the correlation between the pricing kernel and the return; that is

$$SR_t = -R_{ft+1}\sigma_t\left[M_{t+1}\right]Corr_t\left[R_{t+1}, M_{t+1}\right], \tag{3}$$

returns, but only from the history of realized returns. Other studies that relate latent variables with predictability include Lamoureux and Zhou (1996), Ang and Piazzesi (2003) and Dangl and Halling (2012).

[6] A detailed explanation is shown in Appendix A.

where σ_t and $Corr_t$ are the standard deviation and correlation, conditional on information at time t, respectively. The conditional Sharpe ratio of any asset in the economy is time varying as long as the risk-free rate varies or the pricing kernel is conditionally heteroskedastic -that is, $\sigma_t [M_{t+1}]$ changes over time- or if the correlation between the stock market return and the SDF is time varying. In general, each model defines an SDF; therefore, we learn from Eq. (3) that in this setup, Sharpe ratios are model dependent. In this paper, I focus on the conditional Sharpe ratio of the aggregate stock market, which is defined as the instrument that pays the aggregate dividend every period. However, the analysis can be extended to the Sharpe ratios of any traded asset.

The next section presents the LRR model of Bansal and Yaron (2004), with a particular focus on the implications for expected returns, volatilities and Sharpe ratios of the aggregate stock market. This model explains stock price variation as a response to persistent fluctuations in the mean and volatility of aggregate consumption growth by a representative agent with a high elasticity of intertemporal substitution. The tractability of the LRR model allows data to be simulated with relative ease. It provides analytical expressions for expected returns, volatilities and Sharpe ratios for the market portfolio, conditional on the Campbell and Shiller (1988) log-linearizations. Later in the paper I briefly present other asset pricing models and their implications for market Sharpe ratios.

3 The Long-Run Risks Model

Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012a) (BY and BKY hereafter) propose the following stochastic processes for the log-consumption and log-dividend growth, denoted by Δc_{t+1} and Δd_{t+1}, respectively:

$$\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_t + \sigma_t \eta_{t+1} \\
x_{t+1} &= \rho x_t + \varphi_e \sigma_t e_{t+1} \\
\sigma_{t+1}^2 &= \overline{\sigma}^2 + v\left(\sigma_t^2 - \overline{\sigma}^2\right) + \sigma_w w_{t+1} \\
\Delta d_{t+1} &= \mu_d + \phi x_t + \varphi \sigma_t u_{t+1} + \pi \sigma_t \eta_{t+1} \\
w_{t+1}, e_{t+1}, u_{t+1}, \eta_{t+1} &\sim i.i.d.\ \mathcal{N}(0,1),
\end{aligned} \tag{4}$$

where, x_t is a persistently varying component of the expected consumption growth rate and σ_t^2 is the conditional variance of consumption growth, which is time varying and highly persistent, with unconditional mean $\overline{\sigma}^2$. The variance process can take negative values, but it will happen with small probability if its conditional mean is high enough with respect to its variance.

Dividends are correlated with consumption since the growth rate, Δd_{t+1}, shares the same persistent predictable component scaled by a parameter ϕ, and the conditional volatility of dividend growth is proportional to the conditional volatility of consumption growth.

BY solve the LRR model using analytical approximations. They assume assume a representative agent with Epstein-Zin utility with time discount factor δ, coefficient of relative risk aversion γ, and elasticity of intertemporal substitution ψ. The log of the stochastic discount factor, m_{t+1}, for this economy is given by

$$m_{t+1} = \theta \log \delta - \frac{\theta}{\psi} \Delta c_{t+1} + (\theta - 1)\, r_{a,t+1}, \tag{5}$$

where $\theta = (1-\gamma) / (1-\psi)$ and $r_{a,t+1}$ is the return on the consumption claim, or equivalently, the return on aggregate wealth. BY use the Campbell and Shiller (1988) log-linearizations to obtain analytical approximations for the returns on the consumption and dividend cliams. Further details on the model and derivations are explained in Appendix B.

3.1 Implications for Expected Returns, Volatilities and Conditional Sharpe Ratios

Under the long-run risks framework, the equity premium is an affine function of the volatility of consumption growth alone:

$$\mathbb{E}_t \left[r_{m,t+1} - r_{f,t+1} \right] = E_0 + E_1 \sigma_t^2. \tag{6}$$

The model also implies that the conditional variance of the market return is an affine function of σ_t^2 :

$$Var_t \left(r_{m,t+1} \right) = D_0 + D_1 \sigma_t^2. \tag{7}$$

The coefficients E_0, E_1, D_0 and D_1 are known functions of the underlying time-series and preference parameters. The general expressions and details about their derivation are shown in Appendix B.4.

The covariance between the observed market excess return, $r_{m,t+1}$, and the innovation to the volatility process, w_{t+1} is given by

$$cov_t \left(r_{m,t+1}, w_{t+1} \right) = \kappa_{1,m} A_{2,m} \sigma_w. \tag{8}$$

One of the appealing properties of the long-run risk model, is that $A_{2,m} < 0$, for standard calibrations, implying that the LRR model is able to reproduce the negative feedback effect.[7] Another implication of Eq. (8), is that the conditional correlation between excess returns and the innovations to consumption risk is time varying because the conditional variance of stock returns is time varying.

Following Bansal, Kiku, and Yaron (2012b), the cumulative log-return over K time periods is just a sum of K one-period returns,[8]

$$\sum_{k=1}^{K} \left(r_{m,t+k} - r_{f,t+k}\right).$$

The conditional moments are given by

$$\mathbb{E}_t \left[\sum_{k=1}^{K} r_{m,t+k} - r_{f,t+k}\right] = \mathbf{E}_{0,K} + \mathbf{E}_{1,K}\sigma_t^2, \tag{9}$$

and

$$Var_t \left[\sum_{k=1}^{K} r_{m,t+k} - r_{f,t+k}\right] = \mathbf{D}_{0,K} + \mathbf{D}_{1,K}\sigma_t^2, \tag{10}$$

where $\mathbf{E}_{0,K}$, $\mathbf{E}_{1,K}$, $\mathbf{D}_{0,K}$, and $\mathbf{D}_{1,K}$ are known functions of the preference parameters and the number of periods, K, used for time aggregation. If the time unit is the month, then evaluating Eqs. (9) and (10) provides an expression for annual estimates.

The conditional Sharpe ratio of an investment over K time periods is given by the ratio of the conditional mean returns divided by its conditional standard deviation, and is represented by

$$SR_{t,t+K} = \frac{\mathbf{E}_{0,K} + \frac{\mathbf{D}_{0,K}}{2} + \left(\mathbf{E}_{1,K} + \frac{\mathbf{D}_{1,K}}{2}\right)\sigma_t^2}{\sqrt{\mathbf{D}_{0,K} + \mathbf{D}_{1,K}\sigma_t^2}}. \tag{11}$$

Eq. (11) implies that the only source of variation in the conditional Sharpe ratio under the LRR framework is the volatility of consumption growth. Moreover, the conditional Sharpe ratio is stochastic unless σ_t^2 is deterministic. Furthermore, under the standard calibrations by BY, the conditional Sharpe ratio is strictly increasing in the volatility of consumption growth. This implies that the long-run risk framework predicts counter-cyclical Sharpe ratios; that is for bad

[7]Campbell and Hentschel (1992), Glosten, Jagannathan, and Runkle (1993), and Brandt and Kang (2004), among others document the volatility feedback effect; that is, return innovations are negatively correlated with innovations in market volatility.

[8]Time aggregation is an important mechanism for parameter and state inference. Bansal, Kiku, and Yaron (2012b) explicitly consider time aggregation of variables. They find that time aggregation can affect parameter values and they provide evidence that ignoring time aggregation leads to false rejection of the LRR model. Earlier papers that account for time aggregation in estimation in asset pricing context include Hansen and Sargent (1983) and Heaton (1995).

times (high values of the volatility of consumption growth) the Sharpe ratios are high and for good times, (low values of volatility of consumption growth) the conditional Sharpe ratios are low, consistent with the habit formation model of Campbell and Cochrane (1999).

Moreover, Eqs. (9) , (10) and (11) characterize the expected return, volatility and Sharpe ratios of a buy and hold strategy over K time units. These equations define the term structure of risk premia, volatility, and Sharpe ratios of the market portfolio. Moreover, by evaluating Eqs. (9) and (10) in the unconditional value of the volatility of consumption growth, σ^2, we obtain expressions for the unconditional moments of cumulative returns. Similar expressions can be obtained for the cumulative return moments of the risk-free instrument and market portfolio. Details about the derivations are described in Appendix C.

4 Simulation Exercises

In this section, I conduct a simulation study in the spirit of Beeler and Campbell (2012). The objective is to simulate equity returns from the LRR model at a monthly frequency, and then time aggregate them to obtain annual estimates of returns, volatilities and Sharpe ratios. I explain the simulation exercise as follows.

First, I generate four sets of independent standard normal random variables and use them to construct monthly series for consumption, dividends and state variables using the state-space model in Eq.(4).[9] Next, I construct annual consumption and dividend growth by adding twelve monthly consumption and dividend levels, respectively, and then take the growth rate of the sum. The log market returns and risk-free rates are the sum of monthly values, while the log price-dividend ratios use prices measured from the last period of the year. As the price-dividend ratio in the data is divided by the previous year's dividends, the price-dividend ratio in the model is multiplied by the dividend in that month and divided by the dividends over the previous year.

As in BY, BKY and Beeler and Campbell (2012), negative realizations of the conditional variance are censored and replaced with a small positive number.[10] I also retain sample paths along which the volatility process goes negative and is censored.[11] Since the volatility is highly persistent, it is quite likely to have negative values for the conditional variance; indeed, Beeler

[9]The frequency is consistent with the parameters calibrated by BY and BKY, which are provided in monthly terms.

[10]The number is (10^{-14}) and is consistent with the simulation exercise of Beeler and Campbell (2012).

[11]An alternative approach is to replace negative realizations with their absolute values, as in Johnson and Lee (2012).

and Campbell (2012) report that under the BK calibration less than 1% of the volatility simulations are negative for a sample of 100,000 simulations. Each simulation is initialized from the steady-state values and run for a "burn-in" period of ten years.

4.1 Predictive Regressions

The conditional moments of market returns as well as the Sharpe ratio are unobservable. A common approach that has been applied in the empirical literature to circumvent this issue is to project excess stock returns series on a predetermined set of conditioning variables, such as economic or financial indicators observed by the econometrician.

Empirical studies differ in the conditioning information used in projection of excess returns. Among the most commonly used predictor variables are the price-dividend ratios (Fama and French (1988a); Campbell (1991); Hodrick (1992)), short-term interest rates Fama and Schwert (1977); Campbell (1991); Hodrick (1992); Ang and Bekaert (2007)), term spreads and default spreads (Fama and French (1988a)), book market ratios (Lewellen (1999); Vuolteenaho (2000)), proxies for consumption-wealth ratio (Lettau and Ludvigson (2001a,b)), and latent factors obtained from large data sets (Ludvigson and Ng (2007)).[12] Expected returns are calculated by regressing realized returns on the set of predictors and taking the fitted values as estimates.

Conditional volatility may also be measured by a projection onto predetermined conditioning variables, taking the fitted value from this projection as a measure of conditional variance or conditional standard deviation. This type of modeling is commonly used; for example French, Schwert, and Stambaugh (1987) use a time-series model of realized variance to model the conditional variance.

Within the set of techniques to measure conditional volatility by a projection onto predetermined conditioning variables, three approaches are common. One is to take the squared residuals from a regression of excess returns onto a predetermined set of conditioning variables and regress them on to the same set of conditioning variables, using the fitted values from this regression as a measure of conditional variance.[13] Alternatively, volatility can be estimated using high-frequency return data, commonly referred to as realized volatility. This is an ex-post measure that consists of adding up the squared high-frequency returns over the period of interest. The realized volatility is then projected onto time t information variables

[12] Lettau and Ludvigson (2010) and Goyal and Welch (2008) provide a comprehensive review of predictive variables commonly used in the literature.

[13] Campbell (1987) and Breen, Glosten, and Jagannathan (1989) apply these methods in the predictability literature.

to obtain a consistent estimate of the conditional variance of returns.[14] The third approach estimates conditional volatility of excess stock market returns by specifying a parametric form for the conditional volatility, such as the GARCH type of models, or stochastic volatility.[15] The volatility estimates are then obtained from the history of observed returns. For this part of the paper, I focus on the second type of methodology to calculate conditional volatilities of stock returns by projecting the sum of squared monthly returns on a set of predictors.

As for the conditional Sharpe ratio, a standard measure is the ratio of the estimated expected excess return to the estimated volatility, both obtained from separate projections. This approach has been taken by Kandel and Stambaugh (1990), Tang and Whitelaw (2011) and Lettau and Ludvigson (2010), among others.

I model the the conditional moments of annual returns as follows:

$$\mathbb{E}_t\left[R_{t+1} - R_{f,t+1}\right] = X_t\beta_\mu, \tag{12}$$
$$Var_t\left[R_{t+1} - R_{f,t+1}\right] = X_t\beta_\lambda, \tag{13}$$

where X_t is the set of predictor variables observed at time t and $R_{t+1} - R_{f,t+1}$ is the annual excess return on the market. I assume that the predictor variables available at time t are the price-to-dividend ratio, the current excess returns and the risk-free rate, constructed in an annual basis.

The regression equations that correspond to (12) and (13) are

$$R_{t+1} - R_{f,t+1} = X_t\beta_\mu + \varepsilon_{\mu,t+1}, \tag{14}$$
$$v_{t+1} = X_t\beta_\lambda + \varepsilon_{\lambda,t+1}, \tag{15}$$

where $R_t - R_{f,t}$ is the annual excess return on the market portfolio and v_t is the realized variance for year t. The annual excess return is calculated as the sum of the monthly excess log-returns, while the realized variance is the sum of the squared monthly excess log-returns. Both sums are calculated within the same year.

Based on the information available at time t and the parameter estimates from (12) and

[14]This approach is taken by French, Schwert, and Stambaugh (1987), Schwert (1989), Whitelaw (1994), Ghysels, Santa-Clara, and Valkanov (2006), Ludvigson and Ng (2007), Lettau and Ludvigson (2010) and Tang and Whitelaw (2011).

[15]French, Schwert, and Stambaugh (1987), Bollerslev, Engle, and Wooldridge (1988) and Glosten, Jagannathan, and Runkle (1993) have applied this approach in the predictability literature.

(13), the conditional Sharpe ratio is calculated as follows:

$$\widehat{SR}_t = \frac{X_t\widehat{\beta}_\mu + \frac{X_t\widehat{\beta}_\lambda}{2}}{\sqrt{X_t\widehat{\beta}_\lambda}}, \tag{16}$$

where $\widehat{\beta}_\mu$ and $\widehat{\beta}_\lambda$ denote the OLS estimates for β_μ and β_λ respectively.[16]

Figure 1 shows the results of a simulated path of annual returns. Each simulation has 100 annual observations of returns. Panel A shows the time series of expected returns calculated from an OLS regression, Panel B shows the conditional variance estimated from an OLS regression and Panel C contains the conditional Sharpe ratio estimates using the fitted values from the conditional mean and conditional volatility from panels A and B. Finally, Panel D displays the time series of annual Sharpe ratios implied by the BY model. These are obtained by evaluating Eq. (11) in $K = 12$.

[Insert Figure 1 about here]

For this specific simulation, the standard deviation of the Sharpe ratio estimates is 3% while the standard deviation of the model Sharpe ratio is 17%. Moreover, the correlation coefficient between the Sharpe ratio estimates based on OLS methods and the Sharpe ratio implied by the model is 7.9%.

The use of artificial data from a fully specified economy is important because it allows the economic reasons that drive the variation in Sharpe ratios to be isolated. In the first case, the variation in the Sharpe ratio is driven by the volatility of consumption growth. In the second case, the volatility of the Sharpe ratio is driven by consumption risk and measurement error caused by the OLS estimation method. In order to verify the robustness of my results, I repeated the previous exercise 100,000 times via Monte Carlo simulations with sample periods of length 100 years. Table 2 reports the median moments implied by the simulations of the BY calibrations of the LRR calibrations. I look at the empirical first and second moments and at the empirical Sharpe ratios constructed via OLS methods and compare them with the median first and second moments as well as the Sharpe ratios implied by the model.

[Insert Table 2 about here]

From Table 2, we learn that the level of expected returns and volatilities implied by the LRR model are well captured by OLS techniques. Indeed, the difference between expected returns

[16] This definition of Sharpe ratio includes the Jensen's adjustment due to log-returns. However, my results are robust if the Sharpe ratio is defined as the ratio of expected returns to conditional volatility.

and the LRR model counterpart is almost indistinguishable. As for the volatility estimates, OLS techniques do a good job in matching the mean level as well as standard deviation. However, there are some differences worth noting. The standard deviation of the risk premia calculated with OLS techniques is 3.01%, while the standard deviation implied by the model is 0.87%. That is, the standard deviation estimated via OLS methods is more than three times the true standard deviation. Moreover, the median of the correlation coefficient between the risk premia and its OLS estimate is 0.52%. A more serious discrepancy is observed in the estimates of the conditional Sharpe ratio. The model implies a median annual Sharpe ratio of 33.33% while the estimates obtained with projection techniques is 26.45%; the standard deviation of the Sharpe ratio calculated regressions is 15.82%, while, the value implied by the model is 3.53%. The correlation between the "true" Sharpe ratios implied by the model and its OLS estimates is 0.39%.

We learn from this simulation exercise that the use of fitted moments as proxies for the unobserved conditional mean and volatility of stock returns has some obvious drawbacks. First, the dynamics of the conditional mean and volatility are determined by the joint conditional distribution of the first-stage predictors. Thus, with any model misspecification the dynamics of the fitted moments would not need to correspond to the dynamics of the true moments. Even when the predictive models for the conditional mean and volatility are well specified, the effect of errors in variables, which are induced by the first stage-regressions, is not trivial to quantify and has an important effect in the Sharpe ratio volatility estimates. Moreover, OLS methods do not account for time-varying observations or time-varying information sets; therefore OLS methods are not robust to structural changes. In that sense, an econometric technique that accounts for such deficiencies may be a good approach for Sharpe ratio estimation and its dynamic behavior. Filtering techniques are able to overcome these issues.

