Three Essays on Asset Pricing

by

Ji Zhou

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Accounting and Finance

Asper School of Business

University of Manitoba

Winnipeg

Abstract

This thesis consists of three essays. In the first essay, we derive a pricing kernel for a continuous-time long-run risks (LRR) economy with the Epstein-Zin utility function, non-i.i.d. consumption growth, and incomplete information about fundamentals. In equilibrium, agents learn about latent conditional mean of consumption growth and price equity simultaneously. Since the pricing kernel is endogenous and affected by learning, uncertainty about unobserved conditional mean of consumption growth affects risk prices corresponding to shocks in both consumption and dividend growth. We demonstrate our analytical results by applying the model to a profitability-based equity valuation model proposed by Pastor and Veronesi (2003). Calibration of the model demonstrates that the LRR model with learning has potential to fit levels of price-dividend ratios of the S&P 500 Composite Index, equity premium, and the short term interest rate simultaneously.

In essay two, we extend the LRR model with incomplete information proposed in essay one by incorporating inflation and applying the model to the valuation of nominal term structure of interest rate. We estimate the processes of state variables and latent variables using a Bayesian Markov-Chain Monte Carlo method. In the estimation, we rely only on the information in macro-economic data on aggregate consumption growth, inflation, and dividend growth on S&P 500 Composite Index. In this way, parameters and latent state variables are estimated outside the model. Estimation results suggest a mildly persistent LRR component. However, both real and nominal yield curves implied by the LRR model are downward-sloping. We show that the inverted yield curve is due to a negative risk premium, which is determined jointly by covariance between shocks in state variables and shocks in the nominal pricing kernel. Incorporating learning about the mean

consumption growth flattens the yield curve but does not change the sign of the yield curve slope.

In essay three, we study the critique of the conditional affine factor asset pricing models proposed by Lewellen and Nagel (2006). They suggest that two important economic constraints are overlooked in cross-sectional regressions. First, the estimated unconditional slope associated with a risk factor should equal the average risk premium on that factor in a conditional model. Second, the estimated slope associated with the product of a risk factor and an instrument should be equal to the covariance of the factor risk premium with the instrument. We test both constraints on conditional models with time-varying betas and our results confirm the proposition. Also, from the functional relationship between conditional and unconditional betas, we identify an unconditional constraint on unconditional betas for time-varying beta models and develop a testing procedure subject to this constraint. We show that imposing this unconditional constraint changes estimates of unconditional betas and risk prices significantly.

Acknowledgements

First, I would like to express my sincere gratitude to Dr. Alex Paseka for his patience, support, and encouragement throughout my time at the University of Manitoba. Without his guidance and belief, it would not be possible for me to finish this thesis.

I also want to thank the rest of my committee: Dr. Steven Zheng and Dr. Aerambamoorthy Thavaneswaran. They provided invaluable feedback over the course of my thesis work. I am also very grateful to Dr. Marie Racine for agreeing to serve as the external examiner.

I would like to thank my family: my brother Yu for his encouragement and company in hard times and my parents Weirong and Ruihua for their unconditional love and support they have provided me with over the years.

Finally, I would like to say thanks to all my professors and academic support staff during my graduate education at the University of Manitoba.

Dedication

To my father

Table of Contents

List of Tables

List of Figures

Chapter One: General Introduction

This thesis presents three essays on asset pricing. In the first two essays, we extend the long-run risks model of Bansal and Yaron (2004) by incorporating incomplete information on fundamentals and derive an endogenous pricing kernel for the extended model. With the pricing kernel, we study the impact of incomplete information on both the equity premium and the term structure of interest rates. In essay three, we examine the conditional affine factor asset pricing models. More specifically, we identify an unconditional constraint on unconditional betas and develop a testing procedure that incorporates the constraint. We also examine the critique proposed by Lewellen and Nagel (2006).

In long-run risk models, the means of consumption growth and dividend growth are assumed to be non-i.i.d. and contain a small mean-reverting component (the LRR component hereafter). If the LRR component is persistent, current shocks to expected growth have a persistent effect on the expectations about consumption and dividend growth. In equilibrium, investors exposed to the risk require a higher premium for holding risky assets.

In previous studies on the LRR models, the LRR component is assumed to be known to investors. However, in reality, this is not the case. Investors have to learn the value of conditional consumption mean based on information available to them. To this end, in the first essay we model how investors learn about the latent variable. We assume that investors build a minimum mean square estimate of the latent state variable from observations on consumption and dividend growth. The solution of this optimal filtering problem allows us to derive the processes for the state variables under a new probability measure corresponding to information available to investors. We further assume that investors have Epstein-Zin recursive preferences. Under this assumption, we derive the endogenous equilibrium pricing kernel that contains all information about investors' estimate of the LRR component. The estimation risk of the unobserved conditional mean of consumption growth affects all risk prices in the pricing kernel and, consequently, all contributions to the equity premium. Determined by the drift of the pricing kernel, the short rate in our model is also endogenous and is an affine function of the estimated conditional mean of consumption growth and, thus, inherits its mean-reverting property.

To demonstrate our analytical results, we apply our pricing kernel to a profitability-based equity valuation model proposed by Pastor and Veronesi (2003). The equity premium in the combined model has two components: the component corresponding to risk in profitability and the contribution of discount rate risk. We calibrate the model to data on the aggregate consumption and dividends as well as dividend-per-share on the S&P 500 composite index. If observed processes - consumption and dividend growth – have higher (in absolute value) correlations with the latent process, the model can reproduce both the real equity premium (8.55%) and the price-dividend ratio (41.82) with a real short rate of 2.65% per year. These levels are highly sensitive to parameter values such as mean profitability growth. In fact, within the uncertainty bands on mean profitability growth, the model has the potential to reproduce levels of all three variables of interest in the data. Our main focus is on interesting model implications for equity premium properties. When mean consumption growth is not conditionally correlated with consumption and dividend growth, a large proportion of equity premium is due to hedging demands induced by shocks in profitability (10.09%)

since profitability is much more volatile than the model-implied short rate. Surprisingly, the contribution of the short rate risk is negative (-3.42%). This is because of the negative elasticity of the price-dividend ratio to the short rate. The lower estimation risk (due to the higher correlation between observables and latent variables) increases the absolute value of the covariances of the profitability and the short rate with the stochastic discount factor (SDF) but, at the same time, reduces the equity premium's sensitivities to both of these covariances. Overall, the profitability contribution becomes larger (11.32%) and the short rate contribution is less negative (-2.77%). Therefore, the total equity premium is larger as learning becomes more effective.

In essay two, we extend the LRR model with the learning feature proposed in essay one by incorporating inflation and applying the model to the estimation of the term structure of interest rates. We now assume that investors infer the LRR component not only from consumption and dividend growth but also from inflation and its mean[1]. Since the filtered LRR component is a linear combination of observed state variables, the model-implied pricing kernel contains shocks from all four observables. Therefore, the risk premium is jointly determined by the asset payoff's conditional covariance with four observed variables. The nominal short rate, which is the negative of the drift of the nominal pricing kernel, is an affine function of the LRR component and the mean inflation. Just like their counterparts in the real model, both the nominal pricing kernel and the nominal short rate are endogenous in the sense that they are the outcomes of a joint solution of the investors' recursive utility optimization and the filtering problem –

[1] We assume that inflation mean is observed by investors in order to isolate the effect of learning about conditional consumption mean and to preserve closed form solutions.

the estimation of the latent state. As a result, risk premiums are directly tied to parameters of the state processes, which can be estimated using only data on fundamentals and inflation.

To pursue this goal, we propose a Bayesian Markov Chain Monte Carlo method (Bayesian MCMC) to estimate all parameters and state variables simultaneously. More specifically, we estimate the Euler-discretized state variable processes using Gibbs sampling. To improve efficiency, we sample latent variables using the forward filtering backward sampling (FFBS) algorithm developed by Carter and Kohn (1994) and Frühwirth-Schnatter (1994). FFBS is essentially a multi-step Gibbs sampling algorithm designed for linear state-space models with Gaussian error (or mixture of normals). Building on Kalman filter results, FFBS allows us to construct Markov chains for both parameters and state variables that eventually converge to their marginal limiting distributions. The end result is a set of marginal posterior densities for parameters and latent variables that allow us to make inferences on any functions of parameters and states such as the yield curve. This approach renders results free from biases arising in the estimation of state variables keeping parameters fixed at their point estimates such as in the Kalman filter. The bias in this case is due to ignoring the sampling variation in parameter values, which normally occurs in multi-step estimation procedures.

We estimate parameters and state variable processes using information contained only in the aggregate consumption, dividends on S&P 500, and inflation from January 1959 to December 2014. In this way, parameters and latent state variables are estimated outside the model, i.e., we use no price information for parameter and state variable inference. The estimation results suggest a mildly persistent LRR component and random-walk mean inflation at the annual frequency. More importantly, the LRR component is positively correlated with consumption and dividend growth and negatively correlated with inflation and mean inflation. The mean inflation is positively correlated with inflation but negatively correlated with consumption and dividend growth. These correlations jointly determine both the short rate and the risk premium in the model. However, in previous studies, these correlations are largely ignored. For instance, in Bansal and Yaron (2004) (and follow-up papers such as Kiku (2006), Drechsler and Yaron (2011), Bansal and Shaliastovich (2012), and Doh (2013)), the state variables and latent variables are assumed to be uncorrelated. Piazzezi and Schneider (2007) assume that latent variables and state variables are driven by the same set of innovations, effectively assuming perfect correlations. In our estimation approach, these correlation estimates lead to a negative conditional covariance between the pricing kernel and the LRR component and a positive conditional covariance between the pricing kernel and the mean inflation. Given parameter and latent variable estimates, we compute the nominal yield curve both with and without incomplete information. In the complete information case, our results show that the homoscedastic LRR model with inflation fails to reproduce the upward-sloping shape of the observed yield curve. More specifically, a substantial proportion of the risk premium is due to the hedging demand induced by the variation in the LRR component, with a minor contribution from hedging demand due to the shocks in mean inflation. Since the zero-coupon bond return loads negatively on LRR component and mean inflation, the LRR component risk premium is negative and the mean inflation risk premium is positive. Combining two risk premiums we get a negative total risk premium (-21.3 bps per month for the 1-year zero coupon bond) and a downward sloping nominal yield curve. Incorporating learning into the model partially remedies the LRR model's failure in reproducing the correct term structure. However, the magnitude of the effect is small. Learning reduces the LRR component's conditional covariance with the pricing kernel and has a positive overall impact on the total risk premium. Nevertheless, yield curves still retain the negative slope even after we incorporate learning into the model (e.g, the slope increases from -21.3 bps to -20.9 bps for the 1-year zero coupon bond). When we turn on the uncertainty (and learning), another result is a reduction in the mean short rate. The monthly short rates with and without complete information are 0.6357% and 0.6280%, respectively. The reduction is mainly due to changes in the precautionary saving effect and adjustments for inflation in both models. The combination of these two effects shifts the yield curve downward and makes it flatter.

In essay three, we test conditional affine factor asset pricing models. Previous studies show that, conditional factor models are able to explain a large proportion of variation in average stock returns (such as Jagannathan and Wang (1996), Lettau and Ludvigson (2001b), and Santos and Veronesi (2006)). These models suggest that identifying proper instruments to describe time variation in an investment opportunity set in conditional linear factor models is a fruitful avenue for future research in asset pricing. However, Lewellen and Nagel (2006) (hereafter LN) warn that the success of the conditional models may be illusory and should be treated with caution. LN argue that conditional models mentioned above may not be able to produce statistically small unconditional alphas. Their analysis implies that good performance of those conditional models in cross-sectional tests may be due to the fact that some important unconditional constraints are omitted in the estimation procedure.

While Ludvigson (2011) later shows that CCAPM is immune to LN's critique, it is still not clear if models with time-varying beta assumptions are subject to LN's critique. In this Chapter, we follow LN's work and examine the unconditional moment implications of the conditional factor models where factor betas are assumed to be time-varying or, more specifically, affine functions of instruments. Note that the last assumption is not implied in any way by linear SDF models. Instead, it must be considered as an additional constraint on the model. Theoretically, the assumption that betas are affine functions of instruments is likely invalid. The 'affine beta' constraint does not necessarily correctly map back into the SDF affine in factors, which is a central assumption of linear asset pricing models in general and the models we test, in particular. Even if the affine beta assumption is not valid theoretically, it may still be a valid empirical approximation. In this paper, we examine the empirical validity of this assumption. Once the conditional model is conditioned down to its unconditional version, the affine restriction on conditional factor betas leads to unconditional constraints on unconditional betas and risk prices (LN's critique). This feature allows us to assess the empirical validity of the affine beta assumption by comparing estimation results with or without imposing these unconditional constraints. To do so, we develop an empirical testing procedure that allows us to incorporate the constraints on factor betas and risk prices.

To impose the 'affine beta' constraint in empirical tests, we develop a three-stage regression procedure. We first estimate time-varying betas using rolling-window regressions. In the second-stage, we run a regression of the estimated time-varying betas on instruments to retrieve unconditional betas. The last step is to run the cross-sectional regression to estimate factor risk premiums. Our methodology produces the estimates of unconditional betas from estimated time series of time-varying betas based on the functional relationship between them. For demonstration, we apply our test procedure to Santos and Veronesi (2006)'s labor income ratio model using data on aggregate consumption, labor income, market returns, and returns on various portfolios[2]. Our empirical results show that imposing the constraint changes estimates of unconditional betas and factor risk premiums significantly and has a significant negative impact on the model performance measured by the root mean squared error (RMSE) of pricing errors.