4.2 Filtering and Estimation

Let $y_{t+1} = r_{m,t+1} - r_{f,t+1}$ be the continuously compounded monthly excess return. The time-series dynamics of y_{t+1} is represented by

$$y_{t+1} = \mu_t + \lambda_t \varepsilon_{t+1} \text{ with } \varepsilon_{t+1} \sim \mathcal{N}(0,1), \tag{17}$$

where μ_t and λ_t represent the expected return and conditional volatility. Under the LRR model, these are given by

$$\mu_t = E_0 + E_1 \sigma_t^2, \tag{18}$$

and

$$\lambda_t^2 = D_0 + D_1 \sigma_t^2. \tag{19}$$

According to Eq. (4) , the evolution of σ_t^2 is represented by

$$\begin{aligned} \sigma_{t+1}^2 &= \overline{\sigma}^2 + v\left(\sigma_t^2 - \overline{\sigma}^2\right) + \sigma_w w_{t+1}, \\ w_{t+1} &\sim i.i.d.\ \mathcal{N}(0,1), \end{aligned} \tag{20}$$

and the covariance between the observed market excess return, y_{t+1}, and the innovation to the volatility process, w_{t+1} is given in Eq. (8).

4.2.1 Filtering

Eqs. (17) through (20) make up a state-space model. In the terminology of state-space models, Eq. (17) is the measurement or observation equation and Eq. (20) is the transition or state equation. I assume that σ_t^2 is a latent variable; therefore, both the conditional mean and volatility of market returns are unobservable. I also assume that I am able to observe the full history of realized returns. To draw inferences about the dynamic behavior of σ_t^2 as well as return conditional moments, we need to solve a filtering problem.

The solution to the filtering is the distribution of the latent variable σ_t^2 conditional on the history of observed returns. From Eqs. (9) through (11), we learn that expected returns, volatilities and conditional Sharpe ratios can be estimated based on this conditional distribution, for any holding period. Unfortunately, the filtering problem generated by the LRR model is not standard because of the nonlinearities in the measurement equation as well as the non-zero covariance between the observation and transition equations. As a result, the standard Kalman filter (designed for linear Gaussian state-space models) cannot be used directly in the estimation of the model. I instead rely on nonlinear filtering methods to estimate the distribution of σ_t^2, conditional moments of market excess returns and market Sharpe ratios.

Particle Filters

I estimate the latent process σ_t^2, conditional moments and Sharpe ratios via particle filters. The particle filter is a nonlinear filter which works through Monte Carlo methods. The conditional distribution of the state variables is replaced by an empirical distribution drawn by simulation. This method does not require the explicit computation of Jacobians and Hessians, and captures the conditional distribution of the state variable accurately up to a prespecified accuracy level that depends on the number of simulations chosen by the researcher. To implement the particle filter, it is necessary to specify the state-space model.[17] A brief description of the particle filter and its implementation is given in Appendix D.

[17] Doucet, de Freitas, Gordon, and Smith (2001) and Crisan and Rozovskii (2011) describe in detail the properties of the filter and its practical implementation, and van Binsbergen, Fernandez-Villaverde, Koijen, and Rubio-Ramirez (2012) apply the method to estimate a dynamic stochastic general equilibrium model with a particular focus on the term structure of interest rates.

I test the accuracy of the filtered estimates as follows. First, I simulate a path of annual excess returns according to the calibrations by BY. Given the simulated excess returns, I numerically construct the conditional distribution of the volatility of consumption growth, σ_t^2, using Eqs. (17) to (20) as well as the original calibrations by BY. Once, the conditional distribution of the volatility of consumption growth is obtained, I estimate risk premia, conditional variances and Sharpe ratios according to Eqs. (9) to (11). Figure 2 shows a sample simulation of the volatility of consumption growth, conditional moments and Sharpe ratios along with their filtered counterparts. In panel A, I show a path for the volatility of consumption growth; panels B and C show the simulated expected returns and their volatility; panel D shows the simulated annual Sharpe ratio with its filtered estimates. For this specific simulation, the correlation coefficient between the simulated volatility of consumption growth and its filtered value is 60%. As for the expected returns, volatilities and Sharpe ratios, the simulated values have a correlation coefficient above 64% with their filtered counterparts.

[Insert Figure 2 about here]

A possible concern is that filtering is commonly thought of as a smoothing technique, and therefore, if the state variable to be filtered is too volatile, a filtering technique will reduce such volatility and the unconditional moments of interest may not reflect the true state variable dynamics. However, filtering techniques are robust enough to provide accurate estimates even if the true state variable to be filtered is volatile. This is due to the fact that filtered estimates are conditional expectations of the state variables. To evaluate the unconditional moments, it is necessary to account for this fact; thus, I calculate the unconditional mean and variance of the state variables according to the properties of the law of iterated expectations.[18]

To verify the robustness of my results, the simulation exercise was repeated 1,500 times. For each simulation, I obtain time series of expected returns, volatilities and Sharpe ratios as well as their filtered counterparts. I then calculate the unconditional means, variances and correlations between the simulated and filtered series. The results are reported in Table 3. In general, the filters do a good job of capturing the unconditional moments of expected returns, volatilities and Sharpe ratios. Overall, the moments estimated via filtering methods are precise match the true values in at least two decimal places.

[Insert Table 3 about here]

We learn from these simulation exercises that filtering techniques are better able to cap-

[18]The unconditional variance estimate comes from the following identity, which relates conditional and unconditional variances: $Var[X] = Var[\mathbb{E}[X|Y]] + \mathbb{E}[Var[X|Y]]$.

ture the dynamic behavior of the conditional moments and Sharpe ratios than OLS methods. Nonetheless, these filtered estimates rely on a number of assumptions: the state-space model is well specified; realized returns are a noisy measure of expected returns and the volatility of consumption growth is the only unobservable state variable of the system. However, the researcher has full knowledge of its dynamics as well as the functional forms of expected returns and variances. Finally, I assume that the econometrist has full knowledge of the parameter values and the only problem that she faces is the estimation of conditional moments based on the time series of observed returns. In contrast, OLS methods rely on nonstationarity assumptions of the state variables and predictors. By using OLS methods, we approximate expected returns and variances with a linear projection on a set of exogenous predictors and can potentially face a number of well-known econometric problems, such as omitted variables or misspecification. For clarity of exposition, I collect all parameters that define the state-space model into a single parameter vector θ. Each parameter vector characterizes a model; hence, conditional distributions and filtered state variables. As a result, an estimation problem needs to be solved and will be explained in detail as follows.

4.2.2 Estimation

The previous results were obtained by assuming that the set of parameter values is known. This assumption is quite unrealistic, because in reality the researcher is uncertain about the true parameter values. A natural way to approach this issue is by estimating the vector of parameters from the observed data. A common technique for nonlinear dynamic models is quasi-maximum likelihood estimation (QMLE).[19] This approach is discussed in Winschel and Krätzig (2010) and Romero (2012). Details about its implementation are described in Appendix E.

[Insert Figure 3 about here]

I conduct a simulation exercise to better identify the effect of parameter estimation within the filtering exercise. First, I simulate a time series of excess returns from the LRR model, and then I estimate the parameter values via quasi-maximum likelihood methods using the state-space representation implied by the LRR model. The parameter estimates are then used in the filtering estimation procedure.

A sample simulation is illustrated in Figure 3. Panels 3(a) to 3(c) compare conditional Sharpe ratio estimates with their true values. Panel 3(a) shows the empirical estimate obtained via OLS methods, Panel 3(b) displays Sharpe ratio estimates calculated with filtering methods

[19] Some examples are Campbell, Sunderam, and Viceira (2012); van Binsbergen and Koijen (2011) and Calvet, Fisher, and Wu (2010).

by assuming that the true parameter values are known and Panel 3(c) contains the filtered Sharpe ratios using the parameter estimates obtained via quasi-maximum likelihood methods.

For this specific simulation, the time-series average of the model-implied Sharpe ratio is 33%, while the average Sharpe ratio estimates obtained with filtering techniques are 34% and 36%, where the first is obtained by assuming that the true parameter values are known and the second is obtained with the parameter estimates from observed returns. Finally, the average Sharpe ratio obtained with OLS methods is 25%. An explanation for this difference is the model misspecification that is generated from running OLS regressions for expected returns and volatility calculations on a set of predetermined variables. The volatility estimates obtained via filtering methods are 5% and 6%. My results are similar to the estimates obtained from the true simulated data, which is 4%. In contrast, OLS methods deliver a Sharpe ratio volatility estimate of 15%. This exercise illustrates the effect of parameter estimation on the volatility of Sharpe ratios. I show evidence that filtering methods deliver Sharpe ratio volatility estimates consistent with the true model implied values, even if parameter values have to be estimated.

4.3 Other Models

Recent consumption-based asset pricing models have made substantial progress in explaining many asset pricing puzzles across various markets. Even though such models are not often used to study Sharpe ratios or their volatility, they do make theoretical predictions about their values. In standard asset pricing models, the market Sharpe ratio is constant (Sharpe (1964); Lintner (1965); Lucas (1978) and Breeden (1979)) or has negligible variation (Mehra and Prescott (1985) and Weil (1989)). Habit formation preferences can help to capture the counter-cyclicality of the risk premia (Constantinides (1990); Abel (1990) and Campbell and Cochrane (1999)) as well as other features of macro-economic outcomes over the business cycle (Jermann (2010)). Bansal and Yaron (2004) combine the preferences of Epstein and Zin (1989) with stochastic volatility of consumption growth and generate time variation in the conditional volatility of the SDF.

Other studies have found different channels for time variation in risk premia, such as differences in risk aversion (Chan and Kogan (2002), Gomes and Michaelides (2008), Bhamra and Uppal (2010)); rare disasters (Rietz (1988); Barro (2006, 2009) and Wachter (2012)); incomplete markets (Constantinides and Duffie (1996), Gârleanu and Panageas (2011)); participation constraints, (Basak and Cuoco (1998), Guvenen (2009), Chien, Cole, and Lustig (2012)); investment shocks (Papanikolaou (2011)) and heterogeneity in the frequency of shocks to fundamentals (Calvet and Fisher (2007)). A brief summary of the aforementioned models and their asset pricing implications are shown in Table 4.

[Insert Table 4 about here]

The asset pricing implications of the models shown in Table 4 provide a general idea of the model-implied variability of Sharpe ratios. Indeed, this variability could be used as a metric to better assess the performance of a model. For example, an asset pricing model with constant Sharpe ratios would fail in explaining the observed variation in empirical Sharpe ratios. On the other hand, a model that predicts highly volatile Sharpe ratios may exceed the true variability observed in the data. Therefore, the variance of Sharpe ratios can be used as a metric to better assess theoretical asset pricing models. This metric would be in the spirit of the entropy measure recently proposed by Backus, Chernov, and Zin (2012) and studied in Martin (2012).

As a robustness check of the variability generated by OLS methods to calculate Sharpe ratios, I performed a second simulation exercise based on the external habit formation model by Campbell and Cochrane (1999). A brief description of the model and a brief overview of the results are presented below.

4.4 External Habit Formation Model

In the external habit formation model of Campbell and Cochrane (1999), the consumption dynamics are the same as in the standard Lucas model; that is consumption growth rates are assumed to be independent and identically distributed. Furthermore, the agent is assumed to have external habit formation preferences. The habit level is assumed to be a slow-moving and heteroscedastic process. The heteroscedasticity of the habit process, the sensitivity function, can be chosen so that the real interest rate in the model is constant or linear in the habit. Further details can be found in Appendix F.

I use the same calibrated monthly parameters as those in Campbell and Cochrane (1999) to simulate returns from the model and compute annual expected returns, volatilities and Sharpe ratios using standard OLS techniques. I compare these results with the numerical values implied by the model. Consistent with my previous results, I find that the Sharpe ratios based on standard OLS methods are at least twice more volatile than the model-implied variability.

[Insert Figure 4 about here]

The results are plotted in Figure 4. Panel A displays the Sharpe ratio estimates based on OLS methods, while Panel B displays the values of the true Sharpe ratios. Clearly, the Sharpe ratio estimates based on OLS methods are more volatile than the values implied by the habit formation model.

5 Sharpe Ratios in Reduced Form Models

The use of data simulated by means of theoretical models helps to better identify the economic reasons that drive the time-series variation. An alternative form of analyzing returns is via reduced form models, which are statistical models that do not impose any economic structure. These models aim to better describe historical data. Moreover, they do not rely on arbitrary predictors and are not subject to the effects of errors in variables or misspecification.

In this section, I introduce the nonlinear latent VAR representation proposed in Brandt and Kang (2004), in which the first and second conditional moments are considered latent variables identified from the history of returns. In this setup, the Sharpe ratio and its dynamics are obtained endogenously as the ratio of the conditional moments of excess returns. The framework is general enough and can be extended to a setup that includes flexible correlation structures and exogenous predictors.

5.1 Brandt and Kang (2004)

Let y_t be the continuously compounded excess returns with time-series dynamics represented by

$$y_t = \mu_{t-1} + \lambda_{t-1}\varepsilon_t \text{ with } \varepsilon_t \sim \mathcal{N}(0,1) \tag{21}$$

where $\mu_{t-1}$and λ_{t-1} represent the conditional volatility of the excess returns. In addition, it is assumed that the conditional mean and volatility are unobservable and that they follow a first order VAR process in logs:

$$\begin{bmatrix} \ln \mu_t \\ \ln \lambda_t \end{bmatrix} = d + A \begin{bmatrix} \ln \mu_{t-1} \\ \ln \lambda_{t-1} \end{bmatrix} + \eta_t \text{ with } \eta_t \equiv \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \end{bmatrix} \sim \mathcal{N}(0, \Sigma), \tag{22}$$

where

$$\begin{aligned} d &= \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}, \; A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ and} \\ \Sigma &= \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \text{ with } b_{12} = b_{21} = \rho\sqrt{b_{11}b_{22}}. \end{aligned} \tag{23}$$

Following Hamilton (1994), if the VAR is stationary, the unconditional moments for the mean and volatility are given by

$$\mathbb{E}\begin{bmatrix} \ln \mu_t \\ \ln \lambda_t \end{bmatrix} = (I - A)^{-1} d \tag{24}$$

and

$$\text{vec}\left(\text{cov}\begin{bmatrix}\ln\mu_t\\ \ln\lambda_t\end{bmatrix}\right) = (I - (A\otimes A))^{-1}\,\text{vec}\,(\Sigma) \tag{25}$$

where $\otimes$ represents the Kronecker product.

The key elements of the return dynamics presented Eq.(22) are the transition matrix A and the correlation coefficient ρ. The diagonal elements of A capture the persistence of the conditional moments, and the off-diagonal elements reflect the intertemporal feedback between the conditional volatility and the conditional mean. The correlation coefficient ρ captures the contemporaneous correlation between the innovations to the conditional moments. This parameter is of considerable importance since it captures the risk-return trade-off.[20]

The model in Eq. (22) generalizes the permanent and temporary components of Fama and French (1988b) and the standard stochastic volatility model. The equation for the conditional mean is

$$\ln\mu_t = d_1 + a_{11}\ln\mu_{t-1} + a_{12}\ln\lambda_{t-1} + \eta_{1t}, \quad \text{where} \quad \eta_{1t}\sim\mathcal{N}\left(0, b_{11}\right). \tag{26}$$

If $a_{12} = 0$, the dynamics of the conditional mean is similar to the temporary component as in Lamoureux and Zhou (1996). Now, the equation that describes the conditional volatility is

$$\ln\lambda_t = d_2 + a_{21}\ln\mu_{t-1} + a_{22}\ln\lambda_{t-1} + \eta_{2t}, \quad \text{where} \quad \eta_{2t}\sim\mathcal{N}\left(0, b_{22}\right), \tag{27}$$

and corresponds to the standard stochastic volatility model; in particular if $a_{21} = 0$, Eq.(27) is the standard stochastic volatility model as in Andersen and Sørensen (1996) and Kim, Shephard, and Chib (1998). Finally, we learn from Eq. (25) that the unconditional variance is determined by the variance-covariance matrix Σ and the matrix A. For identification purposes, I assume four different specifications for the transition matrix A. First, in model A, I consider the case in which the conditional mean and volatility evolve as in Eqs. (26) and (27) . Models B and C consider $a_{12} = 0$ and $a_{21} = 0$, respectively, allowing for the model of permanent and temporary component in the first case, and the standard stochastic volatility model in the second case. Finally, model D considers the case in which $a_{12} = a_{21} = 0$.

An interesting property is the nonnegativity of expected returns and volatilities. This nonnegativity of the conditional mean guarantees a positive risk premium, as suggested in Merton

[20] Most asset pricing models predict a positive relationship between the market's risk premium and conditional volatility (Merton (1973)). However, the empirical evidence on the sign of the risk-return relation is inconclusive. Indeed, some studies find a positive relation (e.g. Scruggs (1998), Ghysels, Santa-Clara, and Valkanov (2005), Lundblad (2007), Ludvigson and Ng (2007) and Pastor, Sinha, and Swaminathan (2008)), but others find a negative relation (e.g. Campbell (1987), Glosten, Jagannathan, and Runkle (1993), Harvey (2001), Lettau and Ludvigson (2010) and Brandt and Kang (2004)). Others have shown through theoretical studies that the intertemporal mean-variance relationship may not be positive or negative (e.g. Whitelaw (2000) and Ang and Liu (2007)).

(1980), and has been used by Bekaert and Harvey (1995) and Jacquier, Johannes, and Polson (2007), among others. The log-normality specification for the volatility is consistent with Andersen, Bollerslev, Diebold, and Ebens (2001) and Andersen, Bollerslev, Diebold, and Labys (2003), which show that the log-volatility process can be well approximated by a normal distribution, and with Taylor (2008), who proposes to model the logarithm of volatility as an AR(1) process.

5.2 Implied Sharpe Ratio

The latent VAR implies a conditional Sharpe ratio of the form

$$SR_t = \frac{\mu_t + \lambda_t^2/2}{\lambda_t}, \tag{28}$$

where μ_t and λ_t are the conditional mean and volatility of stock returns.[21] It follows that the Sharpe ratio is stochastic if the innovations that affect both the numerator and denominator in Eq.(28) are stochastic and do not cancel each other out. Moreover, the Sharpe ratio is time-varying due to the mean reversion of the two conditional moments. The distribution of the Sharpe ratio corresponds to the sum of two correlated log-normal distributions, which is not standard.

5.3 The data

I study quarterly returns on the value-weighted index market portfolio from CRSP. Excess returns are calculated by subtracting the quarterly yield on a three-month T-bill from the corresponding stock return. I use this yield instead of the monthly yield due to the idiosyncratic variation documented in Duffee (1996). The predictors are the CRSP dividend-price ratio (d-p), calculated as the log-ratio of the CRSP dividends to the price level of the CRSP value-weighted stock index; the relative bill rate ($RREL$), which is the difference between the three-month treasury bill and its four-quarter moving average; the term spread (TRM), the difference between the ten-year treasury bond yield and the three-month treasury bill; the default spread (DEF), the difference between the BAA corporate bond rate and the AAA corporate bond rate and the consumption-wealth ratio proxy (cay).[22] The $RREL$, TRM and DEF are obtained from the Federal Reserve statistical release. Data on the dividend-price ratio is taken from

[21] The squared term in the numerator comes from a Jensen's adjustment for log-returns.