[2] 25 Fama-French size and book-to-market portfolios, 30 industry portfolios, 10 momentum portfolios, and 6 earnings-to-price portfolios.

Chapter Two: Equity Pricing and Risk Premium under Long-Run Risks and Incomplete Information

1. Introduction

The high level of equity premium over the post-war period has long been a major puzzle in theoretical finance. With the power utility function, the observed post-war data implies the representative agent has a relative risk aversion γ larger than 50. With such large risk aversion, the risk free rate should be much higher than what we observe in practice[3]. Moreover, under the observed equity premium, investors should invest a much larger proportion of their total wealth into the equity market than currently observed. (See Cochrane (2005), Chapter 21)

One potential solution to this risk premium puzzle is the long-run risk (LRR) model first developed by Bansal and Yaron (2004) (BY hereafter). In the LRR model, the representative agent is assumed to have Epstein-Zin recursive preferences (Epstein and Zin (1989)). More importantly, means and volatilities of state variables (consumption and dividend growth) are stochastic and mean-reverting. These assumptions lead to extra sources of risk. In equilibrium, investors exposed to the risk require higher premium for holding equities. Eraker (2008) extends LRR model by introducing jumps into the processes of state variable volatilities and assuming that the conditional inflation mean is negatively correlated with consumption and dividend growth[4]. Under these assumptions, Eraker (2008) is able to produce reasonable equity premium levels and a positively

[3] Given the typical values of discount parameter ($\delta = 0.01$), and mean and standard deviation of consumption growth ($E_t(\Delta c) = 0.01, \sigma_t(\Delta c) = 0.01$), the power utility function implies a risk free rate $r_t^f = \delta + \gamma E_t(\Delta c) - \frac{1}{2}\gamma(\gamma+1)\sigma_t^2(\Delta c) = 38\%$.

[4] This assumption suggests that investors consider the high inflation as a signal of low future consumption growth. The theoretical rationale and empirical evidence of this assumption is given by Piazzesi and Schneider (2005).

sloped nominal yield curve with the levels of risk aversion and the elasticity of intertemporal substitution (EIS) of at least 8 and 5, respectively. In a more recent study, Zhou and Paseka (2014) examine the asset-pricing implication of learning about the LRR component. They argue that the LRR component is not directly observed by investors. Instead, investors estimate the LRR component using observations of consumption and dividend growth. As an example, Zhou and Paseka (2014) apply the incomplete information LRR SDF to the earnings-based equity pricing model of Bakshi and Chen (2006) and show that the resulting model has the potential to reproduce the levels of the short rate, the risk premium, and price-to-earnings ratios simultaneously with a risk aversion of 10 regardless of EIS level and economic uncertainty.

Our work in this Chapter adopts Zhou and Paseka (2014)'s model setup. We assume that the LRR component in the conditional means of consumption and dividend growth is not observed by the representative agent. The agent has to learn the value of LRR component based on information available. We assume that the agent forms a minimum mean square estimate of LRR component from observations on consumption and dividend growth. We solve the agent's optimal filtering problem and derive the processes of state variables under a new probability measure corresponding to information available to investors. Given the state variables under the new measure, we derive the process for the SDF. Unlike Zhou and Paseka (2014), we apply the SDF to the profitability-based equity pricing model developed by Pastor and Veronesi (2003)[5]. In both cases, the incomplete information pricing kernel is applied to the complete information version of the underlying equity pricing model, Bakshi and Chen (2006) in

[5] Pastor and Veronesi (2003) study the impact of uncertainty in mean profitability on stock valuation. Since we are interested in pricing implications of learning on LRR component, we only use the base case of Pastor and Veronesi (2003)'s model where mean profitability is assumed to be known.

Zhou and Paseka (2014) or Pastor and Veronesi (2003) in this essay. The model-implied equity price is a function of short rate and earnings, and the model-implied equity premium contains two components corresponding to shocks in short rate and profitability. In Zhou and Paseka (2014) both equity price and premium are determined by earnings growth, mean earnings growth, and short rate.

Jacoby, Paseka, and Wang (2014) (JPW hereafter) also study the impact of learning on asset pricing. JPW adopt the earnings-based equity pricing model of Bakshi and Chen (2006). Earnings growth is modeled as a stochastic process with an unobserved mean, itself a mean-reverting random process. In order to value assets JPW assume an exogenous pricing kernel process. Our model is different from that in JPW along the following dimensions. First, JPW adopt the earnings-based equity pricing model of Bakshi and Chen (2006) while we use the profitability-based model of Pastor and Veronesi (2003). This leads to the difference in contributions to the risk premium. More importantly, JPW study the pricing implications of incomplete information on equity payoffs, while we focus on those of incomplete information on the pricing kernel. More specifically, JPW assume an exogenous pricing kernel, which is not affected by learning. As a result, learning in JPW has no impact on risk prices. It only changes the risk premium contribution corresponding to shocks in earnings growth. In contrast, we assume incomplete information on fundamentals. With the Epstein-Zin recursive utility, we derive the endogenous pricing kernel based on investors' estimates of fundamental processes. Therefore, learning impacts all risk prices in the modeled economy, and, consequently, all contributions to the equity premium even without any parameter uncertainty on the payoff side of the equity pricing model.

In the empirical part of the current paper, we calibrate our model to data on aggregate consumption and dividends, and profitability of the S&P 500 Composite Index. Based on calibrated parameters, we compute the model-implied short rate, equity premium, and price-to-dividend ratio. The model generally reproduces levels of real risk premium (8.55%) and price-to-dividend ratio (41.83)[6] in the data with the real annual short rate of 2.65%. These levels are sensitive to parameter values such as mean profitability growth. However, within the uncertainty bands on mean profitability growth, the model has the potential to reproduce levels of all three variables of interest in the data. Our main focus is on interesting model implications for equity premium properties. When mean consumption growth is not conditionally correlated with consumption and dividend growth, a large proportion of equity premium is due to hedging demands induced by shocks in profitability (10.09%) since the profitability is much more volatile than model-implied short rate. Surprisingly, the contribution of the discount rate portion of the risk premium is negative (-3.42%). This is because of the negative elasticity of the price-dividend ratio to the short rate. The lower estimation risk (due to higher correlation between observables and latent variables) increases the absolute value of the covariances of the profitability and the short rate with the SDF but, at the same time, reduces the equity premium's sensitivities to both of these covariances. Overall, the profitability contribution becomes larger (11.32%) and the short rate contribution is less negative (-2.77%). Therefore, the total equity premium is larger as learning becomes more effective.

[6] Short rate, price-to-dividend ratio, and equity premium are computed assuming the LRR component is positively correlated with state variables. We assume the relative risk aversion $\gamma = 8$. Other model parameters are either set to be in line with their estimated values from previous studies or estimated from the data. See section 4 for details of our calibration.

The rest of the paper is organized as follows. In section 2 we review the related literature. In section 3 we state model assumptions, solve the investor's optimal filtering problem, and derive the value function. In section 4, we derive the SDF, as well as model-implied short rate. In section 5, we apply the SDF to the profitability-based stock pricing model of Pastor and Veronesi (2003) to derive the stock price, price to dividend ratio, and risk premium. We calibrate the model to aggregate consumption and dividend and profitability per share of S&P 500 Composite Index. Finally, we conclude in section 6.

2. Related Literature

2.1 The Long-run Risks Model

The long-run risks model is first proposed by Bansal and Yaron (2004) (BY) as an attempt to explain puzzles in asset pricing (i.e. equity premium puzzle, large asset price volatility, value effect). BY make the following assumptions: the state variables, consumption and dividend growth, contain a small but persistent component in their mean; conditional volatility of consumption and dividend growth are stochastic and mean-reverting; and investors have Epstein-Zin recursive preferences. The first and second assumptions introduce two extra sources of risk on top of the contemporaneous shocks in the consumption growth (short-run risk). Intuitively, since mean consumption and dividend growth are persistent, current shocks in mean consumption growth have a persistent effect on the expectations about consumption and dividend growth. Fluctuations in state variable volatility, which capture the economic uncertainty, lead to fluctuations in risk prices. With these two extra sources of risk, investors are exposed to a higher level of total risk in equilibrium; hence, a higher premium for holding assets.

Epstein-Zin recursive preferences also play an important role in the model. With these recursive preferences, one can decouple risk aversion from intertemporal elasticity of substitution (IES). BY show that, when risk aversion is larger than the reciprocal of the IES, investors prefer early resolution of risk and the risk price for long-run expected growth risk is positive. BY derive the model-implied pricing kernel as well as the risk premium. As expected, the risk premium contains three terms: the contribution corresponding to short-run risk, the one corresponding to shocks in mean consumption growth (long-run risk), and the one corresponding to shocks in the stochastic consumption volatility (economic uncertainty). Calibration results suggest that, with a persistent mean consumption growth, IES $\psi = 1.5$, and risk aversion $\gamma = 10$, the model matches the observed risk premium, short rate, asset price volatility, and captures many of the return and dividend growth predictability dimensions.

Bansal, Gallant and Tauchen (2007) extend the LRR model by incorporating consumption-dividend cointegration. With the assumption that the difference between log dividend and log consumption is stationary, dividend and consumption cannot deviate from each other permanently. Similar cointegration assumption is adopted by Bansal, Dittmar and Kiku (2009) in which the de-meaned log dividend is cointegrated with de-meaned log consumption. While using different estimation approaches (efficient method of moment in Bansal, Gallant and Tauchen (2007) and generalized method of moments (GMM) in Bansal, Dittmar and Kiku (2009)), both studies find their models able to match the market risk premium. Bansal, Dittmar and Kiku (2009) also find that long-run risks drive risk compensation while short-run risks compensation goes to zero at long horizons.

Eraker (2008) studies a LRR model with inflation and jumps in consumption volatility. Under this model setup, the discretized state variables follow affine jump-diffusion processes. Eraker shows that, with Epstein-Zin utility, asset prices are exponential affine functions of state variables. In calibration exercise, when mean inflation impacts the real growth negatively (the negative signaling effect of mean inflation on consumption growth suggested by PS), the model is able to produce equity premium and an upward-slopping nominal term structure that are in line with data when intertemporal elasticity of substitution $\psi = 5$ and risk aversion $\gamma = 8$.

Bansal and Yaron (2004), Bansal, Gallant and Tauchen (2007), Eraker (2008), and Bansal, Dittmar and Kiku (2009) all assume the LRR component is known to investors. In contrast, our model incorporates incomplete information on the LRR component. Doing so, we introduce an extra source of risk, the estimation risk, in to the LRR model, which affects both structure and magnitude of the equity premium.

On the empirical side, there is a wide range of studies examining the LRR model's performance in reproducing the level of equity premium and/or other values of interest: among others, these include Bansal, Khatchatrian, and Yaron (2005), Bansal, Kiku and Yaron (2008, 2010), Malloy, Moskowitz, and Vissing-Jørgensen (2009), Drechsler and Yaron (2010), Aldrich and Gallant (2011), Constantinides and Ghosh (2011), Beeler and Campbell (2012), Hasseltoft (2012), Ferson, Nallareddy and Xie (2013), Ortu, Tamoni, and Tebaldi (2013), Schorfheide, Song, and Yaron (2014), Bansal, Kiku, Shaliastovich, and Yaron (2014), and Jagannathan and Marakani (2015).

Bansal, Kiku and Yaron (2008) and Malloy, Moskowitz, and Vissing-Jørgensen (2009) estimate LRR models using GMM. Bansal, Kiku and Yaron (2008) test the model on the data sample of aggregate consumption, short rate, and market, size, and book-to-market stock portfolios and find a quite persistent LRR component with annual sample autocorrelation equal to 0.67. Given the estimates of the LRR component, the model appears to explain asset prices well including the ‘value’ and ‘size’ premium. Malloy, Moskowitz, and Vissing-Jørgensen (2009) use consumption growth for stockholders as the state variable. Their results suggest that, with the stockholder consumption growth, the LRR model can match observed risk premia with unit IES and risk aversion around 10.

Schorfheide, Song, and Yaron (2014) develop a Bayesian Markov Chain Monte Carlo (MCMC) approach to utilize both annual consumption data from 1929 and monthly data from 1959. Their estimates suggest a persistent LRR component with autocorrelation mean of 0.993. They also show that the model is able to match the various moments of consumption and dividend growth, market return, risk-free rate, and price-to-dividend ratio.

Hasseltoft (2012) estimates the model using simulated method of moments (SMM) and consumption, equity, and bond data. Hasseltoft finds the LRR models can produce equity premium and an upward-slopping nominal term structure in line with data simultaneously.

Ortu, Tamoni, and Tebaldi (2013) decompose the consumption growth time series into components based on their persistence measured by half-life. They report that, the low-frequency component of consumption growth with half-life between 8 and 16 years accounts for an annual risk premium up to 2%.

Bansal, Khatchatrian, and Yaron (2005), Drechsler and Yaron (2010), and Bansal, Kiku, Shaliastovich, and Yaron (2014) examine the importance of economic uncertainty and its impact on risk premiums. Bansal, Khatchatrian, and Yaron (2005) find that consumption volatility level has a significant negative impact on future price to dividend ratio, which implies a positive risk premium on the risk introduced by the dynamic consumption volatility. Such positive volatility risk premium is reported by Bansal, Kiku, Shaliastovich, and Yaron (2014), who also show that the volatility risk plays an important role in producing positive correlation between returns on human capital and financial wealth. Drechsler and Yaron (2010) provide evidence of time-varying economic uncertainty from the derivative market by examining the dynamics of the variance premium[7].

While literature stated above favours the LRR, there are a number of studies furnishing evidence contrary to the LRR model. Constantinides and Ghosh (2011) test the LRR model using GMM and data on consumption, dividends and returns on CRSP market index as well as size and book-to-market portfolios. They report that, while the model can match the unconditional moments of state variables, it mismatches the short-rate, and volatility of p/d ratio, short rate, and market return. Ferson, Nallareddy and Xie (2013) follow Constantinides and Ghosh (2011)'s test approach and study the out-of-sample performance of the LRR models (both with and without consumption-dividend cointegration) in terms of its explanatory power for equity and bonds premiums, size and book-to-market effects, momentum, and reversal. To conduct the out-of-sample test, Ferson, Nallareddy and Xie (2013) first estimate the model over a 36-year estimation

[7] The variance premium is defined by Drechsler and Yaron (2010) as the difference between the Chicago Board Options Exchange's (CBOE) VIX index and the conditional expectation of realized variance.

window, and then evaluate model performance in the year right after the estimation window. Over the annual data sample from 1931 to 2009, Ferson, Nallareddy and Xie (2013) find that the LRR models outperform CAPM and CCAPM in explaining the momentum effect, but fail to explain the credit premium and term premium.

2.2 Incomplete Information and Learning

Previous work on the LRR models assumes that the LRR component in mean consumption growth (and dividend growth) is known by the representative agent. However, this is not the case in practice. More realistically, the agent draws inference on latent states from the observable consumption and dividend growth in the LRR model. This kind of incomplete information and learning model setup is adopted by a long line of research to explain a wide range of questions in finance including Timmermann (1993), Veronesi (2000), Brennan and Xia (2001), Lewellen and Shanken (2002), Pastor and Veronesi (2003, 2006, 2009), Li (2005), Lettau Ludvigson and Wachter (2008), and Ai (2007).

Veronesi (2000) examines the effect of learning on risk premium in an economy where dividend and consumption are the same and unobserved. Surprisingly, with power utility, Veronesi finds that the risk premium decreases with a lower level of information uncertainty.

Brennan and Xia (2001) and Li (2005) consider the model in which a time-varying mean of dividend growth rate is unobserved. Both studies suggest that information uncertainty introduces an extra source of risk and therefore raises the risk premium. Li (2005) also studies Veronesi (2000)'s model and shows that the results of Veronesi (2000) are due to the fact that noisy information makes the perceived future consumption growth more volatile, which increases demand for stocks to hedge the consumption risk.

Unlike Brennan and Xia (2001) and Li (2005), we assume both mean consumption and dividend growth are unobserved. In this sense, our model is more related to that in Veronesi (2000). As a result, our model-implied equity premium decreases with a lower level of estimation error, which is consistent with the results of Veronesi (2000).

Jacoby, Paseka, and Wang (2015) present a dynamic earnings-based stock-valuation model in which the mean earnings growth rate is not observed by investors. The calibration of their model demonstrates that the posterior variance of mean earnings growth estimate generates extra risk premium, and the magnitude of this risk premium is determined jointly by posterior error variance of the estimate and firm characteristics.

Ai (2007) introduces learning into a production-based LRR model, where the productivity of the technology contains a persistent latent component. The representative agent learns about the latent component in productivity from realizations on capital and a noisy signal. When risk aversion is larger than 1, consistent with previous studies, the agent requires a higher level of risk premium with information uncertainty. Ai also shows that, when learning is not perfect, the agent’s optimal estimation of the LRR component is less volatile than the true LRR component. This phenomenon further leads to smaller consumption volatility as well as a more stable risk free rate.

In the work mentioned so far, the volatility of consumption growth is assumed to be known. However, uncertainty about the volatility of consumption growth can also affect the level of equity premium. Lettau Ludvigson and Wachter (2008) (LLW hereafter) assume the unobservable volatility jump between two states. LLW's model is based on the Epstein and Zin (1989) utility function. Assuming the aggregate dividend equals consumption raised to power λ, LLW estimate the posterior probability of the low-volatility state, and find that the probability increased in 1990s. LLW therefore argue that the price run-up in 1990s can be partly described as a rational response to a decline in consumption risk.

Parameter uncertainty and learning can also help to explaining the volatility and predictability puzzle. Timmermann (1993) proposes a model that can explain the excess volatility in the learning framework. The log dividend growth is assumed to be a random walk with unknown mean and variance. Timmermann shows that, at least in the short-run, this model can explain the excess volatility and the predictability.

Lewellen and Shanken (2002) study the predictability in the framework of learning. They confirm that parameter uncertainty and learning can be a source of predictability other than the changes in expected returns with business cycle. Intuitively, if expected dividends are higher than the true mean, prices will be inflated. In future periods, investors generally will receive a negative surprise, which will lead to lower returns. Therefore, when looking back, an econometrician will find that high prices predict relatively low future returns. However, such predictability can only be observed ex-post since reversals are driven by rational behaviours.

Pastor and Veronesi (2003) explore the impact of uncertainty in firm profitability on market to book (M/B) ratio and return volatility. The profitability is defined as the ratio of firm's earnings to the book value of equity and assumed to follow a mean-reverting process with unobserved long-term mean. With the assumption that the market value and the book value of equity converge in future, Pastor and Veronesi show that the M/B ratio is a convex increasing function of mean profitability. Since higher uncertainty in mean profitability increases the tail weight of investors' perceived probability density function (PDF) of mean profitability, it naturally increases the M/B ratio. For the same reason, the idiosyncratic return is more volatile with the higher uncertainty in mean profitability. Given these results, Pastor and Veronesi (2006) show that the high Nasdaq valuations in the late 1990s may be explained by the high uncertainty about average firm profitability. In a similar fashion, Pastor and Veronesi (2009) examine the dynamics of stock valuation (M/B ratio) during a technology revolution with unobserved productivity gain from new technology. Their analysis suggests that the M/B ratio of firms adopting new technology usually increases with perceived productivity gain at the beginning of a technology revolution but this relation becomes negative with mass adoption of technology.

Like Brennan and Xia (2001) and Jacoby, Paseka, and Wang (2015), Pastor and Veronesi (2003) assume an exogenous pricing kernel process. Since the pricing kernel is not affected by learning, risk prices are the same as their counterpart in a complete information model. Learning only affects the risk premium contribution from shocks in the unobserved variables. In contrast, we derive the endogenous pricing kernel from the perceived processes of state variables. Therefore learning has an impact on all prices of risk, and consequently, on all contributions to the equity premium in our model.

3. Long-Run Risk Model with Learning

In this section, we introduce an incomplete information LRR model. We retain main features of the LRR model (BY, Eraker (2008)).

Assumption 1: *Following Eraker (2008) we model the processes of consumption growth and dividend growth with drift being a function of the long-run risk component* x_t

$$dln(C) = dg_c = \left(\mu_c + x_t - \frac{V}{2}\right)dt + \sqrt{V}dw_c, \tag{1}$$

$$dg_d = \left(\mu_d + \phi x_t - \varphi_d^2 \frac{V}{2}\right)dt + \varphi_d \sqrt{V}dw_d, \tag{2}$$

$$dx_t = -\rho x_t dt + \varphi_e \sqrt{V}dw_x, \tag{3}$$

Where w_x *,* w_c *and* w_d *are standard Wiener processes with* $Edw_c dw_d = \rho_{cd}dt$ *,* $Edw_x dw_c = \rho_{xc}dt$*, and* $Edw_x dw_d = \rho_{xd}dt$ *.*

In BY and Eraker (2008), volatility of state variables are time-varying and mean-reverting. The model admits a closed form solution for equity prices only under the assumption of homoscedasticity (i.e. V is constant), which is the assumption we adopt below.

Assumption 2: *The latent LRR component in mean consumption growth is not observed by agents. They only know that* x_t *follows a mean reverting process with zero long-run mean and form their least mean square estimate of* x_t*,* $\hat{x}_t$*, based on observations of consumption and dividend growth.*

Under assumption 2, representative agents determine equilibrium asset prices jointly with the estimation of the conditional mean of consumption growth. Due to the Markovian nature of the model, agents can solve their optimal filtering problem first. Then, treating the estimate of the conditional mean of consumption growth as given, they find equilibrium asset prices. Here, we use the Kalman-Bucy filter to solve agents' optimal filtering problem. Following Liptser and Shiryaev (2013), we present the solution to the filtering problem in the following theorem.

Theorem 1: *(Liptser and Shiryaev (2013)) The investors' best estimate of* x_t, $\hat{x}_t$, *and processes of* g_c *and* g_d *under a new probability measure based on the information set available to agents, are given by*

$$dg_c = \left(\mu_c + \hat{x}_t - \frac{V}{2}\right)dt + \sqrt{V}dw_c^*, \tag{4}$$

$$dg_d = \left(\mu_d + \phi\hat{x}_t - \varphi_d^2\frac{V}{2}\right)dt + \varphi_d\sqrt{V}dw_d^*, \tag{5}$$

$$d\hat{x}_t = -\rho\hat{x}_t dt + \Sigma_x \begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix}, \tag{6}$$

where dw_c^* *and* dw_d^* *are standard Wiener processes under the new measure:*

$$\begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix} = \begin{bmatrix} \dfrac{x_t - \hat{x}_t}{\sqrt{V}}dt + dw_c \\ \dfrac{\phi(x_t - \hat{x}_t)}{\varphi_d\sqrt{V}}dt + dw_d \end{bmatrix},$$

volatility of $\hat{x}_t$, Σ_x, *are*

$$\Sigma_x = \begin{bmatrix} \Sigma_{1x} \\ \Sigma_{2x} \end{bmatrix}^T = \frac{1}{1-\rho_{cd}^2}\begin{bmatrix} S\dfrac{(\varphi_d - \phi\rho_{cd})}{\varphi_d\sqrt{V}} + \varphi_e\sqrt{V}(\rho_{cx} - \rho_{dx}\rho_{cd}) \\ S\dfrac{(\phi - \varphi_d\rho_{cd})}{\varphi_d\sqrt{V}} + \varphi_e\sqrt{V}(\rho_{dx} - \rho_{cx}\rho_{cd}) \end{bmatrix}^T. \tag{7}$$

and S_t *is the posterior variance of* x_t,

$$S_t = E_t[(x_t - \hat{x}_t)^2], \tag{8}$$

which follows a process:

$$dS_t = (-\alpha S_t^2 - 2(\rho + \xi)S_t + \Gamma^2)dt$$

Parameters $\alpha, \xi,$ *and* Γ *are summarized below:*

$$\alpha = \frac{\phi^2 + \varphi_d^2 - 2\phi\varphi_d\rho_{cd}}{V\varphi_d^2(1-\rho_{cd}^2)},$$

$$\xi = \frac{\varphi_e}{\varphi_d(1-\rho_{cd}^2)}\big(\phi(\rho_{cx} - \rho_{dx}\rho_{cd}) + \varphi_d(\rho_{dx} - \rho_{cx}\rho_{cd})\big),$$

$$\Gamma^2 = V\varphi_e^2\left(1 - \frac{1}{1-\rho_{cd}^2}(\rho_{dx} - \rho_{cx}\rho_{cd})^2 - \rho_{cx}^2\right).$$

In this Chapter, we only focus on the properties of the model in the steady-state, $\frac{dS_t}{dt} = 0$. The steady-state posterior variance is then given by $S = \frac{-(\rho+\xi)+\sqrt{(\rho+\xi)^2+\alpha\Gamma^2}}{\alpha}$.The magnitude of S in the steady-state is determined by conditional correlations between the LRR component and state variables. For example, when mean consumption growth is perfectly correlated with consumption and dividend growth, in the steady-state, S is equal to zero, in other words, investors can learn the LRR component perfectly. When these correlations are zero, learning is less efficient, and the estimation error S has the largest steady-state value.

Following previous LRR literature, we assume investors have the stochastic differential utility preferences of Duffie and Epstein (1992). The stochastic differential utility preferences are the continuous-time analogue of the Epstein-Zin recursive preferences, and allow for the separation of risk aversion from the elasticity of intertemporal substitution (EIS).

Assumption 3: *Investors have the stochastic differential utility preferences of Duffie and Epstein (1992):*

$$J(ln(i), g_d, m, \hat{x}, W_t, t) = E_t \int_t^T f(C_s, J_s) ds, \tag{9}$$

with normalized aggregator

$$f(C_t, J_t) = \frac{\beta}{1-\frac{1}{\psi}}(1-\gamma)J_t\left[\left(\frac{C_t}{\left((1-\gamma)J_t\right)^{\frac{1}{1-\gamma}}}\right)^{1-\frac{1}{\psi}} - 1\right], \tag{10}$$

where T is investors' horizon, C_t is consumption at time t, J_t is the recursive utility at time t, β is the discount rate, γ is the relative risk aversion, and ψ is the elasticity of intertemporal substitution.