[22] These predictors are used in the predictability literature. See Goyal and Welch (2008) and Lettau and Ludvigson (2010) for details.

CRSP and the time series of *cay* is taken from Sidney Ludvigson's website.[23] All data are quarterly from the period April 1953 to December 2011.

5.4 Parameter Estimates

The model in Eqs.(21) and (22) is nonlinear since the first equation is nonlinear in the state-variables. The parameters are estimated via quasi-maximum likelihood methods and are shown in Table 5. The first column corresponds to the estimates of model A, the second column shows the estimates for model B, and the third and fourth columns contain the parameter estimates for models C and D, respectively. Given the frequency of returns, expected returns are persistent since the estimates for a_{11} range from 0.6727 to 0.7204.[24] The conditional volatility is more persistent than the conditional mean, for all model specifications.

[Insert Table 5 about here]

The parameter estimates of the models A through D show evidence of a strong and negative risk-return trade-off, measured by the correlation between the innovations to the conditional mean and the volatility of excess returns. The estimates range from -0.1760 to -0.7995, for both the constrained and unconstrained representations, and are statistically significant. This finding is consistent with the negative risk-return relationship found in Brandt and Kang (2004), Campbell and Hentschel (1992) and Campbell (1987). The negative sign of the correlation coefficient between the conditional mean and the volatility of returns amplifies the variability of the Sharpe ratio, whereas a positive correlation between expected returns and volatility makes Sharpe ratios less variable than its mean or even constant.

The estimates show that there is more variation in the mean than in the conditional volatility, since the conditional variance of the innovations to the conditional mean, b_{11}, differs substantially from that of the innovation to the conditional volatility, b_{22}. The off-diagonal elements of the transition matrix A are significant. However, the values for a_{21} are similar across models, while the values for a_{12} differ. The differences in signs of a_{12} and a_{21} are consistent with the results of Whitelaw (1994) and Brandt and Kang (2004), which state that the cross-autocorrelations between the conditional mean and volatility offset each other through time.

[23] I thank Sidney Ludvigson for making the time series data for cay available. This variable is calculated in a quarterly basis. Source: http://www.econ.nyu.edu/user/ludvigsons/

[24] These values correspond to a monthly persistence of roughly 0.87 to 0.89.

5.5 Expected Returns, Volatilities and Sharpe Ratios

Given the parameter estimates in Table 5, I estimate expected returns, volatilities and Sharpe ratios via particle filtering. The left column of Figure 5 presents the filtered estimates of quarterly expected returns (first row), volatility (second row) and Sharpe ratios (third row). Each plot also shows in vertical bars the NBER recession dates. It is clear that the conditional mean, volatility and Sharpe ratio are time varying. The quarterly mean has a standard deviation of less than 1% and it varies from 1% in the third quarter of 1974 to 3% in the last quarter of 2003. The quarterly volatility has a standard deviation of 2% and ranges from 7.3% to 11.6%. Expected returns revert more quickly to their unconditional mean than do conditional volatilities, and this is consistent with the estimates of a_{11} and a_{22}.

Quarterly Sharpe ratios are displayed in the last row of the first column. The Sharpe ratio rises from the peak to the trough of the recession dates in the sample, and is consistent with the empirical results documented by Lustig and Verdelhan (2012), Tang and Whitelaw (2011) and Lettau and Ludvigson (2010). This countercyclical variation of the Sharpe ratio is also consistent with the habit formation models (Constantinides (1990) and Campbell and Cochrane (1999)). Intuitively, at the peak of the business cycle, consumers enjoy consumption levels far above their "habits." As a result, a low Sharpe ratio, or low reward per unit of risk, is required for a consumer to invest in the stock index at the peak of the cycle, in contrast to the trough of a cycle, where consumption levels are closer to those of the habits, which makes consumers more relative risk averse. For an investor willing to invest in the trough of the cycle, the rewards per unit of risk or Sharpe ratios should be substantially high.

[Insert Figure 5 about here]

5.5.1 OLS estimates

I calculate expected returns, volatilities and Sharpe ratios based on OLS techniques for comparison purposes. Table 6 presents the estimates from OLS regressions of quarterly realized excess returns and excess log-returns from the first quarter of 1953 to the last quarter of 2011. The results are generally consistent with those reported in the predictability literature. There is no substantial difference between the regression estimates obtained by using excess returns and excess log-returns. At a one-quarter horizon, cay and $RREL$ show a consistent predictive power for excess returns. Indeed, cay alone explains 3% of next quarter's total variability. Adding the lagged value of excess returns, cay, $d-p$, $RREL$ and TRM explains 8% of the quarter's variation in the next quarter's excess return. The $R-$squared of 8% for log-returns is lower than the values reported in previous studies, but the sample, which includes the 2007-2008 financial crisis, may account for this result.

[Insert Table 6 about here]

The results for the volatility equation are presented in Table 7. In this representation, the lagged volatility, $d - p$, TRM and DEF are significant. The positive serial correlation in realized volatility reflects the autoregressive conditional heteroskedasticity of quarterly returns. The lagged value of volatility alone explains 37% of next the quarter's excess return volatility. Lagged volatility values, cay, $d - p$, $RREL$, and TRM explain altogether 41%. Finally, the high $R-$squared value of 43% in the full volatility equation reflects that realized volatility is much more predictable than excess returns.

[Insert Table 7 about here]

Empirical moments of expected returns, volatilities and Sharpe ratios are displayed in Table 8. The first set of estimates is calculated based on OLS regressions of quarterly realized log-returns for the CRSP value-weighted index on lagged explanatory variables. The second set of estimates is based on the reduced form model by Brandt and Kang (2004), in which the conditional mean and volatility of stock returns are treated as latent variables. This representation guarantees positive values for the conditional mean and volatility of excess returns.

[Insert Table 8 about here]

As in the simulation exercises described in section 4, I find differences worth noting among the estimates. First, expected returns and volatilities calculated via OLS have a quarterly standard deviation of 2%, while the standard deviation of the filtered estimates is 1%. Filtered volatilities are higher, on average, than the ones obtained with OLS methods and are more autocorrelated. The autocorrelation of expected returns obtained with OLS methods is 81%, in contrast with the one estimated from the filtered series, which is less than 59%. This is not surprising, since the regressors used for its estimation are highly persistent. The autocorrelation of the filtered estimates is consistent with the estimated value of a_{11}.

As for the Sharpe ratio estimates, there are major differences worth noting. First, the average quarterly Sharpe ratio estimated via filtering is 26% while the OLS estimate is 30%. As for the standard deviation estimates, the difference is quite substantial. For the OLS estimates, the standard deviation is 42%, which is similar to the 45% reported by Lettau and Ludvigson (2010), while the standard deviation of the filtered Sharpe ratio ranges from 5%. A potential explanation of this difference is the use of standard OLS techniques for its estimation. The reduced form representations do not rely on predetermined conditioning variables to estimate conditional moments; the state variables are identified from the history of returns whereas standard OLS techniques generate fitted moments from a set of predictive regressions as proxies for the unobservable conditional mean and volatility. The fitted moments depend on the joint distribution of the predictors; therefore, any model misspecification would generate fitted moments that do not correspond to the true dynamics of the conditional mean and volatility, and as a result, the Sharpe ratio dynamics. Another important issue is that the ratio of the fitted moments does not adjust for the correlation between expected returns and volatility of stock returns, whereas filtering techniques do.

5.5.2 Alternative Reduced Forms

For comparison purposes, I also analyze an unconstrained version of the representation of Brandt and Kang (2004). The excess returns have time-series dynamics of the form

$$y_t = \mu_{t-1} + \lambda_{t-1}\varepsilon_t \text{ with } \varepsilon_t \sim \mathcal{N}(0,1), \tag{29}$$

where $\mu_{t-1}$and λ_{t-1} represent the conditional volatility of the excess returns. In addition, it is assumed that the conditional mean and the log-volatility are unobservable and that they follow a first order VAR process of the form

$$\begin{bmatrix} \mu_t \\ \ln\lambda_t \end{bmatrix} = d + A \begin{bmatrix} \mu_{t-1} \\ \ln\lambda_{t-1} \end{bmatrix} + \eta_t \text{ with } \eta_t \sim \mathcal{N}(0,\Sigma), \tag{30}$$

where d, A and Σ are defined as in Eq. (23). The main difference between the model representation by Brandt and Kang (2004) and Eqs.(29) and (30) is that expected returns can potentially be negative, as in Lamoureux and Zhou (1996). As in the previous model, I consider four model specifications for the matrix A. The covariance matrix, Σ, has the same structure as Eq. (23). The sign of the correlation coefficient between the conditional mean and the volatility of excess returns has the same sign as the correlation between the conditional mean and the log-volatility.[25]

[Insert Table 9 about here]

Quasi-maximum likelihood estimates of the model with an unconstrained risk premia are shown in Table 9. Under all model specifications, the parameter estimates, are similar to the estimates of the first model. An important difference is that the estimates of the off-diagonal

[25]From Stein's lemma, we have that the conditional covariance between excess returns and the conditional volatility is $cov_{t-1}(\mu_t, \lambda_t) = \mathbb{E}_t[\lambda_t] \cdot cov_{t-1}(\mu_t, \ln\lambda_t)$. Thus, the sign of the correlation coefficient between the conditional mean and the volatility of stock returns is the same as the conditional correlation of the conditional mean and the log-volatility of returns.

elements a_{12} and a_{21} are negative, although a_{12} is not statistically significant.

The right column of Figure 5 displays the filtered estimates of conditional moments and Sharpe ratios for the model with an unconstrained risk premia. The main difference between the constrained and unconstrained representations is that expected returns can take negative values; indeed, expected return estimates took negative values for six quarters of the whole sample. Qualitatively, both latent VAR models show similar dynamic behavior; in fact, the correlation coefficient between the implied Sharpe ratio estimates is 70%.

5.5.3 Exogenous Predictors

The main advantage of the latent VAR approach by Brandt and Kang (2004) is that it allows the study of the dynamics of the conditional mean, volatility and Sharpe ratios without relying on exogenous predictors. At the same time, useful information is potentially discarded, since any correlation structure between predictors and conditional moments is ignored. As a robustness check, I estimate an extended version of the model in which each moment is a function of the same exogenous predictors used in the predictive regressions (cay, $d-p$, $RREL$, and TRM). The model specification is given by

$$y_t = \mu_{t-1} + \lambda_{t-1}\varepsilon_t \text{ with } \varepsilon_t \sim \mathcal{N}(0,1), \tag{31}$$

where

$$\begin{bmatrix} \ln \mu_t \\ \ln \sigma_t \end{bmatrix} = d + A \begin{bmatrix} \ln \mu_{t-1} \\ \ln \sigma_{t-1} \end{bmatrix} + Cx_{t-1} + \eta_t, \text{ with } \eta_t \sim \mathcal{N}\left(0, \Sigma\right), \tag{32}$$

where x_t denotes the de-meaned vector of predictors observed at date t.

Table 10 reports the parameter estimates of the extended model D and also replicates for comparison the results of model D. The estimates of A and Σ are similar across the two models. When I add the exogenous predictors, all the parameter estimates of the base model decrease in magnitude, which means that the exogenous predictors help explain some of the variation in the moments that was left unexplained. Finally, the correlation between the innovations to the mean and volatility decreases in magnitude from -0.7995 to -0.4523, both significant.

[Insert Table 10 about here]

In the mean equation of the extended model, the coefficients of cay, $d-p$, TRM (c_{11}, c_{12} and c_{14}) are positive and the coefficients of $RREL$ and DEF (c_{13} and c_{15}) are negative. In the volatility equation, all coefficients are negative, except for one, DEF. The signs of the coefficients are all consistent with the results of the predictive regressions. However, it is important to note that these results are not directly comparable to standard predictive regressions, since these coefficients correspond to regressions with the conditional moments as dependent variables.

5.5.4 Comparison

Empirical moments of the different Sharpe ratio estimates are displayed in Table 11. The first, second and third sets of Sharpe ratio estimates are based on the latent VAR approach by Brandt and Kang (2004). The first representation is based on Eqs. (21) and (22), while the second representation guarantees a positive volatility only and is based on Eqs. (29) and (30). The third representation is an extended version of the first model in which the conditional moments are positive functions of exogenous predictors and is represented in Eqs.(31) and (32). Finally, the last set of Sharpe ratio estimates is based on the conditional moments calculated from OLS regressions of log-returns on lagged explanatory variables.

[Insert Table 11 about here]

The results from Table 11 show that the average quarterly Sharpe ratios under the first two models are 25% and 26%, respectively. The third model implies a quarterly Sharpe ratio of 31%, while the estimates obtained from OLS methods have a quarterly Sharpe ratio of 30%. The difference is caused by the set of exogenous predictors included within the estimation procedure. The first set of results represents the Sharpe ratio estimates based on the set of observed returns, while the third and fourth correspond to Sharpe ratio estimates using the history of returns and the set of exogenous predictors. The parameter estimates used in the filtering calculations depend on the data used in the estimation process. In the first two models, the parameter and filtered estimates depend on the time series of excess returns, while the last two models depend on the same series of returns as well as on the set of exogenous predictors.

As for the Sharpe ratio volatility implied by the models, there are some differences worth noting. The first two models imply a volatility of 5% and 10%, respectively. The difference is due to the model representation. The first model considers a positive risk premia and the second does not. Since the second model allows for negative Sharpe ratios, there is more variability. As for the third representation, the variability is 25%, which is mainly driven by the inclusion of a set of exogenous predictors that affect the conditional mean and volatility of excess returns. None of these representations deliver a Sharpe ratio volatility of 42% as OLS methods do. The main driver of this difference is the use of conditioning information within the estimation process. In the first two cases, the model representations as well as the history of returns determine the variability of the Sharpe ratio. In contrast, the set of exogenous predictors that are included in the estimation process of the third model and fourth model determines a higher variability of the Sharpe ratio estimates.

6 Implications for Portfolio Choice

In this section, I discuss a standard model from the portfolio-choice literature and its relation to the market Sharpe ratio.

6.1 Portfolio Optimization: One Risky Asset

I consider an investor with mean-variance preferences that optimizes the tradeoff between the mean and the variance of portfolio returns. Two assets are available to an investor at time t. One is risk free, with return $R_{f,t+1}$ from time t to time $t+1$, and the other is risky. The risky asset has simple return R_{t+1} from time t to time $t+1$ with conditional mean $\mathbb{E}_t\left[R_{t+1}\right]$ and conditional variance σ_t^2.

The investor allocates a share α_t of her portfolio into the risky asset. Then the portfolio return is

$$\begin{aligned} R_{p,t+1} &= \alpha_t R_{t+1} + \left(1-\alpha_t\right) R_{f,t+1} \\ &= R_{f,t+1} + \alpha_t \left(R_{t+1} - R_{f,t+1}\right). \end{aligned}$$

The mean portfolio return is $\mathbb{E}_t\left[R_{p,t+1}\right] = R_{f,t+1} + \alpha_t\left(\mathbb{E}_t\left[R_{t+1}\right] - R_{f,t+1}\right)$, while the variance of the portfolio is $\sigma_{pt}^2 = \alpha_t^2 \sigma_t^2$.

If the investor has mean-variance preferences, then she trades off between the mean and variance in a linear fashion. In other words, she maximizes a linear combination of mean and variance with a positive weight on mean and a negative weight on variance,

$$\max_{\alpha_t}\left(\mathbb{E}_t\left[R_{p,t+1}\right] - \frac{\gamma}{2}\sigma_{pt}^2\right).$$

The solution to this optimization problem is

$$\alpha_t = \frac{\mathbb{E}_t\left[R_{t+1}\right] - R_{f,t+1}}{\gamma \sigma_t^2}. \tag{33}$$

The optimal weight for the stock index coincides with the so-called myopic demand and can be interpreted as the product of the relative risk tolerance (i.e., inverse of the relative risk aversion) and the market Sharpe ratio normalized by the volatility of the market returns; that is

$$\alpha_t = \frac{SR_t}{\gamma \sigma_t}. \tag{34}$$

We learn from Eq.(34) that for investors with mean-variance preferences, the optimal allocation in the market portfolio is determined by three elements: the Sharpe ratio of the market portfolio, the conditional volatility of the market portfolio and the risk aversion parameter. Moreover, the variability of portfolio weights is determined by the variability of Sharpe ratios and as well as by the standard deviation of the market portfolio.

Campbell and Viceira (2002) derive a similar expression by assuming an investor with power utility and that the return on an investor's portfolio is lognormal, with the slight difference that the optimal weight in Eq. (33) is adjusted by half the variance of the risky asset; that is,

$$\alpha_t = \frac{\mathbb{E}_t\left[r_{t+1}\right] - r_{f,t+1} + \sigma_t^2/2}{\gamma\sigma_t^2}. \tag{35}$$

Now I implement the model following the standard plug-in approach; that is, I solve the optimization problem assuming that the mean and variance of returns are known. Once the problem is solved, I replace the moments with their estimates obtained via regression or filtering techniques. For simplicity, I assume that the investor ignores estimation risk while making an investment decision.

[Insert Figure 6 about here]

Figure 6 shows the optimal allocations in Eq. (35) using OLS and filtering methods to estimate conditional moments assuming a risk aversion parameter $\gamma = 5$. Clearly, the portfolio weights constructed via OLS methods are more volatile than the ones obtained with the filtered moments. Indeed, the average portfolio weight under the OLS model specification is 1.27 with a standard deviation of 2.13, in contrast to the portfolio weight computed with filtering methods, which is on average 56% with a standard deviation of 12%. Finally, the correlation between the two weights is 15%.

These results have practical implications for portfolio allocation, especially for an investor who faces proportional costs by trading the optimal portfolio of an investor with mean-variance preferences.[26] As the optimal weight is proportional to the market Sharpe ratio, the percentage of wealth traded in each period will depend upon the volatility of the market Sharpe ratio. It is clear that upward-biased estimates of the Sharpe ratio volatility would imply excessive portfolio rebalancing, and therefore more transaction costs.

[26]This fact was first noted by De Miguel, Garlappi, and Uppal (2009) for performance evaluation.

7 Conclusions

In this paper I examine whether estimates of the variability of the Sharpe ratio may be biased due to limitations of the empirical methodology used in its estimation. I provide evidence that measurement error in estimated Sharpe ratios helps to explain the Sharpe ratio volatility puzzle. I further show that this measurement error is caused by the use of standard OLS methods to estimate the ratio. The empirical question I address is important because many studies have used the results implied by OLS methods to calibrate the volatility of the market Sharpe ratio.