For a given consumption process, Duffie and Epstein (1992) prove the Bellman optimality condition and show that it implies that the optimal differential utility satisfies the following PDE:

$$J_t + \hat{\mu}_G^T J_G + \frac{1}{2dt} dG^T J_{GG} dG + f = 0, \tag{11}$$

$$J(T, G) = 0, \tag{12}$$

where $G = (\ln(i), g_c, g_d, m, \hat{x})'$ and $\hat{\mu}_G$ is the mean vector given by

$$\hat{\mu}_G = \begin{bmatrix} \mu_c + \hat{x}_t - \frac{V}{2} \\ \mu_d + \phi \hat{x}_t - \varphi_d^2 \frac{V}{2} \end{bmatrix}.$$

To get an analytical solution of the PDE (11) to (12), we adopt the log-linear approximation suggested by Campbell and Viceira (2003) and Zhu (2006). First, we rewrite the aggregator as follows:

$$f(C_t, J_t) = \frac{\beta}{1 - \frac{1}{\psi}} (1 - \gamma) J_t [F(C_t, J_t) - 1],$$

$$F(C_t, J_t) = \left(\frac{C_t}{\left((1 - \gamma) J_t\right)^{\frac{1}{1-\gamma}}} \right)^{1 - \frac{1}{\psi}}.$$

Campbell and Viceira (2003) show that $\beta F(C_t, J_t)$ is the optimal consumption-wealth ratio:

$$\beta F = e^{c-w}, \tag{13}$$

where c and w are log consumption and log wealth, respectively.

Second, we expand (13) around the unconditional mean of log consumption-wealth ratio, $c_0 - w_0$, using Taylor series expansion to obtain the desired log-linear approximation. This approximation is valid as long as the consumption-wealth ratio is relatively stable:

$$\beta F = h_0 + h\, ln(\beta F), \tag{14}$$

$$\begin{gathered} h = e^{c_0 - w_0}, \\ h_0 = e^{c_0 - w_0}[1 - (c_0 - w_0)]. \end{gathered} \tag{15}$$

Inserting this approximation (14)-(15) into the normalized aggregator (10), we have:

$$f \approx \frac{(1-\gamma)}{1-\frac{1}{\psi}} J[h_0 + hln\beta + hlnF - \beta] \tag{16}$$

$$= h(1-\gamma)J\left[lnC - \frac{1}{(1-\gamma)} lnJ + H\right]$$

where

$$H = \frac{(h_0 + hln\beta - \beta)}{h\left(1-\frac{1}{\psi}\right)} - \frac{1}{(1-\gamma)} ln(1-\gamma).$$

The above approximation (16) allows us to get closed-form solutions to the model. However, this approach has a drawback. With this approximation, the EIS, ψ, has no impact on pricing due to the fact that the EIS only appears in H and has no effect on the aggregator (16) or the value function to the first order of the log-linearization.

With the aggregator (16), we can solve PDE (11) with an exponential-affine solution (see Appendix for details):

$$J(t, g_c, g_d, \hat{x},) = \exp\{\xi_{0t} + +\xi_{1t} g_c + \xi_{2t} g_d + \xi_{3t}\hat{x}\}. \tag{17}$$

where

$$\xi_1 = 1 - \gamma, \tag{18}$$

$$\xi_2 = 0, \tag{19}$$

$$\xi_3 = \frac{1-\gamma}{h+\rho}. \tag{20}$$

4. Stochastic Discount Factor

In this section, we derive the incomplete information stochastic discount factor. According to Duffie and Epstein (1992), with recursive preferences (9), the SDF has the following expression:

$$\pi_t = \exp\left\{\int_0^t f_J dx\right\} f_C. \tag{21}$$

Given the normalized aggregator approximation (16), we apply Itô's lemma to (21) to obtain the full differential of the SDF:

$$\frac{d\pi}{\pi} = f_J dt + \frac{df_C}{f_C} \tag{22}$$

$$= \left(\frac{f}{J} - h\right) dt - \frac{dC}{C} + \frac{dJ}{J} + \frac{(dC)^2}{C^2} - \frac{dC}{C}\frac{dJ}{J}.$$

Equations (17) to (20) imply the following full differential for the value function:

$$dJ = -f dt + J_C C\sqrt{V} dw_c^* + J_d \varphi_d \sqrt{V} dw_d^* + +J_{\hat{x}} \Sigma_x \begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix}, \tag{23}$$

$$\frac{J_d}{J} = 0, \frac{J_C}{J} = \frac{\xi_1}{C}, \frac{J_{\hat{x}}}{J} = \xi_3. \tag{24}$$

From the consumption growth process (4), we also have:

$$\frac{dC}{C} = (\mu_c + \hat{x}_t) dt + \sqrt{V} dw_c^*. \tag{25}$$

Combining equations (22) to (25), we get the expression for the SDF:

$$\frac{d\pi}{\pi} = -[\Omega + \hat{x}_t] dt + \sigma_\pi^T \begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix}, \tag{26}$$

where

$$\Omega = h + \mu_c + (\xi_1 - 1)V + \xi_3\sqrt{V}\Sigma_x \begin{bmatrix} 1 \\ \rho_{cd} \end{bmatrix}, \tag{27}$$

$$\sigma_\pi = \begin{bmatrix} (\xi_1 - 1)\sqrt{V} + \xi_3\Sigma_{1x} \\ \xi_3\Sigma_{2x} \end{bmatrix} = -\begin{bmatrix} \gamma\sqrt{V} + \dfrac{\gamma - 1}{h + \rho}\Sigma_{1x} \\ \dfrac{\gamma - 1}{h + \rho}\Sigma_{2x} \end{bmatrix}, \tag{28}$$

Parameters Σ_{1x} and Σ_{2x} are the (1,1) and (1,2) components of Σ_x. Since h and ρ are in general positive, for a risk averse investor with $\gamma > 1$, the pricing kernel is in general conditionally negatively correlated with the LRR component $\hat{x}_t$.

The short rate is determined by the negative of the drift of SDF (26):

$$\begin{aligned} r_t^f &= -\frac{E\left(\dfrac{d\,\pi}{\pi}\right)}{dt} \\ &= \Omega + \hat{x}_t \\ &= h + \mu_c - \gamma V - \frac{\gamma - 1}{h + \rho}\sqrt{V}\,\Sigma_x \begin{bmatrix} 1 \\ \rho_{cd} \end{bmatrix} + \hat{x}_t. \end{aligned} \tag{29}$$

In steady state, for an infinitely lived agent, Ω is constant. Therefore, r_t^f is a function of only $\hat{x}_t$ in the steady state. Equation (29) also implies that the short rate is low when risk aversion γ is large. This is due to the precautionary saving effect. In our model, consumption growth is stochastic, which introduces uncertainty into future utility level and lowers its expected value. As a hedge to this risk, a risk averse investor saves more today. The demand for the risk free asset pushes up the price and makes the risk free rate lower. With a higher risk aversion, γ, investors worry more about the low utility states in the future. Other things equal, investors want to save more today to smooth their utility stream. The end result is a lower risk free rate.

5. Equity Valuation

To illustrate the impact of incorporating incomplete information on equity valuation, we take the profitability-based stock valuation model of Pastor and Veronesi (2003) as an example. Pastor and Veronesi (2003) study the impact of uncertainty in mean profitability growth on stock valuation measured by the price-earnings ratio of a firm's equity. We assume that stock dividends are a fraction of the book value of the firm's equity (b_t):

$$D_t = \delta b_t. \tag{30}$$

The profitability, p_t, is defined as the firm's earnings (Y_t) divided by the book value of equity (b_t):

$$p_t = \frac{Y_t}{b_t}, \tag{31}$$

We further assume that the profitability growth follows a mean-reverting process and is driven by dw_c^* and dw_d^*:

$$dp_t = \phi_p(\bar{p} - p_t)dt + \Sigma_p^T \begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix}. \tag{32}$$

With this assumption, the model is currently dynamically complete. In principle, we can introduce i.i.d. shocks to process (32) for profitability to make the model incomplete. However, since i.i.d. shocks are uncorrelated with the discount factor by assumption, it will not generate any additional instantaneous risk premium.

The process of the book value of equity b_t is given by:

$$db_t = (p_t - \delta)b_t dt. \tag{33}$$

Since we are interested in the impact of investor's learning on the LRR component in consumption and dividend mean, we apply our incomplete information pricing kernel to the complete information version of Pastor and Veronesi (2003)'s model, i.e., we assume that the long-term mean of profitability $\bar{p}$ is assumed constant and known. We now consider an infinite-horizon economy. With the pricing kernel π_t given in equation (26), the equity price P_t should satisfy the following PDE:

$$E\left(\frac{dP}{P}\right) + \frac{\delta b}{P} dt - r dt = -E\left(\frac{dP}{P}\frac{d\pi}{\pi}\right). \tag{34}$$

We look for a time-t solution in the following form:

$$P(b_t, p_t, r_t, t) = \delta b_t Z(t, p_t, r_t),$$

where δb_t is the dividend per share and $Z(t, p_t, r_t)$ is the price-dividend ratio at time *t*.

Using assumptions (30) to (33), the PDE for the stock price takes the following form (see Appendix for derivation)

$$\frac{Z_t}{Z} + \phi_p(\bar{p} - p)\frac{Z_p}{Z} + \rho(\Omega - r)\frac{Z_r}{Z} + \frac{\Sigma_p'\Sigma\Sigma_p}{2}\frac{Z_{pp}}{Z} + \Sigma_p'\Sigma\Sigma_x\frac{Z_{pr}}{Z} + \frac{\Sigma_x'\Sigma\Sigma_x}{2}\frac{Z_{rr}}{Z} \tag{35}$$

$$+ (p - \delta) - r + \frac{1}{Z} + \frac{Z_p}{Z}\lambda_1 + \frac{Z_r}{Z}\lambda_2 = 0,$$

where $\lambda_1 = \Sigma_p'\Sigma\sigma_\pi$, $\lambda_2 = \Sigma_x'\Sigma\sigma_\pi$, σ_π is the instantaneous volatility of the pricing kernel, Σ_x is the instantaneous volatility of the estimated latent process, and $\Sigma = \begin{bmatrix} 1 & \rho_{cd} \\ \rho_{cd} & 1 \end{bmatrix}$.

Given the affine nature of the short rate process, equations (26), (29), and (34) imply that the solution for the stock price has an exponential-affine form as a function of posterior mean of conditional consumption growth, $\hat{x}$ (see Appendix for further details). Since $\hat{x}$ is affine (see (29)) in the short rate, r, then stock price itself is an affine function of the short rate.

Theorem 2 *The real price of a stock paying an infinite stream of dividends as specified in (30) is given by an integral of an exponential-affine function of the state variables:*

$$P(B_t, p_t, r_t, t) = \delta b_t \int_t^{\infty} exp(A(t,s) + B(t,s)p_t - C(t,s)r_t)ds \tag{36}$$

where functions $A(t,s), B(t,s)$, and $C(t,s)$ are given by:

$$B = \frac{1 - exp\left(-\phi_p(s-t)\right)}{\phi_p},$$

$$C = \frac{1 - exp\left(-\rho(s-t)\right)}{\rho},$$

$$A = m_0(s-t) + m_1\left(\frac{1 - exp\left(-\phi_p(s-t)\right)}{\phi_p}\right) + m_2\left(\frac{1 - exp\left(-\rho(s-t)\right)}{\rho}\right)$$

$$+ m_3\left(\frac{1 - exp\left(-2\phi_p(s-t)\right)}{2\phi_p}\right)$$

$$+ m_4\left(\frac{1 - exp\left(-2\rho(s-t)\right)}{2\rho}\right) + m_5\left(\frac{1 - exp\left(-\left(\phi_p + \rho\right)(s-t)\right)}{\left(\phi_p + \rho\right)}\right),$$

and constants m_0, m_1, m_2, m_3, m_4, and m_5 are defined in Appendix.

The real equity risk premium is defined by the negative covariance of the pricing kernel and equity return and has the following form:

$$\lambda = -\left(\frac{Z_p}{Z}\lambda_1 + \frac{Z_r}{Z}\lambda_2\right), \tag{37}$$

where Z_p and Z_r are the first order partial derivatives of the price to dividend ratio on profitability and short rate, respectively, and constants λ_1 and λ_2 are the pricing kernel's conditional covariances with profitability and short rate, respectively, whose expressions are given in the Appendix.

Proof: see Appendix.

Equation (37) shows that the instantaneous equity risk premium has two components: $\lambda_p = -\frac{Z_p}{Z}\lambda_1$ is the risk premium due to shocks to profitability, and $\lambda_r = -\frac{Z_r}{Z}\lambda_2$ is the risk premium due to shocks to the short rate. If both profitability and the short rate (or equivalently the LRR component $\hat{x}$) are mean-reverting $\left(0 < \phi_p, \rho < 0\right)$, both $B(t,s)$ and $C(t,s)$ in (36) are positive, so we have a negative elasticity of the price-dividend ratio to the short rate $\left(\frac{Z_r}{Z} < 0\right)$ and a positive elasticity of the price-dividend ratio to the profitability growth $\left(\frac{Z_p}{Z} > 0\right)$. From our analysis of the pricing kernel, we know that the short rate (the LRR component) is conditionally negatively correlated with the pricing kernel ($\lambda_2 < 0$). That implies a negative λ_r. Therefore, λ_p has to be large and positive for the model to be able to fit the observed equity premium.

To assess model performance, we do a calibration exercise and compute the equity premium in our model. Our parameter vector contains 15 components:

$$\Theta = \{\gamma, h, \rho, \rho_{cd}, \rho_{cx}, \rho_{dx}, \varphi_d, \phi, \varphi_e, V, \bar{p}, \phi_p, \Sigma_p, \mu_c\}.$$

We start with parameters of fundamental processes (i.e. consumption growth, dividend growth, and the LRR component). There are in total 10 parameters in this set: $h, \rho, \rho_{cd}, \rho_{cx}, \rho_{dx}, \varphi_d, \phi, \varphi_e, V,$ and μ_c. $h = exp(c_0 - w_0)$ is the consumption-wealth ratio. Lustig, Van Nieuwerburgh, and Verdelhan (2013) estimate the log wealth-consumption ratio using a wide range of assets over the period of 1952 to 2011[8].

[8] Lustig, Van Nieuwerburgh, and Verdelhan (2013) estimate the wealth-consumption ratio using a data sample containing the Cochrane and Piazzesi (2005) factor (CP), the nominal short rate (yield on a 3-month Treasury bill), realized inflation, the spread between the yield on a 5-year Treasury note and a 3-month Treasury bill, the price-dividend ratio on the CRSP stock market, the real return on the CRSP stock market, the real return on a factor-mimicking portfolio for consumption growth, the real return on a factor-mimicking portfolio for labor income growth, real per capita consumption growth, and real per capita labor income growth. The factor-mimicking portfolios are formed using 25 size- and value-portfolio returns.

According to their annual estimation results, the log wealth-consumption ratio over that period is 4.63, which implies $h = exp(-4.63) = 0.0098$

Constantinides and Ghosh (2014) (CG hereafter) estimate the LRR model with economic uncertainty using the Generalized Method of Moments (GMM) and data on consumption, dividend, CRSP market index, as well as size and book-to-market portfolios from 1929 to 2009. For $\rho, \rho_{cd}, \rho_{cx}, \rho_{dx}, \varphi_d, \phi, \varphi_e, V$, and μ_c, we adopt GC's estimates. Specifically, we set annual consumption growth to 1.9% per year. Leverage parameter, ϕ, and conditional volatility parameters φ_d and V are matched with the corresponding CG's estimates: $\phi = 4.63$, $\varphi_d = 10.1$, and $V = 3.6 \times 10^{-5}$[9]. We calibrate ρ and φ_e by matching both conditional mean and volatility of the LRR component x in CG with corresponding values in our model. Assuming that the economic uncertainty in LRR model in CG is equal to its estimated value, we have $\rho = -ln(\rho_{CG}) = -ln(0.955) = 0.046$ and $\varphi_e = \psi_x \sqrt{\frac{2\rho}{1-\rho_{CG}^2}}$, where ψ_x and ρ_{CG} are given in Table 6 of CG.

We set the correlation between consumption and dividend growth shocks to be its sample estimate, $\rho_{cd} = -0.1007$[10]. We take relative risk aversion to be equal to 8. In the base case, we assume that consumption and dividend shocks are not correlated with conditional consumption mean, x, i.e., $\rho_{cx} = \rho_{dx} = 0$. Note that ρ_{cx} and ρ_{dx} are crucial in our model in the sense that they determine how precisely investors can learn about the latent state variable x. More precisely, the larger the absolute values of these two

[9] All of these parameter values are from table 6 of CG.
[10] Real Personal Consumption Expenditure data series are downloaded from Federal Reserve Bank of St. Louis's website (https://research.stlouisfed.org/fred2/series/DPCERX1A020NBEA). All items are annual and range from 1977 to 2014. Real dividend of S&P 500 data series is from Robert J. Shiller's website (http://www.econ.yale.edu/~shiller/data.htm). All items are annual and range from 1977 to 2014.

correlations, the smaller the posterior variance of the optimal estimate S. We also consider cases with non-zero values of ρ_{cx} and ρ_{dx}.

Parameters $\delta, \bar{p}, \phi_p, \Sigma_p$ are related to the profitability and equity book value processes. We estimate these parameters from the data. Our data sample includes annual book value per share, dividend per share, and earnings per share of S&P 500 Composite Index over the period 1977 to 2014. Book value per share data come from COMPUSTAT[11], and the other data series are from Robert Shiller's website. The profitability p_t is computed as the ratio of earnings Y_t to the equity book value, B_t. The estimation of the AR(1) model for p_t in (32) produces point estimates for $\bar{p}$ and ϕ_p of 12.99% and 0.7724, respectively[12]. The estimated instantaneous standard deviation of p_t, σ_p, is 0.1831 in our sample. We set the first element of Σ_p, profitability growth shocks' loading on consumption growth shocks, $\Sigma_{p,1}$, equal to 0.12[13] in the base case and compute profitability growth shocks' loading on dividend growth shocks, $\Sigma_{p,2}$ based on the following equation

$$\Sigma_p' \Sigma \Sigma_p = \sigma_p^2.$$

Following equation (30), we compute dividend-book ratio δ as the ratio between the real equity book value and real dividend per share of S&P 500 index. The mean dividend-book ratio in our sample is 0.0576.

[11] Book values per share of S&P 500 index are downloaded from WRDS (COMPUSTAT), ranging from 1977 to 2014. We then deflate the series using CPI from Robert J. Shiller's data.
[12] Details of the estimation are given in Appendix.
[13] We choose 0.12 to make sure $\Sigma_{p,2}$ exists and the transversality condition is satisfied. The transversality condition prohibits price bubbles.

We start our calibration exercise from the base case. Specifically, we set $\rho_{cx} = \rho_{dx} = 0$[14]. To match CG's estimate, $\rho_{CG} = 0.955$, in the discrete model, ρ is set to be equal to $-ln(0.955) = 0.046$. Base case results are reported in the first column of Table 2.1. The model-implied real risk premium is 6.66%, which is lower than the values we see in the data[15]. A large proportion of the risk premium ($\lambda_p = 10.09\%$) is related to the shocks to the profitability. The magnitude of risk premium is jointly determined by its covariance with the SDF and the price-dividend ratio's elasticity with respect to profitability. In our case, the short rate (the LRR component) is more persistent than profitability, shocks to the short rate have much more profound impact on the price-dividend ratio $\left(|Z_r| > |Z_p|\right)$. However, this is more than compensated for by the fact that profitability is much more volatile than the short rate. Since we assume that both processes are driven by the same set of innovations, the covariance between profitability and the SDF is much larger (in scale) than that between the short rate and the SDF. Another interesting observation here is that the risk premium corresponding to shocks in the short rate is negative. One can see the reason for the negative compensation for the discount rate risk from the price function (36) and risk premium function (37). The short rate co-varies negatively with the SDF ($\Sigma_x' \Sigma \sigma_\pi < 0$), and the elasticity of the price-dividend ratio to the short rate is negative $\left(\frac{Z_r}{Z} < 0\right)$. The combination of these two facts leads to a negative discount rate risk premium. The estimate of price to dividend ratio (199.14) is much larger than the sample average (43.77 ranging from 15.92 to 88.12). The model-implied real interest estimate is 2.61%. Given that inflation over the post-war

[14] In our calibration, we assume $\rho_{cx} = \rho_{dx}$. These two correlations can be different in practice. We formally estimate these correlations in chapter three of this thesis.
[15] The annual equity premium in the period of 1982 to 2013 is around 8.76%. Given the inflation in the post-war period has a mean around 3.6%, the implied real risk premium is roughly 8.46%.

period has averaged about 3.6%, the nominal interest rate implied by model is about 6.21%, which is again larger than the observed sample average of 4.27% over the post-war period. The large model-implied short rate is consistent with findings of CG. CG's estimation results show that, over the sample from 1929 to 2009, the LRR model (with economic uncertainty) of Bansal and Yaron (2004) produces significantly larger real short rate than that observed in the data.

Table 2.1 Calibration results with different ρ_{cx}, ρ_{dx}.
This table reports calibrated results of real equity risk premium λ, the price to dividend ratio, Z, the real risk free rate, r^f, the risk premium fraction related to shocks in profitability, Z_p, and the risk premium fraction related to short rate, Z_r, for different values of ρ_{cx}, ρ_{dx}, and $\bar{p}$. We use the following values of the other parameters: $\bar{p} =$ 12.99%, $\phi_p = 0.7724, \phi = 4.63$, $\varphi_d = 10.1,$ $V = 3.6 \times 10^{-5}$, $\mu_c = 1.9\%$, $\gamma = 6,$ $h = \exp(-4.63)$, $\rho_{cd} = -01007,$ $\rho = -ln(\rho_{CG}),$ $\varphi_e = \psi_x\sqrt{\frac{2\rho}{1-\rho_{CG}^2}}$, where $\rho_{CG} = 0.955$ and $\psi_x = 0.616$. In both non-zero correlation cases, $\Sigma_{p,1}$ is 0.12 to satisfy the transversality condition.

	$\rho_{cx} = \rho_{dx} = 0$	$\rho_{cx} = \rho_{dx} = 0.2$	$\rho_{cx} = \rho_{dx} = 0.6$
λ	6.66%	7.49%	8.55%
Z	199.14	76.08	41.82
r^f	2.61%	2.62%	2.65%
λ_p	10.09%	10.57%	11.32%
λ_r	-3.43%	-3.08%	-2.77%

In table 2.1, we also report calibration results for different values of learning parameters ρ_{cx} and ρ_{dx}. In the third column, we show model predictions with $\rho_{cx} = \rho_{dx} = 0.2$. In this case, the real risk premium increases by 83 bps to 7.49% compared to that in the base case. Again, the risk premium is dominated by the risk premium related to the shocks in profitability (10.57%) and the premium for the discount rate risk is still negative (-3.08%). Larger values of learning parameters imply higher covariances of profitability and short rate with the SDF[16], and simultaneously make the price-dividend ratio less sensitive to changes in profitability and the short rate (measured by $\frac{Z_p}{Z}$ and $\frac{Z_r}{Z}$, respectively). Overall, the profitability contribution becomes larger and the short rate contribution is less negative. The nominal risk free rate is roughly the same as that in the base case. The price-dividend ratio drops by more than 50% compared to that in the base case. As we increase the correlations to $\rho_{cx} = \rho_{dx} = 0.6$, the model-implied price-dividend ratio decreases by another 45%. More importantly, the equity premium increases further by 1.06% to 8.55%, which is now in line with what we observed in the data. The equity premium is again dominated by the contribution from the profitability risk (11.32%) with a smaller negative contribution from the discount rate risk (-2.77%). Even though the focus here is simply on the theoretical impact of the conditional correlations of the LRR component with consumption and dividend growth on the equity premium, in chapter three of this thesis, we do estimate the marginal posterior densities of these correlations. The posterior medians of ρ_{cx} and ρ_{dx} are 0.2225 and 0.3820, respectively[17]. If we use these estimated values in calibration, we get the real risk premium of 7.53%, the price-dividend ratio of 24.24, and the real short rate (2.65%), which is the moment that the LRR model fails to match. Given the 2.5% and 97.5% band of ρ_{cx} and ρ_{dx} ([0.1702, 0.2973] and [0.0827,0.6878]), if we assume ρ_{cx} and ρ_{dx} are uncorrelated, we can compute the 2.5% and 97.5% bands of both risk premium and price-dividend ratio. When ρ_{cx} and ρ_{dx} equal their 2.5% (97.5%) band values, the model-implies real risk premium is 5.53% (8.65%) and the price-dividend ratio is 49.26 (18.69).

[16] These covariances are in general non-monotonic functions of learning parameters. However, if the other parameters are fixed at their calibrated values, these covariances are increasing in learning parameters within the range of values we consider.
[17] Both correlations are estimated in chapter three of this thesis.

For both risk premium and price-dividend ratio, sample estimates fall within the 2.5% and 97.5% band of model estimates.

It is worth noting that our estimates of real risk premium and price-dividend ratio are highly sensitive to value of unconditional mean of profitability, $\bar{p}$. For instance, if we assume the true mean profitability growth $\bar{p} = 0.1$[18], the smallest real equity premium the model can produce is around 3.5%[19] per year in the base case. Note that the risk premium is increasing while the risk free rate is decreasing with risk aversion. The model is flexible enough with respect to the risk aversion parameter, γ. For example, if we set $\gamma = 25$, the model can produce a real short rate of 1.99%, which is well within the 90% uncertainty bands of the average short rate in data[20]. Meanwhile the model-implied equity premium ranges from 5.89% to 22.86%, and the price-dividend ratio is between 22.86 and 70.24[21]. Again, for both the risk premium and price-dividend ratio, sample estimates fall within the 2.5% and 97.5% bands of the model estimates.

6. Conclusion

In this Chapter, we derive a model of equity valuation in a LRR economy. We extend the LRR model of Bansal and Yaron (2004) by including investors' learning about the LRR component in mean consumption/dividend growth. We model investors learning process and derive an equilibrium pricing kernel based on investors' optimal estimate of the latent state variable process. We show that there is a closed form solution for equity

[18] Given other parameter values, 0.1 is well within the 50% uncertainty bands of estimated mean profitability growth.

[19] With $\Sigma_{p,1}$ around 0.057.

[20] For a period from 1977 to 2014, the average inflation is 3.79%.Together with a 1.99% real short rate, it implies a 5.78% nominal short rate. Over the same period, the mean nominal risk free rate is 4.98% with the 90% uncertainty bands [3.99%,5.96%].