Based on simulated data from standard asset pricing models, I document that OLS methods produce Sharpe ratio volatility that is larger than the ratio's true variability. Using the OLS approach may also provide conditional moment estimates that do not necessarily correspond to their true values.

Once I have documented the upward bias in the Sharpe ratio's variability generated by OLS methods, I consider if using improved empirical methodologies may better reflect the ratio's true variability. To accomplish this goal, I propose filtering methods as a way to better assess this variation. These techniques explicitly allow for the estimation of time-varying moments by aggregating the entire history of realized returns in a parsimonious way. Moreover, filtering is flexible enough to be used with large information sets without relying on exogenous predictors, while being robust to structural breaks. I also show that filtering techniques better reflect the true variation of Sharpe ratios even when parameter values need to be estimated.

Motivated by the simulation results, I use real data on excess stock returns to compare the Sharpe ratio volatility estimates produced by OLS and filtering methods. I find that filtering methods deliver Sharpe ratio variability estimates that are much smaller than the Sharpe ratio variability estimates implied from OLS methods. The difference in results from the two methodologies arises due to the use of conditioning information within the filtering estimation process.

My findings have significant implications for asset pricing. For example, in a portfolio allocation setting, the optimal portfolio weight is proportional to the market Sharpe ratio. Thus, upward biased estimates of the Sharpe ratio volatility would imply excessive portfolio rebalancing, and therefore more transaction costs.

References

Abel, A. B., 1990. Asset Prices under Habit Formation and Catching up with the Joneses. The American Economic Review 80(2), pp. 38–42.

Andersen, T. G., T. Bollerslev, F. X. Diebold, H. Ebens, 2001. The distribution of realized stock return volatility. Journal of Financial Economics 61(1), 43 – 76.

Andersen, T. G., T. Bollerslev, F. X. Diebold, P. Labys, 2003. Modeling and Forecasting Realized Volatility. Econometrica 71(2), pp. 579–625.

Andersen, T. G., B. E. Sørensen, 1996. GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study. Journal of Business and Economic Statistics 14(3), pp. 328–352.

Ang, A., G. Bekaert, 2007. Stock Return Predictability: Is it There?. The Review of Financial Studies 20(3), pp. 651–707.

Ang, A., J. Liu, 2007. Risk, return, and dividends. Journal of Financial Economics 85(1), 1 – 38.

Ang, A., M. Piazzesi, 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. Journal of Monetary Economics 50(4), 745 – 787.

Back, K., 2010. Asset Pricing and Portfolio Choice Theory. Oxford University Press, .

Backus, D., M. Chernov, S. E. Zin, 2012. Sources of Entropy in Representative Agent Models. Working Paper New York University.

Bansal, R., D. Kiku, A. Yaron, 2012a. An Empirical Evaluation of the Long-Run Risks Model for Asset Prices. Critical Finance Review 1, 183–221.

——— , 2012b. Risks For the Long Run: Estimation with Time Aggregation. Working Paper Duke University.

Bansal, R., A. Yaron, 2004. Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles.. Journal of Finance 59(4), 1481 – 1509.

Barro, R. J., 2006. Rare Disasters and Asset Markets in the Twentieth Century. The Quarterly Journal of Economics 121(3), pp. 823–866.

——— , 2009. Rare Disasters, Asset Prices, and Welfare Costs. The American Economic Review 99(1), pp. 243–264.

Basak, S., D. Cuoco, 1998. An Equilibrium Model with Restricted Stock Market Participation. The Review of Financial Studies 11(2), pp. 309–341.

Beeler, J., J. Y. Campbell, 2012. The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment. Critical Finance Review 1, 141–182.

Bekaert, G., C. R. Harvey, 1995. Time-Varying World Market Integration. The Journal of Finance 50(2), pp. 403–444.

Bhamra, H. S., R. Uppal, 2010. Asset Prices with Heterogeneity in Preferences and Beliefs. Working Paper Edhec Business School and University of British Columbia.

Bollerslev, T., R. Engle, J. Wooldridge, 1988. A capital asset pricing model with time-varying covariances. The Journal of Political Economy pp. 116–131.

Brandt, M. W., Q. Kang, 2004. On the relationship between the conditional mean and volatility of stock returns: A latent VAR approach. Journal of Financial Economics 72(2), 217 – 257.

Breeden, D. T., 1979. An intertemporal asset pricing model with stochastic consumption and investment opportunities. Journal of Financial Economics 7(3), 265 – 296.

Breen, W., L. R. Glosten, R. Jagannathan, 1989. Economic Significance of Predictable Variations in Stock Index Returns. The Journal of Finance 44(5), pp. 1177–1189.

Calvet, L., A. J. Fisher, L. Wu, 2010. Dimension-Invariant Dynamic Term Structures. Working Paper: University of British Columbia.

Calvet, L. E., A. J. Fisher, 2007. Multifrequency news and stock returns. Journal of Financial Economics 86(1), 178 – 212.

Campbell, J., R. Shiller, 1988. The dividend-price ratio and expectations of future dividends and discount factors. Review of Financial Studies 1(3), 195–228.

Campbell, J., A. Sunderam, L. Viceira, 2012. Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds. Working Paper Harvard University.

Campbell, J., L. Viceira, 2002. Strategic Asset Allocation: Portfolio Choice for Long-Term Investors. Oxford University Press, USA, .

Campbell, J. Y., 1987. Stock returns and the term structure. Journal of Financial Economics 18(2), 373 – 399.

——— , 1991. A Variance Decomposition for Stock Returns. The Economic Journal 101(405), pp. 157–179.

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

Campbell, J. Y., L. Hentschel, 1992. No news is good news: An asymmetric model of changing volatility in stock returns. Journal of Financial Economics 31(3), 281 – 318.

Chan, Y. L., L. Kogan, 2002. Catching Up with the Joneses: Heterogeneous Preferences and the Dynamics of Asset Prices.. Journal of Political Economy 110(6), 1255 – 1285.

Chien, Y., H. Cole, H. Lustig, 2012. Is the Volatility of the Market Price of Risk Due to Intermittent Portfolio Rebalancing?. American Economic Review 102(6), 2859–96.

Conrad, J., G. Kaul, 1988. Time-Variation in Expected Returns. The Journal of Business 61(4), pp. 409–425.

Constantinides, G. M., 1990. Habit Formation: A Resolution of the Equity Premium Puzzle. Journal of Political Economy pp. 519 – 543.

Constantinides, G. M., D. Duffie, 1996. Asset Pricing with Heterogeneous Consumers. Journal of Political Economy 104(2), pp. 219–240.

Crisan, D., B. Rozovskii, 2011. The Oxford handbook of Nonlinear Filtering. Oxford University Press, .

Dangl, T., M. Halling, 2012. Predictive regressions with time-varying coefficients. Journal of Financial Economics 106(1), 157 – 181.

De Miguel, V., L. Garlappi, R. Uppal, 2009. Optimal versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy?. The Review of Financial Studies 22(5), pp. 1915–1953.

Doucet, A., N. de Freitas, N. Gordon, A. Smith, 2001. Sequential Monte Carlo Methods in Practice. Springer, .

Duffee, G. R., 1996. Idiosyncratic Variation of Treasury Bill Yields. The Journal of Finance 51(2), pp. 527–551.

Epstein, L. G., S. E. Zin, 1989. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. Econometrica 57(4), pp. 937–969.

Fama, E. F., K. R. French, 1988a. Dividend yields and expected stock returns. Journal of Financial Economics 22(1), 3 – 25.

——— , 1988b. Permanent and Temporary Components of Stock Prices. Journal of Political Economy 96(2), pp. 246–273.

Fama, E. F., G. Schwert, 1977. Asset returns and inflation. Journal of Financial Economics 5(2), 115 – 146.

Ferson, W. E., S. Sarkissian, T. T. Simin, 2003. Spurious Regressions in Financial Economics?. The Journal of Finance 58(4), pp. 1393–1413.

French, K. R., G. Schwert, R. F. Stambaugh, 1987. Expected stock returns and volatility. Journal of Financial Economics 19(1), 3 – 29.

Gallant, A., H. White, 1988. A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models. Basil Blackwell New York, .

Gârleanu, N., S. Panageas, 2011. Young, Old, Conservative and Bold: The Implications of Heterogeneity and Finite Lives for Asset Pricing. Working Paper University of Chicago.

Ghysels, E., P. Santa-Clara, R. Valkanov, 2005. There is a risk-return trade-off after all. Journal of Financial Economics 76(3), 509 – 548.

——— , 2006. Predicting volatility: getting the most out of return data sampled at different frequencies. Journal of Econometrics 131, 59 – 95.

Glosten, L. R., R. Jagannathan, D. E. Runkle, 1993. On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. The Journal of Finance 48(5).

Gomes, F., A. Michaelides, 2008. Asset Pricing with Limited Risk Sharing and Heterogeneous Agents. Review of Financial Studies 21(1), 415–448.

Goyal, A., I. Welch, 2008. A Comprehensive Look at the Empirical Performance of Equity Premium Prediction. The Review of Financial Studies 21(4), pp. 1455–1508.

Guvenen, F., 2009. A Parsimonious Macroeconomic Model for Asset Pricing. Econometrica 77(6), pp. 1711–1750.

Hamilton, J. D., 1994. Time Series Analysis. Princeton University Press, .

Hansen, L. P., R. Jagannathan, 1991. Implications of Security Market Data for Models of Dynamic Economies. Journal of Political Economy 99(2), pp. 225–262.

Hansen, L. P., T. J. Sargent, 1983. Aggregation Over Time and the Inverse Optimal Predictor Problem for Adaptive Expectations in Continuous Time. International Economic Review 24(1), pp. 1–20.

Harrison, J., D. Kreps, 1979. Martingales and arbitrage in multiperiod securities markets. Journal of Economic Theory 20(3), 381–408.

Harvey, C. R., 2001. The Specification of Conditional Expectations. Journal of Empirical Finance 8(5), 573 – 637.

Heaton, J., 1995. An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specifications. Econometrica 63(3), pp. 681–717.

Hodrick, R. J., 1992. Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement. The Review of Financial Studies 5(3), pp. 357–386.

Jacquier, E., M. Johannes, N. Polson, 2007. MCMC maximum likelihood for latent state models. Journal of Econometrics 137(2), 615 – 640.

Jermann, U. J., 2010. The equity premium implied by production. Journal of Financial Economics 98(2), 279 – 296.

Johnson, T., J. Lee, 2012. Systematic Volatility of Unpriced Earnings Shocks. University of Illinois, Urbana-Champaign Working Paper.

Kandel, S., R. F. Stambaugh, 1990. Expectations and Volatility of Consumption and Asset Returns. The Review of Financial Studies 3(2).

Kim, C.-J., C. R. Nelson, 1999. State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications vol. 1. the MIT Press, .

Kim, S., N. Shephard, S. Chib, 1998. Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. The Review of Economic Studies 65(3), 361–393.

Lamoureux, C., G. Zhou, 1996. Temporary components of stock returns: What do the data tell us?. Review of Financial Studies 9(4), 1033–1059.

Lettau, M., S. Ludvigson, 2001a. Consumption, Aggregate Wealth, and Expected Stock Returns. The Journal of Finance 56(3), 815–849.

——— , 2001b. Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time Varying. Journal of Political Economy 109(6), pp. 1238–1287.

——— , 2010. Measuring and Modeling Variation in the Risk-Return Trade-Off. Handbook of Financial Econometrics.

Lettau, M., H. Uhlig, 2002. The Sharpe Ratio and preferences: A Parametric Approach. Macroeconomic Dynamics 6(02), 242–265.

Lettau, M., S. Van Nieuwerburgh, 2008. Reconciling the Return Predictability Evidence. Review of Financial Studies 21(4), 1607–1652.

Lewellen, J., 1999. The time-series relations among expected return, risk, and book-to-market. Journal of Financial Economics 54(1), 5 – 43.

Lintner, J., 1965. Security Prices, Risk, and Maximal Gains From Diversification. The Journal of Finance 20(4), pp. 587–615.

Lucas, Robert E., J., 1978. Asset prices in an exchange economy. Econometrica pp. 1429–1445.

Ludvigson, S. C., S. Ng, 2007. The Empirical Risk-Return Relation: A Factor Analysis Approach. Journal of Financial Economics 83(1), 171 – 222.

Lundblad, C., 2007. The Risk-return Trade-off in the long run: 1836-2003. Journal of Financial Economics 85(1), 123 – 150.

Lustig, H. N., A. Verdelhan, 2012. Business Cycle Variation in the Risk-Return Trade-Off. Journal of Monetary Economics (Forthcoming).

Martin, I., 2012. Simple Variance Swaps. Working Paper Stanford University.

Mehra, R., E. C. Prescott, 1985. The equity premium: A puzzle. Journal of Monetary Economics 15(2), 145 – 161.

Menzly, L., T. Santos, P. Veronesi, 2004. Understanding Predictability. Journal of Political Economy 112(1), pp. 1–47.

Merton, R. C., 1969. Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case. The Review of Economics and Statistics 51(3), pp. 247–257.

Merton, R. C., 1971. Optimum consumption and portfolio rules in a continuous-time model. Journal of Economic Theory 3(4), 373 – 413.

Merton, R. C., 1973. An Intertemporal Capital Asset Pricing Model. Econometrica 41(5), pp. 867–887.

——— , 1980. On estimating the expected return on the market: An exploratory investigation. Journal of Financial Economics 8(4), 323 – 361.

Nielsen, L. T., M. Vassalou, 2004. Sharpe Ratios and Alphas in Continuous Time. The Journal of Financial and Quantitative Analysis 39(1), pp. 103–114.

Papanikolaou, D., 2011. Investment Shocks and Asset Prices. Journal of Political Economy 119(4), pp. 639–685.

Pastor, L., M. Sinha, B. Swaminathan, 2008. Estimating the Intertemporal Risk-Return Tradeoff Using the Implied Cost of Capital. The Journal of Finance 63(6), pp. 2859–2897.

Pástor, L., R. F. Stambaugh, 2009. Predictive Systems: Living with Imperfect Predictors.. Journal of Finance 64(4), 1583 – 1628.

Rietz, T. A., 1988. The equity risk premium a solution. Journal of Monetary Economics 22(1), 117 – 131.

Romero, A., 2012. Filtering via Taylor Series Approximations. Working Paper: University of British Columbia.

Rytchkov, O., 2012. Filtering out Expected Dividends and Expected Returns. Quarterly Journal of Finance (Forthcoming).

Schwert, G. W., 1989. Why Does Stock Market Volatility Change Over Time?. The Journal of Finance 44(5), pp. 1115–1153.

Scruggs, J. T., 1998. Resolving the Puzzling Intertemporal Relation between the Market Risk Premium and Conditional Market Variance: A Two-Factor Approach. The Journal of Finance 53(2), pp. 575–603.

Sharpe, W. F., 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. The Journal of Finance 19(3), pp. 425–442.

Tang, Y., R. F. Whitelaw, 2011. Time-Varying Sharpe Ratios and Market Timing. Quarterly Journal of Finance 1(3), 465 – 493.

Taylor, S., 2008. Modelling Financial Time Series. World Scientific, .

van Binsbergen, J., J. Fernandez-Villaverde, R. Koijen, J. F. Rubio-Ramirez, 2012. The Term Structure of Interest Rates in a DSGE Model with Recursive Preferences. Journal of Monetary Economics: Forthcoming.

van Binsbergen, J., R. Koijen, 2011. Likelihood-based Estimation of Exactly-Solved Present-Value Models. Working Paper: University of Chicago.

van Binsbergen, J. H., R. S. J. Koijen, 2010. Predictive Regressions: A Present-Value Approach.. Journal of Finance 65(4), 1439 – 1471.

Vuolteenaho, T., 2000. Understanding the Aggregate Book-to-Market Ratio. Unpublished paper, Harvard University.

Wachter, J. A., 2005. Solving models with external habit. Finance Research Letters 2(4), 210 – 226.

———, 2012. Can time-varying risk of rare disasters explain aggregate stock market volatility?. Journal of Finance, Forthcoming.

Weil, P., 1989. The equity premium puzzle and the risk-free rate puzzle. Journal of Monetary Economics 24(3), 401 – 421.

Whitelaw, R., 2000. Stock market risk and return: An equilibrium approach. Review of Financial Studies 13(3), 521–547.

Whitelaw, R. F., 1994. Time Variations and Covariations in the Expectation and Volatility of Stock Market Returns. The Journal of Finance 49(2), pp. 515–541.

Winschel, V., M. Krätzig, 2010. Solving, Estimating, and Selecting Nonlinear Dynamic Models Without the Curse of Dimensionality. Econometrica 78(2), 803–821.

Appendices

A Sharpe Ratios in Asset Pricing

Harrison and Kreps (1979) show that the absence of arbitrage implies the existence of a stochastic discount factor (SDF) or pricing kernel, denoted by M_t, that prices all assets in the economy.[27] More specifically, the conditional expectation of the product of the stochastic discount factor and the gross asset return (R_t) must be equal to one; that is,

$$\mathbb{E}_t \left[M_{t+1} R_{t+1} \right] = 1, \tag{36}$$

where the conditional expectation is based on the information available at time t. Since Eq.(36) holds for any asset in the economy, it must hold for the one-period risk-free interest rate $\left(R_{ft+1} \right)$; consequently, the risk-free rate can be written as the inverse of the conditional expectation of the stochastic discount factor,

$$R_{f,t+1} = \frac{1}{\mathbb{E}_t \left[M_{t+1} \right]}. \tag{37}$$

Another implication of Eq.(36) is that the expected risk premium on any asset is given by the negative of the product of the risk-free rate and the conditional covariance of the stochastic discount factor with the gross return; that is,

$$\mathbb{E}_t \left[R_{t+1} - R_{ft+1} \right] = -R_{ft+1} Cov_t \left(R_{t+1}, M_{t+1} \right). \tag{38}$$

The conditional Sharpe ratio of an asset at time t, denoted by SR_t, is defined as the ratio of the conditional mean excess return to the conditional standard deviation of its return, that is

$$SR_t = \frac{\mathbb{E}_t \left[R_{t+1} - R_{ft+1} \right]}{\sigma_t \left[R_{t+1} - R_{ft+1} \right]}. \tag{39}$$

Then, the conditional Sharpe ratio is proportional to the risk-free rate, the volatility of the pricing kernel and the correlation between the pricing kernel and the return; that is,

$$\frac{\mathbb{E}_t \left[R_{t+1} - R_{ft+1} \right]}{\sigma_t \left[R_{t+1} - R_{ft+1} \right]} = -R_{ft+1} \sigma_t \left[M_{t+1} \right] Corr_t \left[R_{t+1}, M_{t+1} \right], \tag{40}$$

where σ_t and $Corr_t$ are the standard deviation and correlation; respectively, both conditional on information at time t. The conditional Sharpe ratio of any asset in the economy is time varying as long as the risk-free rate varies or the pricing kernel is conditionally heteroskedastic; that is, if $\sigma_t \left[M_{t+1} \right]$ changes over time or if the correlation between the stock market return and

[27] See Back (2010) for a detailed and concise explanation.

the stochastic discount factor is time varying.