[21] With $\Sigma_{p,1} = 0.0165$, ρ_{cx} is from 0.1702 to 0.2973, and ρ_{dx} from 0.0827 to 0.6878.

prices and the instantaneous equity premium under the new risk-neutral measure corresponding to the information available to investors.

We derive an equilibrium pricing kernel that contains all pricing information in the incomplete information environment. In our model, since the pricing kernel reflects incomplete information about fundamentals in equilibrium, the estimation risk of the unobserved conditional mean of consumption growth affects all risk prices in the pricing kernel and, consequently, all contributions to the equity premium.

Further, in contrast to previous work in this line of literature, in our model the short rate is determined endogenously from the drift of the stochastic discount factor. As a result, the short rate is an affine function of the estimated conditional mean of consumption growth and, thus, inherits its mean-reverting property. One of the interesting aspects of the endogeneity of the short rate is its negative contribution to the overall equity premium.

We apply our SDF to the base case model of Pastor and Veronesi (2003). The equity premium in the combined model has two components: the component corresponding to profitability risk and the contribution of the discount rate risk. We calibrate the combined model to aggregate consumption, dividends, earnings, and book value processes for S&P 500 composite index. With a risk aversion parameter of 8, the model is able to reproduce the level of real equity premium (8.55%) and price to dividend ratio (41.82) when the LRR component is positively correlated with both consumption and dividend growth. The model implied short rate is around 2.65% per year. These levels are sensitive to parameter values such as mean profitability growth. However, within the uncertainty bands of mean profitability growth, the model has the potential to fit levels of all three variables of interest in the data.

Our work in this chapter is extendable in the following dimensions: First, our application to the base case of Pastor and Veronesi (2003) is only a calibration exercise. Thus, all conclusions are tentative. To see whether the model indeed has a good fit to the actual data or not, we need to formally estimate the model using the real data, which we do in chapter three. Second, the model can be applied to the valuation of any asset in the modeled economy. In chapter three, we examine the model's ability to explain the observed real and nominal term structure of interest rates. Last but not least, we may include other mechanisms, such as economic uncertainty and jumps, into the current model and study the impact of learning on asset pricing in the extended model albeit at the cost of losing closed form solutions.

A2.1 Derivation of the Value Function J

With the log-linear approximation (16), the PDE (11) is a parabolic PDE with coefficients affine in states, which implies an exponential affine solution to the PDE:

$$J(t, g_c, g_d, \hat{x},) = \exp\{\xi_{0t} + +\xi_{1t} g_c + \xi_{2t} g_d + \xi_{3t}\hat{x}\}. \tag{38}$$

Upon the substitution of (38) into (11) the PDE becomes an identity that must hold for arbitrary values of state variables:

$$\partial_t \xi_{0t} + G^T \partial_t \begin{bmatrix} \xi_{1t} \\ \xi_{2t} \\ \xi_{3t} \end{bmatrix} + \hat{\mu}_G^T \begin{bmatrix} \xi_{1t} \\ \xi_{2t} \\ \xi_{3t} \end{bmatrix} + \frac{1}{2dt} tr\left(dGdG^T \frac{J_{GG}}{J}\right) + \frac{f}{J} = 0. \tag{39}$$

Since $\frac{1}{2dt} tr(dGdG^T)$ is a function of only diffusion parameters and $\frac{1}{2dt} tr\left(\frac{J_{GG}}{J}\right)$ is a function of ξs, the fourth term on the left-hand side of PDE (39) contains no state variables.

With the log-linear approximation of the normalized aggregator given by equation (16), we also have

$$\frac{f}{J} = h(1-\gamma)\left[H - \frac{\xi_{0t}}{1-\gamma} + G'\begin{pmatrix} 1 - \frac{\xi_{1t}}{1-\gamma} \\ -\frac{\xi_{2t}}{1-\gamma} \\ -\frac{\xi_{3t}}{1-\gamma} \end{pmatrix}\right]. \tag{40}$$

Since equation (39) must hold for all values of state variables, coefficients on the state variables must all be zero. Therefore, after collecting terms containing state variables, we have the following identity:

$$G^T\left[\partial_t \begin{bmatrix} \xi_{1t} \\ \xi_{2t} \\ \xi_{3t} \end{bmatrix} + \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & \phi \\ 0 & 0 & -\rho \end{pmatrix}^T \begin{bmatrix} \xi_{1t} \\ \xi_{2t} \\ \xi_{3t} \end{bmatrix} + h(1-\gamma)\begin{pmatrix} 1 - \frac{\xi_{1t}}{1-\gamma} \\ -\frac{\xi_{2t}}{1-\gamma} \\ -\frac{\xi_{3t}}{1-\gamma} \end{pmatrix}\right] = 0. \tag{41}$$

Identity (41) implies the following ODEs

$$\partial_t \xi_{1t} + h(1-\gamma) - h\xi_{1t} = 0, \tag{42}$$

$$\partial_t \xi_{2t} - h\xi_{2t} = 0, \tag{43}$$

$$\partial_t \xi_{3t} + \xi_{1t} + \phi\xi_{2t} - (\rho + h)\xi_{3t} = 0, \tag{44}$$

subject to boundary condition

$$\xi_{1t}(T) = \xi_{2t}(T) = \xi_{3t}(T) = 0 \tag{45}$$

Solving the system (42) – (44) for an infinitely lived agent, we obtain solutions (18) – (20).

A2.2 Derivation of the stock price, P_t

We derive the share price using standard SDE arguments based on a stochastic discount factor (SDF) approach.

With the pricing kernel given in (26), the stock price must satisfy the following PDE:

$$E\left(\frac{dP}{P}\right) + \frac{\delta B}{P}dt - rdt = -E\left(\frac{dP}{P}\frac{d\pi}{\pi}\right). \tag{46}$$

We look for a time-t stock price solution in the following form:

$$P(B_t, p_t, r_t, t) = \delta b_t Z(t, p_t, r_t),$$

where δb_t is the dividend per share and $Z(t, p_t, r_t)$ is the price-dividend ratio at time *t*.

Applying Itô's rule to P_t:

$$\frac{dP}{P} = \frac{dZ}{Z} + \frac{db}{b} + \frac{dZ}{Z}\frac{db}{b}$$

$$= \frac{1}{Z}\left(Z_t dt + Z_p dp + Z_r dr + \frac{1}{2}\left(Z_{pp}dp^2 + 2Z_{pr}dpdr + Z_{rr}dr^2\right)\right)$$

$$+ (p - \delta)dt,$$

$$\frac{dP}{P}\frac{d\pi}{\pi} = \left(\frac{Z_p}{Z}\Sigma_p'\Sigma\sigma_\pi + \frac{Z_r}{Z}\Sigma_x'\Sigma\sigma_\pi\right)dt.$$

Collecting all the terms in (46), taking the expectation, and dividing throughout by *dt*, we arrive at the PDE for the equity price:

$$\frac{Z_t}{Z} + \phi_p(\bar{p} - p)\frac{Z_p}{Z} + \rho(\Omega - r)\frac{Z_r}{Z} + \frac{\Sigma_p'\Sigma\Sigma_p}{2}\frac{Z_{pp}}{Z} + \Sigma_p'\Sigma\Sigma_x\frac{Z_{pr}}{Z} + \frac{\Sigma_x'\Sigma\Sigma_x}{2}\frac{Z_{rr}}{Z}$$

$$+ (p - \delta) - r + \frac{1}{Z} + \frac{Z_p}{Z}\lambda_1 + \frac{Z_r}{Z}\lambda_2 = 0,$$

where $\lambda_1 = \Sigma_p' \Sigma \sigma_\pi$, $\lambda_2 = \Sigma_x' \Sigma \sigma_\pi$, σ_π is the instantaneous standard deviation of the pricing kernel, Σ_x is the instantaneous volatility of the estimated latent process, and $\Sigma = \begin{bmatrix} 1 & \rho_{cd} \\ \rho_{cd} & 1 \end{bmatrix}$.

The above PDE satisfies Feynman-Kac conditions, and therefore allows us to write the solution as follows:

$$Z(t,s,p,r) = \int_t^{\infty} exp(A(t,s) + B(t,s)p_t - C(t,s)r_t)ds. \tag{47}$$

Inserting (47) into the PDE (46), together with the fact that PDE (46) should hold for arbitrary values of p_t and r_t, we get the following ordinary differential equations for functions $A(t,s)$, $B(t,s)$, and $C(t,s)$:

$$B_t - \phi_p B + 1 = 0,$$

$$C_t - \rho C + 1 = 0,$$

$$A_t + \phi_p \bar{p} B - \rho \Omega C + \frac{\Sigma_p' \Sigma \Sigma_p}{2} B^2 - \Sigma_p' \Sigma \Sigma_x BC + \frac{\Sigma_x' \Sigma \Sigma_x}{2} C^2 - \delta + \lambda_1 B - \lambda_2 C = 0,$$

with the initial conditions $A(s,s) = 0, B(s,s) = 0$, and $C(s,s) = 0$.

The solution to the ODE system is:

$$B = \frac{1 - exp\left(-\phi_p(s-t)\right)}{\phi_p},$$

$$C = \frac{1 - exp\left(-\rho(s-t)\right)}{\rho},$$

$$A = m_0(s-t) + m_1\left(\frac{1-exp\left(-\phi_p(s-t)\right)}{\phi_p}\right) + m_2\left(\frac{1-exp\left(-\rho(s-t)\right)}{\rho}\right)$$

$$+ m_3\left(\frac{1-exp\left(-2\phi_p(s-t)\right)}{2\phi_p}\right)$$

$$+ m_4\left(\frac{1-exp\left(-2\rho(s-t)\right)}{2\rho}\right) + m_5\left(\frac{1-exp\left(-\left(\phi_p+\rho\right)(s-t)\right)}{\left(\phi_p+\rho\right)}\right),$$

where

$$m_0 = \frac{1}{\phi_p^2\rho^2}\left(\frac{\Sigma_p'\Sigma\Sigma_p}{2}\rho^2 + \frac{\Sigma_x'\Sigma\Sigma_x}{2}\phi_p^2 + \phi_p\rho\left(\rho\lambda_1 - \phi_p\lambda_2\right) + (\bar{p}-\delta-\Omega)\phi_p^2\rho^2\right.$$

$$\left. - \Sigma_p'\Sigma\Sigma_x\phi_p\rho\right),$$

$$m_1 = -\frac{1}{\phi_p^2\rho}\left(\Sigma_p'\Sigma\Sigma_p\rho + \bar{p}\phi_p^2\rho - \Sigma_p'\Sigma\Sigma_x\phi_p + \phi_p\rho\lambda_1\right),$$

$$m_2 = \frac{1}{\phi_p\rho^2}\left(-\Sigma_x'\Sigma\Sigma_x\phi_p + \Sigma_p'\Sigma\Sigma_x\rho + \Omega\phi_p\rho^2 + \phi_p\rho\lambda_2\right),$$

$$m_3 = \frac{1}{\phi_p^2}\frac{\Sigma_p'\Sigma\Sigma_p}{2},$$

$$m_4 = \frac{1}{\rho^2}\frac{\Sigma_x'\Sigma\Sigma_x}{2},$$

$$m_5 = -\frac{1}{\phi_p\rho}\Sigma_p'\Sigma\Sigma_x.$$

For the integral (47) to exist, the integrand should be declining with s sufficiently fast. Since B and C are bounded, a transversality condition on the model parameters then requires that function A be negative and unbounded at large s. To derive the transversality condition, we notice that any term containing an expression of the form

$\frac{1-exp(-\vartheta(s-t))}{\vartheta}$ with positive ϑ is bounded. After ignoring all bounded and collecting all unbounded terms, we have

$$A \xrightarrow{leading\ terms} m_0(s-t).$$

The transversality condition requires that the leading terms must be negative for the price integral to exist:

$$m_0 < 0.$$

A2.3 Estimating Parameters $\bar{p}, \phi_p, \Sigma_p$ in Profitability Process (32)

Following equation (31), we compute the real annual profitability of the S&P 500 index as the ratio between the real annual earnings per share and the real book value per share at the end of each year. The real monthly earnings (annualized) data series are available on Robert J. Shiller's website together with the CPI used to deflate the earnings series. To find the real annual earnings at period t, we take the average of earnings from period $t-11$ to period t. The book values of S&P 500 index data series are downloaded from COMPUSTAT and deflated by the same price index as the annual earnings series. Table A2.3.1 reports the basic characteristics of real earnings, real book value, and real profitability.

Table A2.3.1 Summary statistics for real earnings, book value per share, and real profitability.
This table reports the summary statistics of real annual earnings per share (Y_t), real book value per share (b_t), and real annual profitability(p_t) of S&P 500 index. The sample period is 1977 to 2014. Std. error is the sample standard deviation. ρ is the first-order autocorrelation coefficient.

Variable	Y_t	b_t	p_t
Mean	51.38	399.31	13.02%
Std.error	20.45	141.31	2.62%
Min.	18.13	271.12	3.22%
Max.	93.79	728.23	16.77%
ρ	0.7141	0.9824	0.4621

To estimate the profitability process specified by equation (32), we first discretize the process as:

$$p_t = \phi_p^d \bar{p} + \left(1 - \phi_p^d\right) p_{t-1} + \sigma_p \epsilon_t, \tag{48}$$

where

$$\phi_p^d = 1 - exp\left(-\phi_p\right),$$
$$\sigma_p = \Sigma_p' \Sigma \Sigma_p / \sqrt{\frac{2\phi_p}{1 - exp\left(-2\phi_p\right)}}$$

With the time series of real annual profitability, we estimate the AR(1) model (48) by regressing the profitability on its lag 1 value and a constant. Table A2.3.2 summaries the OLS estimation results.