Now, the maximum of the right-hand side of Eq.(40) over all returns defines a lower bound for the standard deviation of any stochastic discount factor depending on the risk-free rate. Since the correlation coefficient is between -1 and 1, we have

$$\frac{\mathbb{E}_t\left[R_{t+1}\right]-R_{ft+1}}{\sigma_t\left[R_{t+1}-R_{ft+1}\right]} \leq R_{ft+1}\sigma_t\left[M_{t+1}\right] \equiv SR_t^{\max}, \quad \text{for all assets.} \tag{41}$$

Eq. (41) implies the Hansen and Jagannathan (1991) bound, which is an upper bound to the absolute value of the conditional Sharpe ratios of any asset in the economy, given a specific discount factor. The maximum Sharpe ratio, $SR_t^{\max}$, is achieved if there exists an asset in the economy which is perfectly negatively correlated with M_{t+1}. In general, the Sharpe ratios of all the assets in the economy are bounded by the right-hand side of Eq.(41) but when markets are complete there exists an asset that achieves the upper bound, and the inequality becomes an equality.[28] Moreover, a very volatile SDF is necessary to understand high Sharpe ratios. The conditional variance of the SDF can be thought of as the variance of the investor's marginal utility of consumption in the next period.[29] Therefore, from Eq.(40) we learn that each model has an implication for the dynamic behavior of the market Sharpe ratio, since each model implies a functional form for the SDF.

The use of log-returns is a common practice in the empirical literature. A standard approximation of the Sharpe ratio based on continuously compounded returns is given by

$$SR_t = \frac{\mathbb{E}_t[r_{t+1}] - r_{f,t+1} + \frac{\sigma_t^2[r_{t+1}]}{2}}{\sigma_t\left[r_{t+1}\right]}, \tag{42}$$

where r_{t+1} denotes the continuously compounded return of an asset, $r_{f,t+1}$ denotes the continuously compounded risk-free rate and $\sigma_t\left[r_{t+1}\right]$ denotes the standard deviation of the return of an asset. The numerator in Eq. (42) includes the Jensen adjustment for log-returns.[30]

[28] A detailed discussion of this result is shown in Lettau and Uhlig (2002).

[29] Hansen and Jagannathan (1991) provide a comprehensive analysis of this bound, allowing for many risky assets and no risk-free asset, and derive implications of the positivity of the stochastic discount factor.

[30] The difference between Eqs. (41) and (42) is almost negligible for short return horizons, as reported by Brandt and Kang (2004). Nielsen and Vassalou (2004) analyze the difference between discrete and continuously compounded versions of Sharpe ratios and propose this adjustment for performance evaluation. Campbell and Viceira (2002) discuss in detail this approximation in a portfolio optimization framework.

B The Solution to the Long-Run Risks Model

This section provides solutions for the consumption and dividend claim for the Bansal, Kiku, and Yaron (2012a) endowment process,

$$\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_t + \sigma_t \eta_{t+1} \\
x_{t+1} &= \rho x_t + \varphi_e \sigma_t e_{t+1} \\
\sigma_{t+1}^2 &= \overline{\sigma}^2 + v\left(\sigma_t^2 - \overline{\sigma}^2\right) + \sigma_w w_{t+1} \\
\Delta d_{t+1} &= \mu_d + \phi x_t + \varphi \sigma_t u_{t+1} + \pi \sigma_t \eta_{t+1} \\
w_{t+1}, e_{t+1}, u_{t+1}, \eta_{t+1} &\sim i.i.d.\ \mathcal{N}(0,1).
\end{aligned} \tag{43}$$

The Euler equation for this economy is

$$E_t\left[\exp\left(\theta \log \delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta - 1) r_{a,t+1} + r_{i,t+1}\right)\right] = 1, \tag{44}$$

where $r_{a,t+1}$ is the log-return on the consumption claim and $r_{i,t+1}$ is the log-return on any asset. All returns are given by the approximation from Campbell and Shiller (1988), $r_{i,t+1} = \kappa_{0,i} + \kappa_{1,i} z_{i,t+1} - z_{i,t} + \Delta d_{i,t+1}$.

Let $Y_t' = [1, x_t, \sigma_t^2]$ denote a vector of state variables and the log price-consumption ratio be given by $z_t = A'Y_t$, where A denotes a vector of coefficients $A' = [A_0, A_1, A_2]$. In general, for any other asset i, define the coefficients in the same manner: $A_i' = [A_{0,i}, A_{1,i}, A_{2,i}]$. This section calculates the price of the consumption claim as well as the dividend claim $z_{t,m} = A_m' Y_t$. The coefficients that characterize z_t and $z_{t,m}$ are obtained by the method of undetermined coefficients and by the fact that the Euler equation must hold for all values of Y_t'.

The risk premium on any asset is

$$\begin{aligned}
E_t\left[r_{i,t+1} - r_{f,t}\right] + \frac{1}{2} Var_t\left[r_{i,t+1}\right] &= -Cov_t\left(m_{t+1}, r_{i,t+1}\right) \\
&= \sum_{j=n,e,w} \lambda_j \beta_{i,j} \sigma_{j,t}^2,
\end{aligned} \tag{45}$$

where $\beta_{i,j}$ is the beta and $\sigma_{j,t}^2$ the volatility of the j^{th} risk source, and the λ_j represents the price of each risk source.

B.1 Consumption Claim

The risk premium for the consumption claim is

$$E_t\left[r_{a,t+1}-r_{f,t}\right]+\frac{1}{2}Var_t\left[r_{a,t+1}\right]=\lambda_n\beta_{a,n}\sigma_t^2+\lambda_e\beta_{a,e}\sigma_t^2+\lambda_w\beta_{a,w}\sigma_w^2, \tag{46}$$

where $\beta_{a,n}=1$, $\beta_{a,e}=\kappa_1A_1\varphi_e$ and $\beta_{a,w}=\kappa_1A_2$. The conditional variance of the consumption claim is equal to

$$Var_t\left[r_{a,t+1}\right]=\left(\beta_{a,n}^2+\beta_{a,e}^2\right)\sigma_t^2+\beta_{a,w}^2\sigma_w^2. \tag{47}$$

The coefficients A' for the log price-consumption ratio z_t are

$$\begin{aligned}
A_0 &= \frac{\ln\delta+\mu_c\left(1-\frac{1}{\psi}\right)+\kappa_0+\beta_{a,w}\overline{\sigma}^2\left(1-v\right)+\frac{1}{2}\theta\beta_{a,w}^2\sigma_w^2}{\left(1-\kappa_1\right)},\\
A_1 &= \frac{1-\frac{1}{\psi}}{1-\kappa_1\rho},\\
A_2 &= \frac{\frac{\theta}{2}\left[\left(1-\frac{1}{\psi}\right)^2+\beta_{a,e}^2\right]}{\left(1-\kappa_1 v_1\right)}.
\end{aligned} \tag{48}$$

B.2 Dividend Claim

The innovation to the market return, denoted by $r_{m,t+1}-E_t\left(r_{m,t+1}\right)$, is

$$r_{m,t+1}-E_t\left(r_{m,t+1}\right)=\varphi\sigma_tu_{t+1}+\beta_{m,\eta}\sigma_t\eta_{t+1}+\beta_{m,e}\sigma_te_{t+1}+\beta_{m,w}\sigma_ww_{t+1}, \tag{49}$$

where $\beta_{m,\eta}=\pi$, $\beta_{m,e}=\kappa_{1,m}A_{1,m}\varphi_e$ and $\beta_{m,w}=\kappa_{1,m}A_{2,m}$, which implies that the risk premium on the dividend claim is

$$E_t\left[r_{m,t+1}-r_{f,t}\right]+\frac{1}{2}Var_t\left[r_{m,t+1}\right]=\lambda_\eta\beta_{m,\eta}\sigma_t^2+\lambda_e\beta_{m,e}\sigma_t^2+\lambda_w\beta_{m,w}\sigma_w^2. \tag{50}$$

Finally, the coefficients A'_m for the log price-dividend ratio are as follows

$$\begin{aligned}
A_{0,m} &= \frac{\left[\begin{array}{c}\theta\ln\delta+\mu_c\left(\theta-\frac{\theta}{\psi}-1\right)-\lambda_w\overline{\sigma}^2\left(1-v\right)+\left(\theta-1\right)\left[\kappa_0+A_0\left(\kappa_1-1\right)\right]\\ \kappa_{0,m}+\beta_{m,w}\overline{\sigma}^2\left(1-v\right)+\mu_d+\frac{1}{2}\left[\beta_{m,w}-\lambda_w\right]^2\sigma_w^2\end{array}\right]}{\left(1-\kappa_{1,m}\right)},\\
A_{1,m} &= \frac{\phi-\frac{1}{\psi}}{1-\kappa_{1,m}\rho},\\
A_{2,m} &= \frac{\left(1-\theta\right)A_2\left(1-\kappa_1v_1\right)+\frac{1}{2}\left[\left(\pi-\lambda_n\right)^2+\left[\beta_{m,e}-\lambda_e\right]^2+\varphi^2\right]}{\left(1-\kappa_{1,m}v\right)}.
\end{aligned} \tag{51}$$

B.3 Risk-Free Interest Rate

The risk-free rate is derived from the Euler equation applied to the risk-less asset:

$$\begin{aligned} r_{f,t+1} &= -\log E_t\left[\exp\left(m_{t+1}\right)\right] \\ &= -\theta\ln\delta + \frac{\theta}{\psi}E_t\left[\Delta c_{t+1}\right] + (1-\theta)\,E_t\left[r_{a,t+1}\right] \\ &\quad -\frac{1}{2}Var_t\left[\frac{\theta}{\psi}\Delta c_{t+1} + (1-\theta)\,r_{a,t+1}\right]. \end{aligned} \tag{52}$$

By subtracting $(1-\theta)\,r_{f,t+1}$ from both sides of Eq.(52) and if $\theta \neq 0$, then we can divide by θ, yielding to an expression for the risk-free rate

$$r_{f,t+1} = -\ln\delta + \frac{1}{\psi}E_t\left[\Delta c_{t+1}\right] + \frac{(1-\theta)}{\theta}E_t\left[r_{a,t+1} - r_{f,t+1}\right] - \frac{1}{2\theta}Var_t\left(m_{t+1}\right), \tag{53}$$

where

$$\begin{aligned} E_t\left[\Delta c_{t+1}\right] &= \mu_c + x_t \\ E_t\left[r_{a,t+1} - r_{f,t+1}\right] &= \left(\lambda_n\beta_{a,n} + \lambda_e\beta_{a,e} - \frac{\left(\beta_{a,n}^2 + \beta_{a,e}^2\right)}{2}\right)\sigma_t^2 + \left(\lambda_w\beta_{a,w} - \frac{\beta_{a,w}^2}{2}\right)\sigma_w^2. \\ Var_t\left(m_{t+1}\right) &= \left(\lambda_\eta^2 + \lambda_e^2\right)\sigma_t^2 + \lambda_w^2\sigma_w^2, \end{aligned}$$

as a result

$$r_{f,t+1} = A_{0,f} + A_{1,f}x_t + A_{2,f}\sigma_t^2, \tag{54}$$

where

$$\begin{aligned} A_{0,f} &= -\ln\delta + \frac{\mu_c}{\psi} + \frac{(1-\theta)}{\theta}\left(\lambda_w\beta_{a,w} - \frac{\beta_{a,w}^2}{2}\right)\sigma_w^2 - \frac{\lambda_w^2\sigma_w^2}{2\theta}, \\ A_{1,f} &= \frac{1}{\psi}, \\ A_{2,f} &= \frac{(1-\theta)}{\theta}\left(\lambda_n\beta_{a,n} + \lambda_e\beta_{a,e} - \frac{\left(\beta_{a,n}^2 + \beta_{a,e}^2\right)}{2}\right) - \frac{\left(\lambda_\eta^2 + \lambda_e^2\right)}{2\theta}. \end{aligned}$$

B.4 Return on the Market Portfolio

Recall that the rate of return on the market portfolio is

$$r_{m,t+1} = \kappa_{0,m} + \kappa_{1,m}z_{m,t+1} - z_{m,t} + \Delta d_{t+1}, \tag{55}$$

where the dynamics are characterized by the following equations

$$\begin{aligned} z_{m,t} &= A_{0,m} + A_{1,m} x_t + A_{2,m} \sigma_t^2 \qquad (56) \\ x_{t+1} &= \rho x_t + \varphi_e \sigma_t e_{t+1} \\ \sigma_{t+1}^2 &= \overline{\sigma}^2 + v\left(\sigma_t^2 - \overline{\sigma}^2\right) + \sigma_w w_{t+1}, \\ \Delta d_{t+1} &= \mu_d + \phi x_t + \varphi \sigma_t u_{t+1} + \pi \sigma_t \eta_{t+1}. \end{aligned}$$

Now, since each of the components of the market return follows a normal distribution, then the market return has a normal distribution with conditional mean

$$\begin{aligned} E_t\left[r_{m,t+1}\right] &= \kappa_{0,m} + \kappa_{1,m} E_t\left[z_{m,t+1}\right] - z_{m,t} + E_t\left[\Delta d_{t+1}\right] \\ &= \kappa_{0,m} + \kappa_{1,m}\left(A_{0,m} + A_{1,m}\rho x_t + A_{2,m}\left(\overline{\sigma}^2 + v\left(\sigma_t^2 - \overline{\sigma}^2\right)\right)\right) \\ &\quad - A_{0,m} - A_{1,m} x_t - A_{2,m}\sigma_t^2 + \mu_d + \phi x_t \\ &= \kappa_{0,m} + \left(\kappa_{1,m} - 1\right) A_{0,m} + \kappa_{1,m} A_{2,m}\left(1 - v\right)\overline{\sigma}^2 + \mu_d \\ &\quad + \left[A_{1,m}\left(\kappa_{1,m}\rho - 1\right) + \phi\right] x_t + A_{2,m}\left(\kappa_{1,m} v - 1\right)\sigma_t^2 \\ &= B_0 + B_1 x_t + B_2 \sigma_t^2, \qquad (57) \end{aligned}$$

where

$$\begin{aligned} B_0 &= \kappa_{0,m} + \left(\kappa_{1,m} - 1\right) A_{0,m} + \kappa_{1,m} A_{2,m}\left(1 - v\right)\overline{\sigma}^2 + \mu_d \\ B_1 &= \phi - A_{1,m}\left(1 - \kappa_{1,m}\rho\right) = \frac{1}{\psi}, \\ B_2 &= A_{2,m}\left(\kappa_{1,m} v - 1\right). \end{aligned}$$

Now, the variance of the market portfolio is given by

$$\begin{aligned} Var_t\left[r_{m,t+1}\right] &= \kappa_{1,m}^2 Var_t\left[z_{m,t+1}\right] + Var_t\left[\Delta d_{t+1}\right] \\ &= \kappa_{1,m}^2\left(A_{1,m}^2 \varphi_e^2 \sigma_t^2 + A_{2,m}^2 \sigma_w^2\right) + \left(\varphi^2 + \pi^2\right)\sigma_t^2 \\ &= D_0 + D_1 \sigma_t^2, \end{aligned}$$

where $D_0 = \left(\kappa_{1,m} A_{2,m} \sigma_w\right)^2$ and $D_1 = \kappa_{1,m}^2 A_{1,m}^2 \varphi_e^2 + \varphi^2 + \pi^2$.

B.5 Linearization Parameters

For any asset, the linearization parameters are determined endogenously by the following system of equations as discussed in Bansal, Kiku, and Yaron (2012a) and Beeler and Campbell (2012):

$$
\begin{aligned}
\overline{z_i} &= A_{0,i}\left(\overline{z_i}\right) + A_{2,i}\left(\overline{z_i}\right)\sigma^2, \\
\kappa_{1,i} &= \frac{\exp\left(\overline{z_i}\right)}{1+\exp\left(\overline{z_i}\right)}, \\
\kappa_{0,i} &= \ln\left(1+\exp\left(\overline{z_i}\right)\right) - \kappa_{1,i}\overline{z_i}.
\end{aligned} \tag{58}
$$

The solution is determined numerically by iteration until reaching a fixed point of $\overline{z_i}$. The dependence of $A_{0,i}$ and $A_{2,i}$ on the linearization parameters has been discussed in previous sections.

C Excess Returns Conditional Moments Implied by the Long-Run Risks Model

C.1 Expected Returns

The expected excess returns for period k are defined as

$$r_{m,t+k+1} - r_{f,t+k+1},\ \ k = 0, 1, 2, ...$$

Now, the conditional excess risk premium for any period has a closed-form expression given by

$$E_t \left[r_{m,t+k+1} - r_{f,t+k+1}\right] = E_{0,k+1} + E_{1,k+1}\sigma_t^2, \tag{59}$$

where

$$\begin{aligned} E_{0,k+1} &= E_0 + E_1 \left(1 - v^k\right) \overline{\sigma}^2, \\ E_{1,k+1} &= E_1 v^k,\ \ k = 0, 1, 2, ..., \\ E_0 &= B_0 - A_{0,f}, \\ E_1 &= B_2 - A_{2,f}. \end{aligned}$$

C.2 Variance of Excess Returns

Now, for any time period k, the conditional variance of the future excess returns is given by

$$Var_t \left[r_{m,t+k+1} - r_{f,t+k+1}\right],\ \text{for}\ k = 0, 1, 2, ...$$

Its closed-form expression is given by

$$Var_t \left[r_{m,t+k+1} - r_{f,t+k+1}\right] = D_{0,k+1} + D_{1,k+1}\sigma_t^2, \tag{60}$$

where

$$\begin{aligned} D_{0,k+1} &= D_0 + D_1 \left(1 - v^k\right) \overline{\sigma}^2 + E_1^2 \sigma_w^2 \frac{1 - v^{2k}}{1 - v^2} \\ D_{1,k+1} &= v^k D_1, \\ D_0 &= \left(\kappa_{1,m} A_{2,m} \sigma_m\right)^2, \\ D_1 &= \left(\kappa_{1,m} A_{1,m} \varphi_e\right)^2 + \varphi^2 + \pi^2. \end{aligned}$$

C.2.1 Autocovariance of Excess Returns

Now, let $0 \leq k < p$. Then the autocovariance of excess returns is

$$cov_t \left(r_{m,t+k+1} - r_{f,t+k+1}, r_{m,t+p+1} - r_{f,t+p+1}\right) = E_1^2 \sigma_w^2 v^{p-k} \left(\frac{1 - v^{2k}}{1 - v^2}\right) + E_1 \kappa_{1,m} A_{2,m} \sigma_w^2 v^{p-k-1}.$$

C.3 Aggregate Excess Returns

Now, the expected excess returns during K periods are given by the sum of the one-period excess returns,

$$\sum_{k=1}^{K} \left(r_{m,t+k} - r_{f,t+k}\right).$$

Its conditional mean is

$$E_t \left[\sum_{k=1}^{K} r_{m,t+k} - r_{f,t+k}\right] = \mathbf{E}_{0,K} + \mathbf{E}_{1,K} \sigma_t^2,$$

where

$$\begin{aligned} \mathbf{E}_{0,K} &= K E_0 + E_1 \overline{\sigma}^2 \left[K - \frac{\left(1 - v^K\right)}{(1 - v)}\right], \\ \mathbf{E}_{1,K} &= E_1 \frac{\left(1 - v^K\right)}{(1 - v)}. \end{aligned}$$

C.4 Variance of Aggregate Excess Returns

The conditional variance is

$$Var_t \left[\sum_{k=1}^{K} r_{m,t+k} - r_{f,t+k}\right] = \mathbf{D}_{0,K} + \mathbf{D}_{1,K} \sigma_t^2.$$

where

$$\begin{aligned}
\mathbf{D}_{0,K} &= KD_0 + KD_1\overline{\sigma}^2 - D_1\overline{\sigma}^2\frac{\left(1-v^K\right)}{(1-v)} \\
&\quad +\frac{E_1^2\sigma_w^2}{1-v^2}\left[K-\frac{\left(1-v^{2K}\right)}{(1-v^2)}\right] \\
&\quad +2\left[K-\frac{1-v^K}{1-v}\right]\left[\begin{array}{c}\frac{E_1^2\sigma_w^2}{(1-v^2)}\left[\frac{v}{1-v}\right] \\ +E_1\kappa_{1,m}A_{2,m}\sigma_w^2\left[\frac{1}{1-v}\right]\end{array}\right] \\
&\quad -\frac{2E_1^2\sigma_w^2}{(1-v^2)}\frac{v^3\left(1-v^{K-1}\right)\left(1-v^K\right)}{(1-v)^2(1+v)}, \\
\mathbf{D}_{1,K} &= \frac{\left(1-v^K\right)}{(1-v)}D_1.
\end{aligned}$$

D The Particle Filter

The particle filter is a sequential Monte Carlo algorithm, that is, a sampling method for approximating a distribution that makes use of its temporal structure. A "particle representation" of distributions is used. In particular, I will be concerned about the distribution $f(S_t|y_t, y_{t-1}, ..., y_0)$ where S_t is the unobserved state at time t, and $y_t, y_{t-1}, ..., y_0$ is the sequence of observations from time 0 to time t.

For the standard Kalman filters, this distribution $f(S_t|y_t, y_{t-1}, ..., y_0)$ follows a multivariate normal distribution due to the linearity in the measurement and its distribution $(f(S_t|S_{t-1}))$ and state equations. The particle filter is more general, and is based on a less restrictive framework. The only requirements are that the conditional distributions should be tractable, in the sense that I only need to be able to evaluate these distributions, and draw samples from $f(S_t|S_{t-1})$ or $f(S_t|S_{t-1}, y_t)$.

In a few cases, such as the linear and Gaussian filters, without restrictive linear Gaussian assumptions regarding the transition and sensor models, $f(S_t|y_t, y_{t-1}, ..., y_0)$ cannot be written in a simple form. Instead, I will represent it using a collection of N weighted samples or particles, $\left\{S_t^{(i)}, \pi_t^{(i)}\right\}_{i=1}^N$ where $\pi_t^{(i)}$ is the weight of particle $S_t^{(i)}$. A particle representation of this density is given by

$$f(S_t|y_t, y_{t-1}, ..., y_0) \simeq \sum_i \pi_t^{(i)} \delta\left(S_t - S_{t-1}^{(i)}\right).$$

Consider the integral that needs to be evaluated at each filtering step; then

$$f(S_t|y_t, y_{t-1}, ..., y_0) = \alpha f(y_t|S_t) \int f(S_{t-1}|y_{t-1}, y_{t-2}, ..., y_0) f(S_t|S_{t-1}) \, dS_{t-1},$$

which is a recursive definition to compute the filtered distribution $f(S_t|y_t, y_{t-1}, ..., y_0)$ given the distribution $f(S_{t-1}|y_{t-1}, y_{t-2}, ..., y_0)$.

With a particle representation for $f(S_{t-1}|y_{t-1}, y_{t-2}, ..., y_0)$, the recursive equation can be approximated as

$$f(S_t|y_t, y_{t-1}, ..., y_0) \simeq \alpha f(y_t|S_t) \sum_i \pi_{t-1}^{(i)} f\left(S_t|S_{t-1}^{(i)}\right).$$

The only element to be decided is the set of particles for representing the distribution $f(S_t|y_t, y_{t-1}, ..., y_0)$. One choice is to use importance sampling. The particle filter can be thought as an application of the importance sampler on this distribution. The technique of importance sampling is used for generating random samples of a distribution $f(x)$. Suppose $f(x)$ is a density from which it is difficult to draw samples, but it is easy to evaluate $g(x_i)$ for some particular x_i. Then, an approximation to $f(x)$ is given by

$$f(x) \simeq \sum_i \pi^{(i)} \delta\left(x - x^{(i)}\right),$$

where

$$\pi^{(i)} = \frac{f(x)}{g(x_i)}.$$

Note that any distribution $g(\cdot)$, known as a proposal distribution, can be used here, in particular, a uniform sampling of state space x. However, with such a uniform sampling strategy, most samples will be wasted, having small $\pi^{(i)}$ values. Instead, I use a more direct proposal distribution, my approximation to $f(S_t | y_t, y_{t-1}, ..., y_0)$. With this proposal distribution, the weights $\pi^{(i)}$ end up being relatively simple due to cancellation.

Concretely, the particle filter consists of the following steps:

1. Draw N samples $S_t^{(j)}$ from the proposal distribution $q(S_t)$

$$S_t^{(j)} \sim q(S_t) = \sum_i \pi_{t-1}^{(i)} f\left(S_t | S_{t-1}^{(i)}\right)$$

   by drawing a random variable from a uniform distribution defined on $(0,1)$, choosing the corresponding particle i and then sampling from $f\left(S_t | S_{t-1}^{(i)}\right)$. This transition model is typically a linear Gaussian model, but any model from which samples can easily be drawn will suffice.

2. Set the weight $\pi_t^{(j)}$ as the likelihood

$$\pi_t^{(j)} = f\left(y_t | S_t^{(j)}\right)$$

   .The samples $\left\{S_t^{(j)}\right\}$ are drawn from $f(S_t | y_t, y_{t-1}, ..., y_0)$. Re-weighting them in this fashion accounts for evidence y_t.

3. Normalize the weights $\left\{\pi_t^{(j)}\right\}$:

$$\pi_t^{(j)} := \frac{\pi_t^{(j)}}{\sum_k \pi_t^{(k)}}$$

Note also that there is an optimal proposal distribution, which is not the one used here. The optimal proposal distribution, minimizing variance in weights π, is $f(S_t | S_{t-1}, y_t)$. The most important property of the particle filter is its ability to handle any nonlinearity. However, it has difficulties when S_t is high-dimensional. Essentially, the number of particles N required to adequately approximate the distribution grows exponentially with the dimensionality of the state space.

E Quasi-Maximum Likelihood Estimation

Since the measurement equation considered in each of the models is nonlinear, one possibility is to rely on Taylor series approximations to obtain extended forms of the Kalman filter. The transition and measurement equations analyzed in the previous section are expressed as follows:

$$y_t = \mu\left(S_{t-1}\right) + \lambda\left(S_{t-1}\right)\varepsilon_t, \tag{61}$$

$$S_t = AS_{t-1} + \eta_t, \tag{62}$$

where ε_t follows a standard normal distribution and η_t is a $d-$dimensional noise vector with variance-covariance matrix Σ. The deterministic functions $\mu\left(S_t\right)$ and $\lambda\left(S_t\right)$ define the conditional mean and volatility of excess returns and are characterized by each of the models.

I use Gaussian approximations to filter the mean and covariance of the states and measurement series. More specifically, the linearity of the state vector implies that the first and second conditional moments of the state vectors are

$$\begin{aligned} S_{t+1|t} &= AS_{t|t}, \qquad (63)\\ P_{t+1|t} &= AP_{t|t}A' + \Sigma, \qquad (64) \end{aligned}$$

where $S_{t+1|t}$ and $P_{t+1|t}$ are the time t predicted values of the conditional mean and covariance matrix of the state vector, respectively. These moments allow us to generate a predicted mean $y_{t+1|t}$ and covariance matrix $P^{yy}_{t+1|t}$ of the measurement series, given by

$$\begin{aligned} y_{t+1|t} &= \mathbb{E}\left[\mu\left(S_t\right) + \lambda\left(S_t\right)\varepsilon_{t+1}\left|y_t, y_{t-1}..., y_0\right.\right], \qquad (65)\\ P^{yy}_{t+1|t} &= Var\left[\mu\left(S_t\right) + \lambda\left(S_t\right)\varepsilon_{t+1}\left|y_t, y_{t-1}, ..., y_0\right.\right]. \end{aligned}$$

Finally, the covariance between the observed and unobserved variables, $P^{sy}_{t+1|t}$, is

$$P^{sy}_{t+1|t} = Cov\left[S_{t+1}, \mu\left(S_t\right) + \lambda\left(S_t\right)\varepsilon_{t+1}\left|y_t, y_{t-1}, ..., y_0\right.\right]. \tag{66}$$

Using these conditional moments, we apply the Kalman update, represented by the following set of recursive equations to obtained values for the conditional mean $S_{t+1|t+1}$ and covariance $P_{t+1|t+1}$:

$$\begin{aligned} K_{t+1} &= P^{sy}_{t+1|t}\left(P^{yy}_{t+1|t}\right)^{-1}, \qquad (67)\\ S_{t+1|t+1} &= S_{t+1|t} + K_{t+1}\left(y_{t+1} - y_{t+1|t}\right),\\ P_{t+1|t+1} &= P_{t+1|t} - K_{t+1}P^{yy}_{t+1|t}K_{t+1}. \end{aligned}$$

The first attempt to estimate the moments in Eqs. (65) through (67) uses closed-form expression, if available. An alternative way is to use Taylor series expansions of $\mu(S_t)$ and $\lambda(S_t)$ around $S_{t+1|t}$, for an arbitrary number of terms. Romero (2012) develops a nonlinear filter based on Gaussian approximations that uses high-order Taylor series to compute the moments involved in the Kalman filter update represented by Eq(67).

E.1 Quasi-Maximum Likelihood Function

Once the conditional mean, $y_{t+1|t}$, and conditional covariance, $P^{yy}_{t+1|t}$, for each observation are obtained, a quasi log-likelihood function for each observation is constructed assuming that each observation y_{t+1}, is normally distributed with mean, $y_{t+1|t}$, and volatility $P^{yy}_{t+1|t}$. Let θ denote the vector of parameters that are used to perform the Kalman filter. The log-likelihood for each observation, denoted by $l_t(\theta)$, is calculated as

$$l_t(\theta) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\left(P^{yy}_{t+1|t}\right) - \frac{1}{2}\frac{\left(y_{t+1} - y_{t+1|t}\right)^2}{P^{yy}_{t+1|t}}. \tag{68}$$

Finally, we choose the parameter values θ that maximize[31]

$$L(\theta) = \Sigma_{t=1}^{T} l_t(\theta). \tag{69}$$

[31] See Gallant and White (1988) for a detailed theoretical justification of quasi-maximum likelihood estimation.

F External Habit Formation Model

This section presents the model by Campbell and Cochrane (1999) in discrete time and its extension in Wachter (2005). A representative investor is assumed to have state-dependent preferences. More specifically, an investor has utility over consumption relative to a reference point X_t and maximizes

$$\mathbb{E}\left[\sum_{t=0}^{\infty}\delta^t \frac{(C_t - X_t)^{1-\gamma} - 1}{1-\gamma}\right], \tag{70}$$

where $\delta > 0$ is the time preference parameter and $\gamma > 0$ is the curvature parameter.

Each investor is concerned with her consumption relative to that of others. Habit X_t is defined through surplus consumption S_t, where

$$S_t \equiv \frac{C_t - X_t}{C_t}. \tag{71}$$

One can interpret S_t as a business cycle indicator. In economic booms, consumption substantially exceeds the external habit and the surplus, S_t, is large; and in recessions consumption barely exceeds the external habit, and the external habit is relatively small.

It is assumed that $s_t = \log S_t$ follows the process

$$s_{t+1} = (1-\phi)\,\overline{s} + \phi s_t + \lambda\,(s_t)\,(\Delta c_{t+1} - E_t\,[\Delta c_{t+1}])\,, \tag{72}$$

where $\overline{s}$ is the unconditional mean of s_t, ϕ is the persistence and $\lambda\,(s_t)$ is the sensitivity of the changes in consumption. The unconditional mean and the sensitivity function are defined in terms of primitive parameters. It is assumed that aggregate consumption growth is log-normal with independent and identically distributed innovations; that is

$$\Delta c_{t+1} = g + v_{t+1}, \tag{73}$$

where $c_t = \log C_t$ and $v_{t+1} \sim N\left(0, \sigma_v^2\right)$ is an i.i.d. sequence. The process for s_t is heteroscedastic and perfectly conditionally correlated with innovations in consumption growth. The sensitivity function $\lambda\,(s_t)$ is specified so that the real risk-free rate is linear, and for $s_t \approx \overline{s}$, x_t is a deterministic function of past consumption. Consequently, we have

$$\lambda\,(s_t) = \begin{cases} 1/\overline{S}\sqrt{1 - 2\,(s_t - \overline{s})} - 1, & \text{if } s_t \leq s_{\max} \\ 0 & \text{otherwise,} \end{cases} \tag{74}$$

$$\overline{S} = \sigma_v \sqrt{\frac{\gamma}{1 - \phi - b/\gamma}}, \tag{75}$$

where b is a preference parameter that determines the behavior of the risk-free rate and $s_{\max} = \overline{s} + \frac{1}{2}\left(1 - \overline{S}^2\right)$. In Campbell and Cochrane (1999), b is chosen to be zero and produce a constant real risk-free rate, while Wachter (2005) shows that values of $b > 0$, imply a risk-free rate that is linear in s_t;

F.1 Stochastic Discount Factor

Since the habit is external, the investor's inter-temporal marginal rate of substitution is given by

$$M_{t+1} = \delta \left(\frac{S_{t+1}}{S_t}\right)^{-\gamma} \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}. \tag{76}$$

Moreover, any asset return R_{t+1} must satisfy

$$\mathbb{E}_t \left[M_{t+1} R_{t+1}\right] = 1. \tag{77}$$

F.2 Risk-Free rate and Maximum Sharpe Ratio

Let $R_{f,t+1}$ denote the one-period risk-free return between t and $t+1$, and $r_{f,t+1} = \log\left(R_{f,t+1}\right)$; as a result, from Eqs.(76) and (77) imply that

$$\begin{aligned} r_{f,t+1} &= -\log\left(\mathbb{E}_t\left[M_{t+1}\right]\right) \\ &= -\log(\delta) + \gamma g + \gamma(1-\phi)(\overline{s} - s_t) - \frac{\gamma^2\sigma_v^2}{2}\left(1 + \lambda(s_t)\right)^2 \\ &= -\log(\delta) + \gamma g - \frac{\gamma(1-\phi) - b}{2} + b(\overline{s} - s_t), \end{aligned} \tag{78}$$

where the last equality comes from substituting the definition of $\lambda(s_t)$. This definition implies a risk-free rate linear in s_t.

Conditional on the information at time t, the one-period stochastic discount factor, defined in Eq.(76) is the exponential of a normally distributed random variable that has variance $\gamma^2\left[1 + \lambda(S_t)\right]^2 \sigma^2$. As a result, the Hansen-Jagannathan bound implies that

$$\sqrt{\exp\left(\gamma^2\left[1 + \lambda(S_t)\right]^2 \sigma^2\right) - 1}$$

is an upper bound on the Sharpe ratio of any portfolio. If λ is a decreasing function of S_t, then the upper bound on Sharpe ratios will be counter-cyclical: higher in recessions than in booms.

F.3 Price-Dividend Ratio

The aggregate market is represented as the claim to the future consumption stream. If P_t denotes the ex-dividend price of this claim, then Eq. (77) implies that in equilibrium P_t satisfies

$$E_t \left[M_{t+1} \left(\frac{P_{t+1} + C_{t+1}}{P_t} \right) \right] = 1, \tag{79}$$

which can be rewritten as

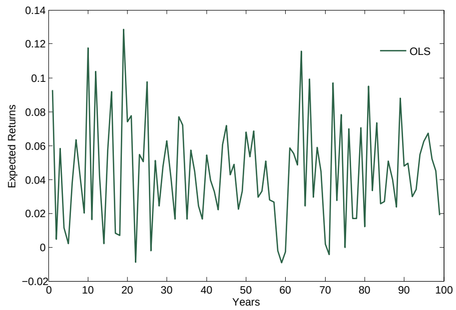
$$E_t \left[M_{t+1} \left(1 + \frac{P_{t+1}}{C_{t+1}} \right) \frac{C_{t+1}}{C_t} \right] = \frac{P_t}{C_t}.$$

Because C_t is the dividend paid by the aggregate market, P_t/C_t is the price-dividend ratio. The price-dividend ratio can be computed numerically using numerical methods; Wachter (2005) provides an efficient method for its computation.

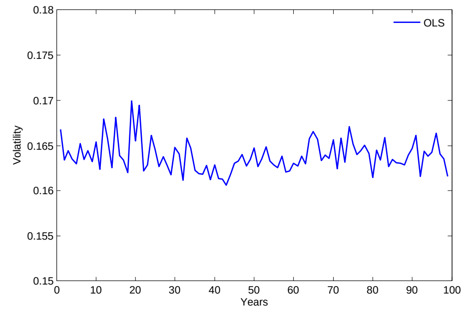
Returns on the aggregate market are defined as

$$R^m_{t+1} = \left(\frac{P_{t+1}/C_{t+1} + 1}{P_t/C_t} \right) \frac{C_{t+1}}{C_t}.$$

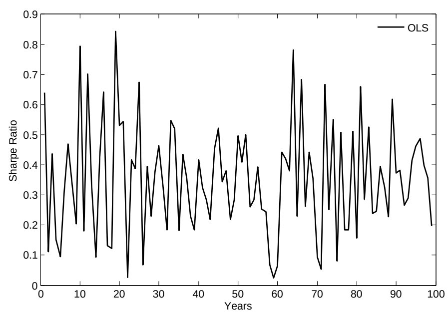
The main difficulty lies in solving the model (79) for the price-dividend ratio as a function of s_t. Once the price-dividend ratio is calculated numerically, Monte Carlo simulations can be performed to obtain accurate estimates of expected returns, volatilities and Sharpe ratios for different holding periods. Details about the simulations are explained in Wachter (2005).