Table A2.3.2 Estimation results of profitability process (32).
This table reports estimation results of profitability process specified by equation (48). The data sample contains real annual profitability (p_t) of S&P 500 Composite Index from 1977 to 2014. $\hat{\sigma}$ is the estimated standard deviation of innovations, which is computed as $\hat{\varepsilon}'\hat{\varepsilon}/36$, where $\hat{\varepsilon}$ is the vector of residuals.

	Coefficient	Standard Error	t-stat
Intercept	0.0699	0.0202	3.46
p_{t-1}	0.4619	0.1520	3.04
R^2	0.2135	$\hat{\sigma}$	0.1306

Given the estimation results in table A2.3.2, we have

$$\left(1 - \phi_p^d\right) = 0.4619,$$
$$\phi_p^d \bar{p} = 0.0699,$$
$$\sigma_p = 0.1306.$$

Recall that

$$\left(1 - \phi_p^d\right) = exp\left(-\phi_p\right),$$
$$\Sigma_p' \Sigma \Sigma_p = \sigma_p \sqrt{\frac{2\phi_p}{1 - exp\left(-2\phi_p\right)}}$$

These constraints allow us to solve for ϕ_p, $\bar{p}$, and $\Sigma_p' \Sigma \Sigma_p$:

$$\phi_p = 0.7724,$$
$$\bar{p} = 0.1299,$$
$$\Sigma_p' \Sigma \Sigma_p = 0.1831.$$

Chapter Three: Term Structure under Long-Run Risks and Incomplete Information

3.1 Introduction

In this Chapter, we develop and estimate a dynamic model of the term structure of interest rates based on the long run risks (LRR) model (see, e.g., Bansal and Yaron (2004), Piazzesi and Schneider (2007), Bansal, Gallant and Tauchen (2007), Eraker (2008), Hasseltoft (2012), and Bansal and Shaliastovich (2012) among others). Our model allows for learning about the unobservable mean-reverting mean of both consumption and dividend growth rates. Our model setup enables us to study how the shape of the yield curve (more specifically, the slope) depends on extra risk premiums induced by learning about the conditional mean of consumption and dividend growth[22] and correlations between the observed and latent state variables. In the empirical part, we estimate model parameters as well as latent variables jointly using a Bayesian Markov Chain Monte Carlo (Bayesian MCMC) approach. Since the Bayesian MCMC approach produces marginal density of model parameters and latent variables, it allows us to study the tension between the time series properties of the state variables and the cross sectional properties of yields.

Previous work on LRR models assumes that the long-run risk component - the conditional mean of consumption growth - is observed by investors. However, in practice, investors have to learn the value of the conditional consumption mean from the observed information. To accommodate this feature of reality, we model how investors learn about the latent variable. We assume that investors minimize the mean square error of the

[22] We later use terms ‘the conditional mean’ and ‘the LRR component’ interchangeably.

estimate of the latent state variable based on observations on consumption and dividend growth. When volatility of the latent state variable is constant (corresponding to economic uncertainty being turned off in the model of Bansal and Yaron (2004)), the model admits closed form solutions for bond prices and yields.

Following Eraker (2008) and Piazzesi and Schneider (2007), we incorporate inflation into our model, with the inflation rate and its mean following an arithmetic Brownian motion and a mean-reverting process, respectively. Under these conditions, the resulting nominal pricing kernel contains all necessary nominal pricing information in the incomplete information environment. Since the filtered LRR component is a linear combination of observed state variables, the priced risk space is effectively defined by only four shocks (those on consumption and dividend growth, inflation, and mean inflation, respectively) in the incomplete information environment. While the utility is only a function of consumption growth, other shocks enter utility and later the SDF indirectly. Specifically, mean inflation and the LRR component show up in utility as components of consumption growth. Shocks in the LRR component then bring in the shocks in dividend growth and inflation as a result of filtering. Inflation shocks also enter via the nominalization of the SDF. Now all four shocks appear in the utility function and the SDF and jointly determine the total risk premium level. The model-implied nominal short rate is jointly determined by five terms: the average consumption-wealth ratio, the mean consumption growth, the mean inflation, the premium on consumption growth risk, and the premium on the inflation risk. The premium on consumption growth risk term is closely related to precautionary savings. The risk-free asset serves as a hedge for the risk in utility induced by uncertainty in consumption growth. As a result, this premium is in general negative. The inflation risk premium, however, depends on the conditional covariance between shocks in inflation and the nominal SDF.

Just like their counterparts in the real model, both the nominal pricing kernel and the nominal short rate are endogenous in a sense that they are the outcomes of a joint solution of the investors' recursive utility optimization and the filtering problem – the estimation of the latent state. As a result, risk premiums are directly tied to parameters of the state processes, which can be estimated using only data on fundamentals and inflation (with the exception of the risk aversion parameter whose estimation, in principle, would require price data). This endogeneity imposes more stringent empirical restrictions on our model than in the case of models with exogenous pricing kernel. In particular, given parameter estimates (and, thus, pricing kernel) based only on information from the time series of fundamentals, we can study price implications, such as the shape of the yield curve, of that information.

To meaningfully pursue this goal, we need to develop an estimation procedure of the model's parameters and state variables that allows studying the empirical behavior of various functions of these parameters and state variables such as the model-implied nominal yield curve. In particular, we want to see whether the model can still generate the upward-sloping nominal yield curve as in Eraker (2008) and how the slope depends on the correlation between mean inflation and real growth rates in fundamentals. Further, we want to study the impact of learning about mean consumption growth on the shape of the nominal yield curve and on the ability of the model to fit yields at different maturities. To handle these tasks, we propose a Bayesian estimation of the model, which does allow us to construct posterior bands on any functions of model parameters and state variables.

Our model is complex in that it contains a large number of unknown parameters and several latent state variables. As a result, we can no longer use a deterministic maximum likelihood method such as the Kalman filter (e.g., De Jong (2000)) to estimate the model. Instead, we turn to numerical maximum likelihood methods. The Bayesian approach compares favourably, e.g., to the Kalman filter, even if it is feasible, which estimates parameters conditional on state variable estimates, thus ignoring sampling variation in those state variables. Under Bayesian Markov Chain Monte Carlo methods (Bayesian MCMC) all parameters and state variables are estimated simultaneously so that any sampling variation in the state variables is integrated out of the parameter estimates.

Gibbs sampler is our estimation method of choice. It is easy to implement but requires all posteriors to be standard. One can easily obtain standard posteriors by the Euler discretization of the state equations resulting in Gaussian distributions of state variables and those of drift parameters and the inverse-Wishart distribution of variance matrices. The most efficient algorithm for latent state estimates is the forward filtering backward sampling (FFBS) algorithm, which allows for simultaneous sampling of all latent states (see West and Harrison (1997) and Kim and Nelson (1999)).

We analyze the nominal yield curve based on our LRR incomplete-information pricing kernel. To estimate the model we use monthly data on aggregate consumption and dividends on S&P 500 Composite Index as well as Consumer Price Index (CPI). In this way, parameters and latent state variables are estimated outside the model, i.e., we use no price information for parameter and state variable inference.

The estimation results suggest that the LRR component is mildly persistent with a median monthly persistence parameter of 0.9297. The inflation mean, however, is close to a random walk at the annual frequency. More importantly, the LRR component is moderately positively correlated with consumption and dividend growth (0.2225 and 0.3820, respectively), while mean inflation is positively correlated with inflation (0.7549). The innovations in the two latent states variables, the mean consumption growth and mean inflation, are negatively correlated (with a median of -0.3584).These correlations, together with other correlations between state variables, are crucial in our model in the sense that they jointly determine the learning efficiency, level of short rate, level of risk premium, and term structure properties. However, in previous studies, these correlations are largely ignored. For instance, in Bansal and Yaron (2004) (and follow-up papers such as Kiku (2006), Drechsler and Yaron (2010), Bansal, Kiku and Yaron (2011), Bansal and Shaliastovich (2012), and Doh (2013)), observed and latent state variables are assumed to be uncorrelated. Piazzezi and Schneider (2006) assume that latent variables and state variables are driven by the same set of innovations, effectively assuming perfect correlations. In contrast, we estimate these correlations and the estimates suggest a negative conditional covariance between the pricing kernel and the LRR component and a positive conditional covariance between the pricing kernel and the mean inflation.

With parameter estimates at hand, we compute the model-implied nominal yield curve with or without information uncertainty. In the complete information case, a larger proportion of the risk premium is due to the hedging demand induced by the variation in the LRR component, with a minor contribution from hedging demand due to shocks in the mean inflation. Since the zero-coupon bond return loads negatively on LRR component and mean inflation, the LRR component risk premium is negative and the mean inflation risk premium is positive. Combining two risk premiums we get a negative total risk premium and a downward sloping nominal yield curve. For a 1-year zero-coupon bond, the LRR component risk premium has a median of -23 bps per month, and the mean inflation risk premium is just 1.9 bps per month, and therefore the total risk premium is -21.3 bps.

The slope of yield curves doesn't change sign after we incorporate learning into the model. However, learning reduces the LRR component's conditional covariance with the pricing kernel (from -0.0291% to -0.0286%) and has a positive, although minor, overall impact on the total risk premium (from -21.3 bps to -20.9 bps for the 1-year zero coupon bond). When we turn on the uncertainty (and learning), another result is a reduction in the mean short rate. The short rate with or without complete information is 0.6357% or 0.6280%, respectively. The reduction is mainly due to changes in the sum of precautionary saving effect and adjustment for inflation in the two models (53.75 bps and 52.82 bps, respectively). The combination of these two effects shifts the yield curve downward and makes it flatter. The intuition behind these changes is straightforward. Since the mean consumption growth is not observed, investors estimate it by projecting its innovations on the priced innovations in the observed state variables. In general, the estimated mean consumption growth has different conditional correlations with observed state variables than conditional correlations of true mean consumption growth with observed state variables. Since both the short rate and the risk premium are functions of these correlations, their estimates change with changes in the correlations.

In short, our study provides empirical evidence against the homoscedastic version of the LRR model (without economic uncertainty) with inflation. If the model parameters are estimated outside the model without using the price data, the LRR model produces a negative risk premium and a downward sloping yield curve. Once we add incomplete information about conditional consumption mean and learning to the LRR model, the total risk premium becomes less negative and the yield curve flattens. However, the impact of information uncertainty on the risk premium is small and insufficient to fundamentally change the overall model implications for the term structure of interest rates.

Our study is closely related to Piazzesi and Schneider (2007), Eraker (2008), and Doh (2012), who also study the nominal yield curve in the LRR framework. Our model is distinctly different from Piazzesi and Schneider (2007), Eraker (2008), and Doh (2012) along two important dimensions – theoretically and empirically.

First, in our model, the latent variable (the LRR component) is unobserved and requires estimation by investors. LRR component innovations are different but, in general, correlated with the innovations in the observed variables. Under the incomplete information assumption, the filtered LRR component is a linear combination of observables. This leads to correction of the risk premiums of the corresponding observables (such as consumption growth in our case). This correction, however, depends on the posterior error variance. In Eraker (2008), latent variables are assumed to be known. In Piazzesi and Schneider (2007) latent variables are just linear combinations of observables. In other words, the posterior error variance is zero in the steady state. Therefore, the steady state correction to the risk premium is small and the impact of incomplete information can only occur when there is a structural change. At another extreme, Doh (2012) assumes innovations in latent states and observables are uncorrelated. With such independence assumption, learning is most inefficient, which implies a large posterior error variance and a large risk premium correction for information uncertainty.

Second, on the empirical side, we structure our estimation approach so that all model parameter estimates with the exception of the risk aversion parameter are based only on the information from the time series of the state variables under the actual measure. We use no price data (such as yields) to estimate these parameters. As a consequence, we can study whether the properties of the model-implied nominal yield curve are in agreement with the properties of the state variables. In that sense, our empirical results, at least in part, follow recommendations of Lewellen, Nagel, and Shanken (2010).

The rest of the paper is organized as follows. In section 2, we review the related literature on LRR term structure and Bayesian estimation of state space models. In section 3, we state the assumptions of the model, derive the optimal posterior estimate of the latent conditional mean of consumption growth, and solve for the value function jointly with learning problem. Further, in section 4 we derive the pricing kernel (stochastic discount factor) with adjustment for learning. Section 5 lays out theoretical details of the term structure. In section 6 we introduce the estimation approach. In section 7, we report the model estimation results using data on aggregate consumption and dividend as well as Consumer Price Index (CPI). Finally, we conclude in section 8.

3.2 Related Literature

3.2.1 Term Structure in LRR Economy

While Bansal and Yaron (2004) show that the real yield curve in the LRR economy is downward sloping, it is still unclear whether the LRR model can reproduce the correct level and shape of the nominal yield curve. To answer this question, several studies incorporate inflation into the LRR model and examine the properties of the model-implied nominal yield curve; among others, these include Piazzesi and Schneider (2007), Eraker (2008), Doh (2012), Hasseltoft (2012), and Bansal and Shaliastovich (2013).