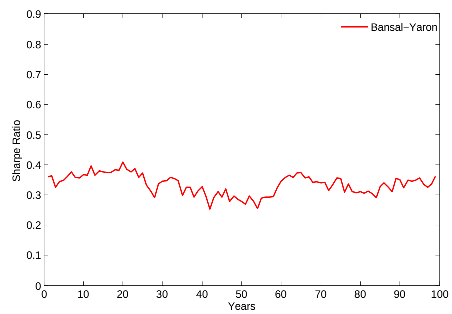
(a) Expected Returns (OLS)



(b) Conditional Volatility (OLS)



(c) Conditional Sharpe Ratios (OLS)



(d) Conditional Sharpe Ratios (BY)

Figure 1

This figure shows the results of a simulated path of annual returns using the calibration by Bansal and Yaron (2004). Each simulation has 100 annual observations of returns. Fitted values for the conditional mean and variance were constructed using predictor variables. Panel A shows a random path of annual returns with the fitted OLS values. Panel B shows the realized variance constructed with realized returns along with its OLS fitted values in dotted lines. Panel C contains the conditional Sharpe ratio estimates based on the OLS fitted values of the conditional mean and conditional volatility; Panel D contains the Sharpe ratios implied by the model.

Table 1

Long-Run Risks Parameters

Endowment Process Parameters	Symbol	BY Calibration
Mean Consumption Growth	μ_c	0.0015
LRR Persistence	ρ	0.979
LRR Volatility Multiple	φ_e	0.044
Mean Dividend Growth	μ_d	0.0015
Dividend Leverage	ϕ	3
Dividend Volatility Multiple	φ	4.5
Dividend Consumption Exposure	π	0
Baseline Volatility	$\overline{\sigma}$	0.0078
Volatility of Volatility	σ_w	0.0000023
Persistence of Volatility	ν	0.987

Preference Parameters	Symbol	BY Calibration
Risk Aversion	γ	10
EIS	ψ	1.5
Time Discount Factor	δ	0.998

Endowment Process:

$$\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_t + \sigma_t \eta_{t+1} \\
x_{t+1} &= \rho x_t + \varphi_e \sigma_t e_{t+1} \\
\sigma_{t+1}^2 &= \overline{\sigma}^2 + v\left(\sigma_t^2 - \overline{\sigma}^2\right) + \sigma_w w_{t+1} \\
\Delta d_{t+1} &= \mu_d + \phi x_t + \varphi \sigma_t u_{t+1} + \pi \sigma_t \eta_{t+1} \\
w_{t+1}, e_{t+1}, u_{t+1}, \eta_{t+1} &\sim i.i.d.\ \mathcal{N}(0,1).
\end{aligned}$$

This table displays the model parameters for Bansal and Yaron (2004) (BY). The endowment process is described above. All parameters are given in monthly terms. The standard deviation of the long-run innovations is equal to the volatility of consumption growth times the long-run volatility multiple, and the standard deviation of dividend growth innovations is equal to the volatility of consumption growth times the volatility multiple for dividend growth. Dividend consumption exposure is the magnitude of the impact of the one-period consumption shock on dividend growth. Dividend leverage is the exposure of dividend growth to long-run risks.

Table 2

Long-Run Risks Moment Comparison

Moment	OLS Regressions	Model
Expected Returns	0.0417	0.0417
Standard Deviation	0.0301	0.0087
Correlation	0.0052	
Volatility	0.1653	0.1641
Standard Deviation	0.0092	0.0167
Correlation	0.0434	
Conditional Sharpe Ratio	0.2645	0.3333
Standard Deviation	0.1582	0.0353
Correlation	0.0039	

This table displays moments calculated for the Bansal and Yaron (2004) model from annual datasets. Columns 1 and 2 display the results using years as time interval. The moment displayed is the median from 100,000 finite sample simulations of length 100 years. The returns on equity and the risk-free rate are aggregated to a yearly level by adding the log-returns within the year.

Table 3

Long-Run Risks Moment Comparison

Moment	Filtering	Model
Expected Returns	0.0418	0.0417
Standard Deviation	0.0080	0.0087
Correlation	0.5721	
Volatility	0.1645	0.1650
Standard Deviation	0.0168	0.0167
Correlation	0.5651	
Conditional Sharpe Ratio	0.3341	0.3333
Standard Deviation	0.0322	0.0353
Correlation	0.5694	

This table displays moments calculated for the Bansal and Yaron (2004) model. Columns 2 to 5 display the results using years as time interval. The moment displayed is the median from 1500 finite sample simulations of length 100 years. The returns on equity and the risk-free rate are aggregated to a yearly level by adding the log-returns within the year.

Table 4

Summary of Models in the Literature

	Preferences			Time-Varying		
	CRRA	**Recursive**	**Habit**	**Equity Premium**	**Volatility**	**Sharpe Ratios**
Lucas (1978)	✓					
Breeden (1979)	✓					
Mehra and Prescott (1985)	✓			✓	✓	✓
Rietz (1988)	✓			✓	✓	✓
Weil (1989)		✓		✓	✓	✓
Constantinides (1990)			✓			
Abel (1990)			✓	✓	✓	
Campbell and Cochrane (1999)			✓	✓	✓	✓
Chan and Kogan (2002)			✓	✓	✓	✓
Menzly, Santos, and Veronesi (2004)			✓	✓	✓	✓
Bansal and Yaron (2004)		✓		✓	✓	✓
Barro (2006)	✓					
Calvet and Fisher (2007)		✓		✓	✓	✓
Barro (2009)		✓				
Jermann (2010) *				✓	✓	✓
Papanikolaou (2011)		✓		✓	✓	✓
Wachter (2012)		✓		✓	✓	✓
Chien, Cole, and Lustig (2012)	✓			✓	✓	✓

This table compares features of asset pricing models which have been used to price the aggregate stock market: Lucas (1978), Breeden (1979), Mehra and Prescott (1985), Rietz (1988), Weil (1989), Constantinides (1990), Abel (1990), Campbell and Cochrane (1999), Chan and Kogan (2002), Menzly, Santos, and Veronesi (2004), Bansal and Yaron (2004), Barro (2006), Calvet and Fisher (2007), Barro (2009), Jermann (2010), Papanikolaou (2011), Wachter (2012),Chien, Cole, and Lustig (2012). The comparison table is divided into two panels. The first panel focuses on the features of the model (preferences, endowment and technology), while the second focuses on the pricing implications of the various models.

Table 5

Quasi-Maximum Likelihood Parameter Estimates

Positive Risk Premia

	Model A	Model B	Model C	Model D
a_{11}	0.6727	0.7029	0.7204	0.7079
	(0.0066)	(0.0041)	(0.0834)	(0.1211)
a_{21}	-0.0894	-0.1521		
	(0.0279)	(0.0184)		
a_{12}	0.3215		-0.4948	
	(0.0011)		(0.0938)	
a_{22}	0.8310	0.7400	0.9182	0.8730
	(0.0025)	(0.0114)	(0.1798)	(0.1142)
b_{11}	0.2897	0.1350	0.0944	0.1924
	(0.0063)	(0.0194)	(0.5390)	(0.1674)
b_{22}	0.0020	0.0001	0.0055	0.0072
	(0.0070)	(0.0054)	(0.1111)	(0.8430)
ρ	-0.3073	-0.1760	-0.7989	-0.7995
	(0.0009)	(0.0004)	(0.2773)	(0.0029)
$\overline{\mu}$	0.0131	0.0131	0.0131	0.0131
	(0.0000)	(0.0159)	(0.0675)	(0.1065)
$\overline{\sigma}$	0.0857	0.0857	0.0857	0.0857
	(0.0000)	(0.0005)	(0.0166)	(0.5172)
Q-lik	245.37	245.32	244.93	244.69

This table presents the quasi-maximum likelihood estimates of the models of the form

$$y_t = \mu(S_{t-1}) + \lambda_(S_{t-1})\varepsilon_t,$$

and

$$S_t = AS_{t-1} + \underline{\eta_t} \text{ with } \underline{\eta_t} \sim N\left(0, \Sigma\right),$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \Sigma = \begin{bmatrix} b_{11} & \rho\sqrt{b_{11}b_{22}} \\ \rho\sqrt{b_{11}b_{22}} & b_{22} \end{bmatrix}.$$

$\mu(S_t) = \overline{\mu}\exp(S_{1t})$ and $\sigma(S_t) = \overline{\sigma}\exp(S_{2t})$. The estimates are for quarterly returns on the value-weighted CRSP index in excess of the three-month Treasury bill from the second quarter of 1953 to the fourth quarter of 2011. Standard errors are reported in parentheses.

Table 6

Regressions on Quarterly Data

No	Constant	Lag	Cay	d-p	RREL	TRM	DEF	R^2
Panel A: Excess Returns: 1953:2 - 2011:4								
1953:2 - 2011:4								
1	0.01	0.07						0.01
	(2.76)	(1.18)						
2	0.02		0.79					0.03
	(2.92)		(2.57)					
3	0.01	0.08	0.80					0.03
	(2.81)	(1.28)	(2.71)					
4	0.14		0.76	0.02				0.04
	(1.84)		(2.30)	(1.61)				
5	0.13	0.08	0.67	0.02	-1.46			0.07
	(1.86)	(1.31)	(2.16)	(1.64)	(-2.64)			
6	0.16	0.06	0.61	0.03	-1.29	0.84		0.08
	(2.12)	(0.99)	(1.99)	(1.96)	(-2.20)	(1.59)		
7	0.16	0.06	0.61	0.03	-1.30	0.84	-0.07	0.08
	(1.87)	(0.99)	(1.99)	(1.83)	(-2.30)	(1.59)	(-0.05)	
Panel B:Log Excess Returns:								
1953:2 - 2011:4								
1	0.01	0.08						0.01
	(2.14)	(1.31)						
2	0.01		0.82					0.03
	(2.34)		(2.65)					
3	0.01	0.09	0.83					0.03
	(2.18)	(1.41)	(2.81)					
4	0.15		0.78	0.03				0.04
	(1.96)		(2.35)	(1.77)				
5	0.15	0.09	0.70	0.03	-1.37			0.07
	(2.00)	(1.43)	(2.23)	(1.81)	(-2.38)			
6	0.17	0.07	0.65	0.03	-1.21	0.79		0.08
	(2.25)	(1.14)	(2.07)	(2.11)	(-1.99)	(1.49)		
7	0.18	0.07	0.63	0.03	-1.26	0.83	-0.38	0.08
	(2.07)	(1.14)	(2.01)	(2.04)	(-2.17)	(1.56)	(-0.28)	

This table reports estimates from OLS regressions of quarterly realized returns and log-returns for the CRSP VW index on lagged explanatory variables for the second quarter of 1953 to the fourth quarter of 2011. The conditioning variables are lagged realized volatility (*Lag*); the consumption, wealth, income ratio (*cay*); log dividend-price ratio ($d - p$); the relative bill rate (*RREL*); the term spread, the difference between the ten-year Treasury bond yield and the three-month Treasury bond yield (*TRM*); the Baa-Aaa default spread (*DEF*). The t-stats were constructed with heteroscedasticity-consistent standard errors.

Table 7
Regressions on Quarterly Data

No	Constant	Lag	Cay	d-p	RREL	TRM	DEF	R^2
Panel C: Realized Volatility of Excess Returns: 1953:2 - 2011:4								
1	0.03	0.61						0.37
	(5.16)	(7.39)						
2	0.07		-0.32					0.02
	(16.76)		(-2.02)					
3	0.03	0.60	-0.23					0.38
	(5.35)	(7.54)	(-2.76)					
4	-0.08		-0.28	-0.03				0.12
	(-2.16)		(-1.69)	(-3.90)				
5	-0.05	0.54	-0.23	-0.02	-0.27			0.41
	(-2.74)	(6.84)	(-2.58)	(-3.97)	(-1.00)			
6	-0.06	0.54	-0.22	-0.02	-0.28	-0.07		0.41
	(-3.02)	(6.84)	(-2.48)	(-4.39)	(-1.06)	(-0.51)		
7	-0.10	0.46	-0.16	-0.02	-0.18	-0.19	1.34	0.43
	(-3.64)	(5.33)	(-1.71)	(-4.57)	(-0.75)	(-1.28)	(2.83)	
Panel D: Realized Volatility of Log Excess Returns: 1953:2 - 2011:4								
1	0.03	0.61						0.36
	(5.15)	(7.38)						
2	0.07		-0.32					0.02
	(16.7)		(-2.01)					
3	0.03	0.60	-0.23					0.37
	(5.34)	(7.53)	(-2.75)					
4	-0.08		-0.28	-0.03				0.12
	(-2.18)		(-1.68)	(-3.92)				
5	-0.05	0.54	-0.23	-0.02	-0.27			0.41
	(-2.76)	(6.83)	(-2.57)	(-3.99)	(-0.99)			
6	-0.06	0.54	-0.22	-0.02	-0.28	-0.07		0.41
	(-3.04)	(6.83)	(-2.46)	(-4.41)	(-1.05)	(-0.48)		
7	-0.10	0.46	-0.16	-0.02	-0.17	-0.19	1.34	0.42
	(-3.65)	(5.32)	(-1.71)	(-4.58)	(-0.73)	(-1.24)	(2.81)	

This table reports estimates from OLS regressions of quarterly realized volatility of returns and log-returns for the CRSP VW index on lagged explanatory variables for the second quarter of 1953 to the fourth quarter of 2011. The conditioning variables are lagged realized volatility (Lag); the consumption, wealth, income ratio (cay); log dividend-price ratio ($d-p$); the relative bill rate ($RREL$); the term spread, the difference between the ten-year Treasury bond yield and the three-month Treasury bond yield (TRM); the Baa-Aaa default spread (DEF). The t-stats were constructed with heteroscedasticity-consistent standard errors.

Table 8

Summary Statistics of Expected Returns, Volatilities and Sharpe Ratio Estimates

		Mean	St. Dev.	Min	Max	AC(1)
	μ_t	0.01	0.02	-0.04	0.07	0.80
OLS Methods	σ_t	0.07	0.02	0.02	0.22	0.79
	SR_t	0.30	0.42	-0.61	1.84	0.81
	μ_t	0.02	0.01	0.01	0.03	0.59
Brandt and Kang (2004)	σ_t	0.09	0.01	0.07	0.12	0.85
	SR_t	0.25	0.05	0.14	0.41	0.61

This table reports descriptive statistics of the estimates of expected returns, volatilities and Sharpe ratios. The first set of conditional moments are estimated from OLS regressions of quarterly realized log-returns for the CRSP VW index on lagged explanatory variables for the first quarter of 1953 to the last quarter of 2011. The second set of estimates are based on the reduced form model by Brandt and Kang (2004) in which the conditional mean and volatility of stock returns are treated as latent variables.

Table 9

Quasi-Maximum Likelihood Parameter Estimates

Unconstrained Risk Premia

	Model A	Model B	Model C	Model D
a_{11}	0.5276	0.5532	0.5090	0.5282
	(0.0498)	(0.0002)	(0.0440)	(0.0026)
a_{21}	-0.4967	-0.4154		
	(0.0206)	(0.0001)		
a_{12}	-0.0165		-0.0247	
	(0.1388)		(0.3278)	
a_{22}	0.8426	0.8551	0.8859	0.8221
	(0.1521)	(0.0000)	(0.9465)	(0.0068)
b_{11}	0.0002	0.0004	0.0001	0.0004
	(0.0014)	(0.0012)	(0.0029)	(0.0020)
b_{22}	0.0088	0.0048	0.0091	0.0132
	(0.0091)	(0.0013)	(0.1097)	(0.1982)
ρ	-0.7994	-0.7491	-0.7999	-0.7678
	(0.0409)	(0.0002)	(0.2009)	(0.0012)
$\overline{\mu}$	0.0131	0.0131	0.0131	0.0131
	(0.0259)	(0.0300)	(0.0013)	(0.1253)
$\overline{\sigma}$	0.0857	0.0857	0.0857	0.0857
	(0.0045)	(0.0036)	(0.0001)	(0.0106)
Q-lik	246.12	245.72	246.01	245.52

This table presents the quasi-maximum likelihood estimates of the models of the form

$$y_t = \mu(S_{t-1}) + \lambda_(S_{t-1})\varepsilon_t,$$

and

$$S_t = AS_{t-1} + \underline{\eta_t} \text{ with } \underline{\eta_t} \sim N\left(0, \Sigma\right),$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \Sigma = \begin{bmatrix} b_{11} & \rho\sqrt{b_{11}b_{22}} \\ \rho\sqrt{b_{11}b_{22}} & b_{22} \end{bmatrix}.$$

$\mu(S_t) = \overline{\mu} + S_{1t}$ and $\sigma(S_t) = \overline{\sigma}\exp(S_{2t})$. The estimates are for quarterly returns on the value-weighted CRSP index in excess of the three-month Treasury bill from the second quarter of 1953 to the fourth quarter of 2011. Standard errors are reported in parentheses.

Table 10

Quasi-Maximum Likelihood Parameter Estimates

Parameters	Model D		Extended Model D	
	Estimate	S.E.	Estimate	S.E.
a_{11}	0.7079	0.1211	0.5135	0.3421
a_{21}	-	-	-	-
a_{12}	-	-	-	-
a_{22}	0.8730	0.1142	0.7649	0.1381
b_{11}	0.1924	0.1674	0.0049	0.5434
b_{22}	0.0072	0.8430	0.0006	0.0690
ρ	-0.7995	0.0029	-0.4523	0.0882
$\overline{\mu}$	0.0131	0.1065	0.0131	0.0021
$\overline{\sigma}$	0.0857	0.5172	0.0857	0.0009
c_{11}	-	-	7.7812	2.6462
c_{12}	-	-	1.0911	0.8892
c_{13}	-	-	-38.8899	1.5632
c_{14}	-	-	0.4021	0.7437
c_{15}	-	-	-39.1056	0.1624
c_{21}	-	-	-1.3562	1.7051
c_{22}	-	-	-0.1460	0.0644
c_{23}	-	-	-0.1245	5.6614
c_{24}	-	-	-5.4407	1.4016
c_{25}	-	-	10.7989	0.4681
Q-lik	244.69		263.46	

This table presents the Quasi-maximum likelihood estimates of the model of the form

$$y_t = \mu(S_{t-1}) + \lambda_(S_{t-1})\varepsilon_t,$$

and

$$S_t = Cx_t + AS_{t-1} + \underline{\eta_t} \text{ with } \underline{\eta_t} \sim N(0, \Sigma),$$

where

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} \\ c_{21} & c_{22} & c_{23} & c_{24} & c_{25} \end{bmatrix}, A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \Sigma = \begin{bmatrix} b_{11} & \rho\sqrt{b_{11}b_{22}} \\ \rho\sqrt{b_{11}b_{22}} & b_{22} \end{bmatrix}.$$

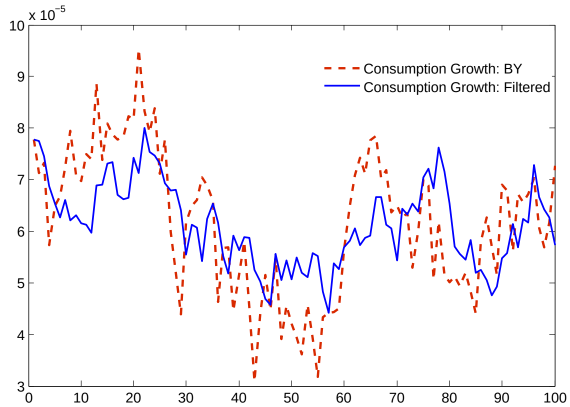
$\mu(S_t) = \overline{\mu}\exp(S_{1t})$ and $\sigma(S_t) = \overline{\sigma}\exp(S_{2t})$. The vector of conditioning variables x_t contains the de-meaned consumption, wealth, income ratio (cay); log dividend-price ratio ($d-p$); the relative bill rate ($RREL$); the term spread, the difference between the ten-year Treasury bond yield and the three-month Treasury bond yield (TRM); and the Baa-Aaa default spread (DEF). Heteroscedasticity-consistent standard errors are reported. The estimates are for quarterly returns on the value-weighted CRSP index in excess of the three-month Treasury bill from the second quarter of 1953 to the fourth quarter of 2011.

TABLE 11

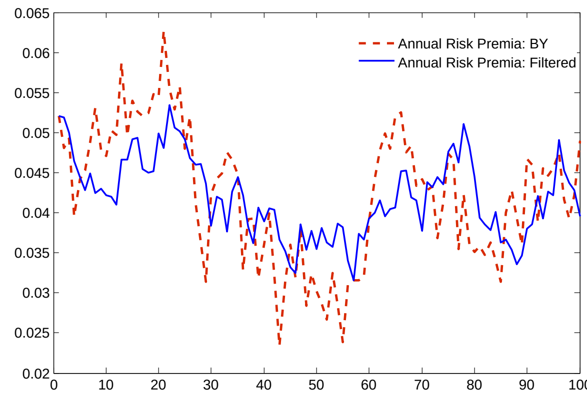
Summary Statistics of Sharpe Ratio Estimates

	Mean	St. Dev.	Min	Max	AC(1)
BK	0.25	0.05	0.14	0.41	0.61
BK (Unconstrained)	0.26	0.10	-0.05	0.49	0.71
BK (Extended)	0.31	0.25	0.07	1.58	0.88
OLS Methods	0.30	0.42	-0.61	1.84	0.81

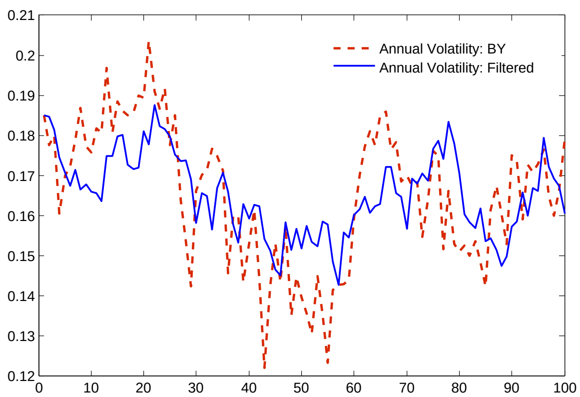
This table reports descriptive statistics of the estimates of Sharpe ratios based on quarterly realized log-returns for the CRSP VW index for the first quarter of 1953 to the last quarter of 2011. The first, second and third sets of Sharpe ratio estimates are based on the reduced form model by Brandt and Kang (2004) (BK) in which the conditional mean and volatility of stock returns are treated as latent variables. The first representation guarantees positive values for the conditional mean and volatility, while the second representation guarantees a positive volatility only. The third representation is an extended version in which the conditional moments are positive functions of exogenous predictors. Finally, the last set of Sharpe ratio estimates is based on the conditional moments estimated from OLS regressions of log-returns on lagged explanatory variables.



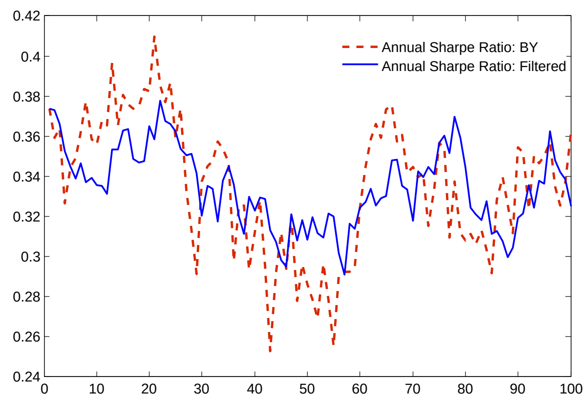
(a) Consumption Growth



(b) Risk Premia



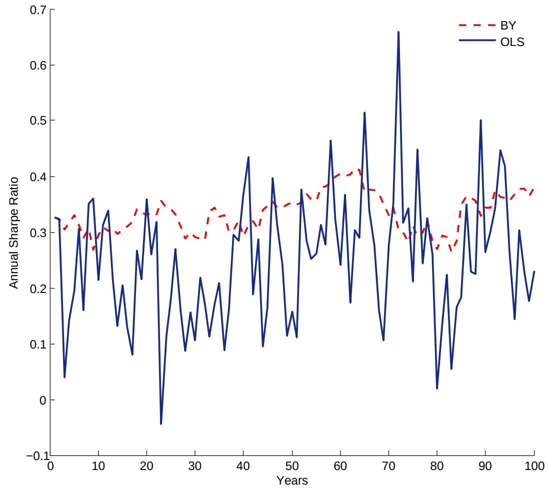
(c) Conditional Volatility



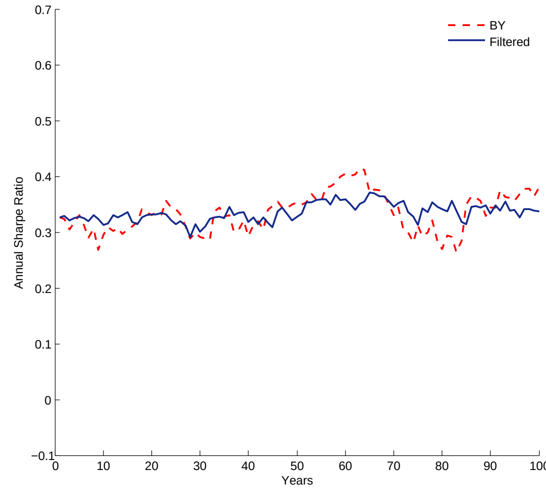
(d) Conditional Sharpe Ratio

Figure 2

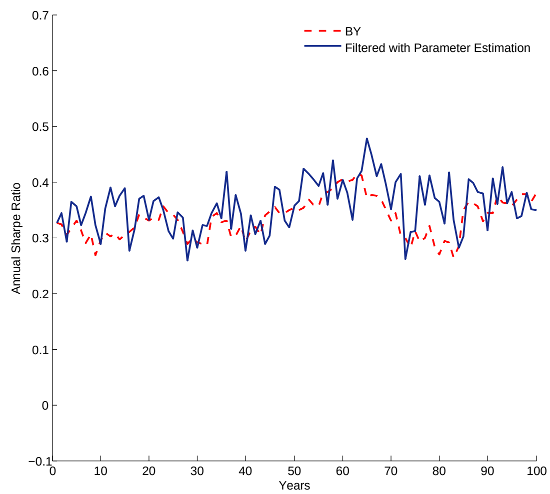
This figure shows the results of a simulated path of the volatility of consumption growth using the calibration by Bansal and Yaron (2004). Each simulation has 100 annual return observations. Panel A shows a random path of monthly returns of the volatility of consumption growth. The dotted line represents the filtered values of σ_t^2. Panel B shows the simulated risk premia along with its filtered values in dotted lines. Panel C contains the simulated standard deviation of the risk premia as well as its filtered values. Panel D contains the simulated conditional Sharpe ratio along with its filtered values. The dashed lines are assumed to be unobservable to the econometrist, while the continuous lines are the filtered values.



(a) OLS Fitted Values



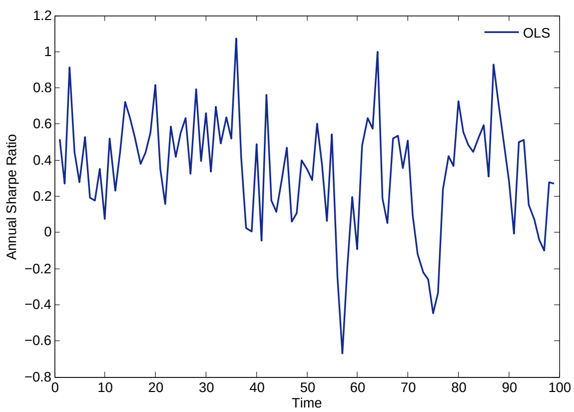
(b) Filtered Values (known parameter values)



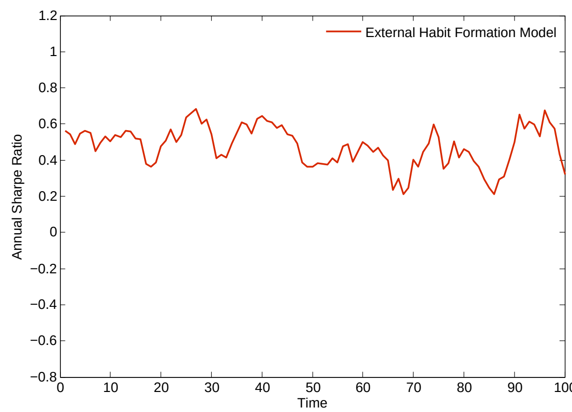
(c) Filtered Values (unknown parameter values)

Figure 3

This figure shows the results of a simulated path of the volatility of consumption growth using the calibration by Bansal and Yaron (2004). Each simulation has 100 annual return observations. Panel A contains the conditional Sharpe ratio estimates based on the OLS fitted values. Panel B contains the filtered Sharpe ratio estimates implied by the long-run risks model; the dotted lines represent the annual Sharpe ratio implied by the model which are assumed to be unobservable to the econometrist; Panel C contains the filtered Sharpe ratio estimates implied by the long-run risks model based on parameter estimates obtained via quasi-maximum likelihood. The dotted lines represent the annual Sharpe ratio implied by the model, which are assumed to be unobservable to the econometrist. The simulations were performed with the calibrated parameter values from Bansal and Yaron (2004).



(a) OLS



(b) External Habit Formation

Figure 4

This figure shows the results of a simulated path of the volatility of consumption growth using the calibration by Campbell and Cochrane (1999). Each simulation has 100 annual return observations. Panel A contains the conditional Sharpe ratio estimates based on the OLS fitted values; Panel B contains the filtered Sharpe ratio estimates implied by the external habit formation model. The simulations were performed with the calibrated parameter values from Campbell and Cochrane (1999).

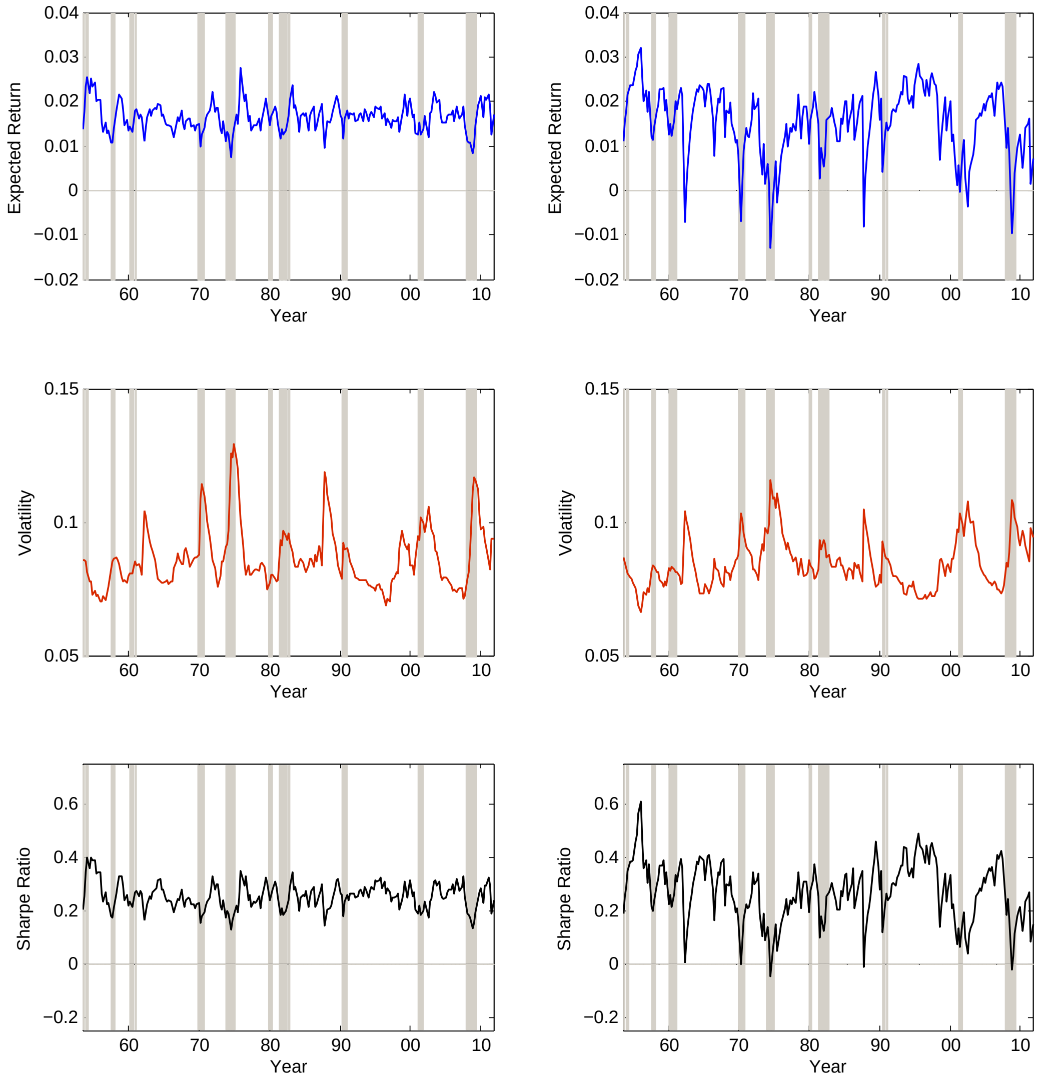


FIGURE 5

This figure shows the estimates of the conditional mean, volatility and Sharpe ratio. The figures show the quarterly estimates of the conditional mean, μ_t, conditional volatility, σ_t and Sharpe ratio, SR_t, obtained via filtering techniques. The left column corresponds to the model with a positive risk premia and the right column contains the filtered estimates of the model with an unconstrained risk premia. The vertical bars represent the NBER recession dates.

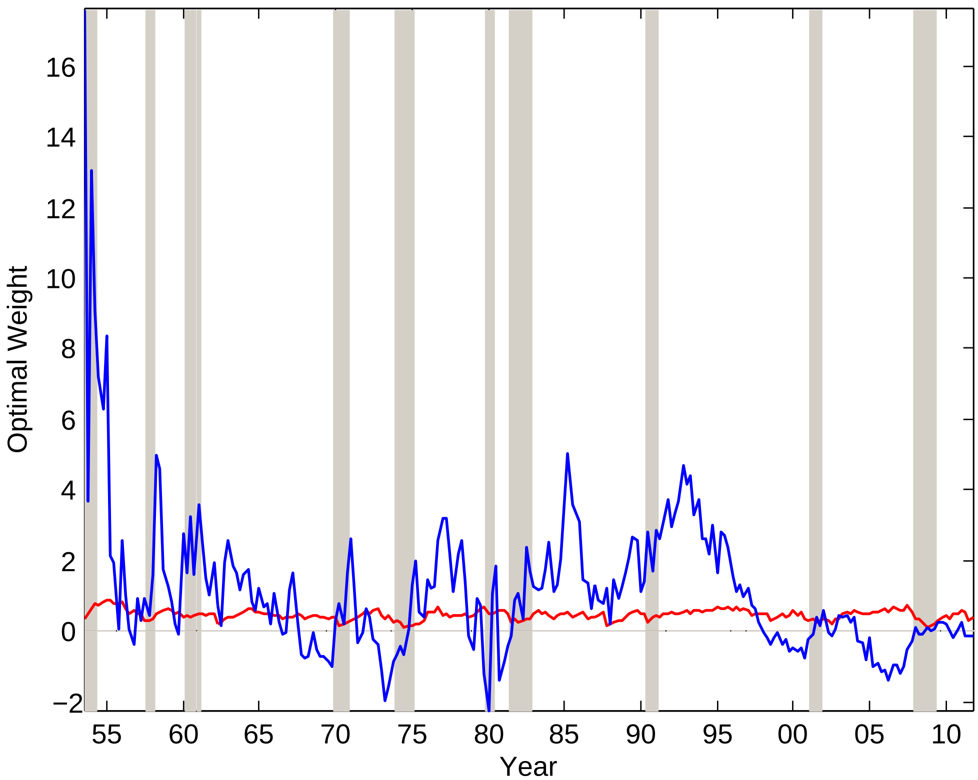


FIGURE 6

This figure shows the portfolio weights estimates based on the conditional mean, volatility and Sharpe ratio. The figure shows the time series of optimal weights, $w_t = (\mu_t + \sigma_t^2/2)/(\gamma\sigma_t^2)$, where γ represents the risk aversion parameter, and μ_t and σ_t are the quarterly estimates of the conditional mean and conditional volatility respectively. The figure shows the optimal weights based on OLS techniques (blue) and the model based on nonlinear latent variables, assuming a positive risk premium (red) and $\gamma = 5$. The vertical bars represent the NBER recession dates.