Piazzesi and Schneider (2007) (PS) consider a homoscedastic LRR model (no economic uncertainty) in which representative investor's belief about state variables, consumption growth and inflation, have mean-reverting means, and are driven by the same set of innovations as their latent means. With IES=1, PS show that the demeaned nominal yields are equal to expected demeaned nominal consumption growth rates over the lifetime of the bond.

PS estimate the state variable processes on aggregate consumption growth and inflation data ranging from 1952:2 to 2005:4 by using maximum likelihood estimation. They find that the perceived mean consumption growth has a negative loading on the lagged perceived mean inflation. Intuitively, this means that the representative investor considers inflation as a negative indicator of future consumption growth. In order to reproduce an upward-sloping nominal yield curve, PS pick a large discount factor $(\beta = 1.005)$ and a high risk aversion $(\gamma = 59)$. In subsequent exercises, PS impose further assumptions on their model: 1) investor's information set also includes short rate and term spread; 2) investor's expectations of state variables at period t are formed based only on previous observations of state variables, and he/she pays more attention to more recent observations; and 3) the investor is uncertain about his/her estimation of mean μ_z. To incorporate assumption 1, PS estimate an extended model with short rate and term spread as extra state variables. Assumption 2 is handled by running a sequential estimation with a constant gain adaptive learning scheme. Finally, PS incorporate assumption 3 into the model estimation by treating μ_z as an extra latent variable. PS find that assumption 1 and/or 2 alone don't have much impact on the model implied yield. However, with all 3 assumptions, the model is able to capture the sluggish behavior in interest rates in early 1980s.

PS assume a perfect correlation between latent and state variables. In this regard, their model can be considered a special case of our model. Learning about unobserved states leads to the correction to the risk premiums of observables. The magnitude of the correction, however, depends on the posterior error variance, which is a function of correlations between unobserved and observed states. A perfect correlation implies the posterior error variance is zero in steady state (investors can learn the unobserved states perfectly), and the impact of learning on the total risk premium is zero. This explains why PS find that learning does not affect the model implication significantly unless there is a structural change in the model. In our model, without the perfect correlation assumption, the posterior error variance is non-zero in general. The impact of learning on the risk premium is then more profound even in the steady state.

Doh (2012) estimates an adjusted version of PS's model using a Bayesian methodology with the same dataset as that of PS but from 1953:Q1 to 2006:Q4. In Doh's model, innovations on latent and observed state variables are independent and latent variables' variances are assumed to be stochastic. Since in the LRR model bond yields can also be written as affine functions of latent state variables, Doh includes term structure data as extra state variables in model estimations. Doing so allows Doh to estimate macro-economic parameters such as subjective discount factor, IES, and risk aversion simultaneously with other model parameters. IES and risk aversion are estimated to be around 1.02 and 9.5, respectively. Mean consumption growth and mean inflation appear persistent. Also, posterior means of yield curve moments provide a good fit to corresponding sample moments.

Bansal and Shaliastovich (2013) extend PS's model by assuming that mean consumption growth and mean inflation have stochastic volatilities. Together with the Epstein-Zin recursive utility, the model implies affine real and nominal yields. BS estimate the model on quarterly observations of nominal yields of one to five years to maturity and the forecasts of real consumption growth and inflation from the Survey of Professional Forecasts from 1969 to 2010 by using the maximum likelihood method. By setting the mean and standard deviation of consumption growth and inflation to corresponding values in the data, and fixing the subjective discount factor at 0.994, BS show that the model is able to not only produce state variables and a yield curve that are in line with the data but also provide a good fit to the bond yield predictability feature in the data.

Bansal and Shaliastovich (2013) is echoed by Hasseltoft (2012) in which the LRR model is estimated using simulated method of moments (SMM). Hasseltoft (2012)'s results suggest that, with a negative correlation between inflation and consumption growth, the LRR model is able to produce both equity premiums and term premiums that are in line with the data.

Eraker (2008) further extends the LRR models by incorporating jumps into the process of stochastic variance – a measure of economic uncertainty in LRR models.

Under this model setup, the discretized state variables follow affine jump-diffusion processes. Eraker shows that, with Epstein-Zin utility, asset prices are exponential affine functions of state variables. In calibration exercise, when mean inflation impacts the real growth negatively (the negative signaling effect of mean inflation on consumption growth suggested by PS), the model is able to produce an equity premium and an upward-sloping nominal term structure that are in line with the data when intertemporal elasticity of substitution $\psi = 5$ and risk aversion $\gamma = 8$.

Different from our model, Eraker (2008), Doh (2012), Hasseltoft (2012), and Bansal and Shaliastovich (2013) all assume the LRR component is known to investors. Moreover, while Eraker (2008) only calibrates his model, Doh (2012) and Hasseltoft (2012) estimate their models using both macro-economic data and price data. In contrast, we estimate our model using only macro-economic data. In that sense, our tests are more stringent and our empirical results, at least in part, follow recommendations of Lewellen, Nagel, and Shanken (2010).

3.2.2 Estimation of State Space Models

Previous empirical work on estimating LRR models maybe categorized into 3 groups by the estimation method used: method of moments (MM) such as GMM (e.g., Constantinides and Ghosh (2011) and Ferson, Nallareddy, and Xie (2013)) and SMM (e.g., Hasseltoft (2012) and Bansal, Gallant and Tauchen (2007)); maximum likelihood method (MLE) (e.g., Piazzesi and Schneider (2007), Bansal and Shaliastovich (2013)); and Bayesian approach (e.g., Aldrich and Gallant (2010), Schorfheide, Song, and Yaron (2013), and Doh (2012)). These approaches have their own pros and cons. In general, MM requires fewer assumptions on the distributions of state variables than the Bayesian approach and MLE, for MM only matches certain moments of state variables to data without taking a stand on the joint distribution of state variables. However, this advantage of MM is not used in previous studies on LRR models. In the literature, the LRR model assumptions imply Gaussian states. Likewise, the assumption that latent variables are affine functions of observables also requires Gaussian structure (see Appendix A.2.3 of Constantinides and Ghosh (2011) and equation (5a) to (5f) in Ferson, Nallareddy, and Xie (2013)). Once we take a stand on the model distribution, the Bayesian approach and MLE are more effective in the sense that the whole distribution rather than certain moments are used in estimation. More importantly, the Bayesian approach returns finite-sample marginal distributions of model parameters and latent variables while MM produces only asymptotic distributions. This feature of Bayesian approach frees us from making in-sample corrections to parameter estimates, which are usually hard to compute. Finally, since all estimated densities of model parameters and state variables retrieved under the Bayesian approach are marginal, one can easily construct the exact marginal density of any function of model parameters and state variables. This enables us to see how well the model can fit the characteristics of functions of interest. The drawback of MLE is that it provides no convenient way to evaluate the model. MLE only produces conditional densities of model parameters and latent state variables conditioning on all other parameters and state variables. It is often infeasible to find the marginal densities of functions of model parameters and latent state variables. In contrast, such tasks can be easily accomplished when the model is estimated using the Bayesian approach as noted above.

The Bayesian approach is widely used in the estimation of generalized (non-Gaussian) state space models. Sampling approaches are developed by early studies such as Metropolis et al. (1953), Hastings (1970) (for Metropolis-Hastings algorithm), Geman and Geman (1984) (for Gibbs Sampling), and Gelfand and Smith (1990) (for combination of Gibbs and Metropolis-Hastings). Equipped with these sampling approaches, Carlin, Polson, and Stoffer (1992), and Geweke and Tanizaki (2001), among others, consider estimation of nonlinear and non-Gaussian state-space models. Carlin, Polson, and Stoffer (1992) handle the nonlinearity and non-normality by introducing nuisance parameters into the conditional variances in the state and/or observational processes. Doing so allows model estimation using Gibbs Sampling. Geweke and Tanizaki (2001) incorporate Metropolis-Hastings algorithm into model estimation so that the approach can handle a general form of state-space models. For linear Gaussian models, Carter and Kohn (1994) and Fruhwirth-Schnatter (1995) develop a so-called forward filtering backward sampling (FFBS) sampler in which latent state variables and parameters are sampled jointly (West and Harrison (1997) chapter 15).

General multivariate state-space models are not identified. In a recent study, Bai and Wang (2015) examine the minimum identifying restrictions for linear Gaussian dynamic factor models:

$$X_t = \Lambda_0 f_t + \Lambda_1 f_{t-1} + \cdots + \Lambda_s f_{t-s} + e_t,$$
$$f_t = \Phi_1 f_{t-1} + \cdots + \Phi_h f_{t-h} + \varepsilon_t,$$

where f_t is $q \times 1$, $e_t \sim i.i.d.N(0, R)$, and $\varepsilon_t \sim i.i.d.N(0, Q)$.

Bai and Wang (2015) provide two types of identifying restrictions:

1. DFM1: $Q = I_q$ and the first $q \times q$ block of Λ_0 is a lower-triangular matrix with positive diagonal elements, and

2. DFM2: the first $q \times q$ block of Λ_0 is identity matrix.

To incorporate the identification restrictions above as well as additional model-implied restrictions into the model estimation, Bai and Wang (2015) develop a Bayesian algorithm where Forward Filtering Backward Sampling (FFBS) is used to estimate states jointly. To ensure the numerical stability of the proposed algorithm, a square-root Kalman filter (see Bierman (1977) and Evenson (2009)) is used to compute the conditional distribution of states in FFBS.

3.3 Long-Run Risk Model with Learning

We model an endowment economy with investors with recursive preferences as in Duffie and Epstein (1992) (recent discrete-time examples of such models include Bansal and Yaron (2004) and Eraker (2008)). We retain main features of the LRR models such as non-i.i.d. consumption and dividend growth. The processes for consumption and dividend growth as well as the properties of inflation process are exogenous to the model. We list model assumptions below.

Assumption 1: *Following Eraker (2008) we model the processes of consumption growth and dividend growth with drift being a function of the long-run risk component x_t and mean inflation m_t*

$$dln(C) = dg_c = \left(\mu_c + Q_c m_t + x_t - \frac{V}{2}\right) dt + \sqrt{V} dw_c, \tag{1}$$

$$dg_d = \left(\mu_d + Q_d m_t + \phi x_t - \varphi_d^2 \frac{V}{2}\right) dt + \varphi_d \sqrt{V} dw_d, \tag{2}$$

where w_c and w_d are standard Wiener processes with $Edw_c dw_d = \rho_{cd} dt$.

Assumption 2: *Both processes of inflation and mean inflation are observed by investors.*

$$\frac{di}{i} = m_t dt + \sigma_i dw_i,$$

$$dln(i) = \left(m_t - \frac{\sigma_i^2}{2}\right) dt + \sigma_i dw_i, \tag{3}$$

$$dm = k(m^* - m_t)dt + \sigma_m dw_m, \tag{4}$$

where w_i *and* w_m *are standard Wiener processes with* $Edw_i dw_c = \rho_{ic} dt$, $Edw_i dw_d = \rho_{id} dt$, $Edw_i dw_m = \rho_{im} dt$, $Edw_m dw_c = \rho_{cm} dt$, $Edw_m dw_d = \rho_{md} dt$.

Assumption 3: *The long-run risk component* x_t *is not observed by investors. However, it is known that it follows a mean reverting process*

$$dx_t = -\rho x_t dt + \varphi_e \sqrt{V} dw_x, \tag{5}$$

where w_x *is a standard Wiener process with* $Edw_x dw_c = \rho_{xc} dt$, $Edw_x dw_d = \rho_{xd} dt$, $Edw_i dw_x = \rho_{ix} dt$, *, and* $Edw_m dw_x = \rho_{mx} dt$.

In practice, mean inflation m_t is not observable to investors. However, we assume that it is observable for two reasons. First, filtering x_t and m_t simultaneously leads to a recursive solution to the optimal filtering problem while we want to retain a closed-form solution. Second, assumption 2 allows us to isolate the impact of incomplete information about the variable of interest, x_t, on risk premiums and the term structure of interest rates. In the empirical work of this paper, we estimate both x_t and m_t simultaneously using FFBS.

Assumption 4: *All state variable processes are homoscedastic.*

We make the last assumption to retain the closed form solution for our bond pricing exercise.

According to assumption 3, the LRR component x_t is unobservable to the representative agent. We model the agent's optimal estimate of x_t as an expectation conditional on previous observations on consumption and dividend growth, inflation, and mean inflation. The agent estimates x_t and takes the optimal estimate of x_t as given when determining the equilibrium asset prices due to the Markov property of the model. We use the Karman-Bucy filter to derive the process of an agent's best estimate of x_t, $\hat{x}_t$. With investors' optimal estimate of x_t, $\hat{x}_t$, the model is reduced to a full-information model under the new measure corresponding to the information set available to investors. Based on results of Liptser and Shiryaev (2013), solution to our optimal filtering problem is summarized in the following result.

Result. *(Liptser and Shiryaev (2013)) The investors' best estimate of* x_t, $\hat{x}_t$, *is given by*

$$d\hat{x}_t = -\rho\hat{x}_t dt + \Sigma_x \begin{bmatrix} dw_i^* \\ dw_c^* \\ dw_d^* \\ dw_m^* \end{bmatrix}, \tag{6}$$

where the posterior variance of x_t, S_t, *and volatility of* $\hat{x}_t$, Σ_x, *are given by*

$$S_t = E_t[(x_t - \hat{x}_t)^2], \tag{7}$$

$$\Sigma_x = \begin{bmatrix} \Sigma_{1x} \\ \Sigma_{2x} \\ \Sigma_{3x} \\ \Sigma_{4x} \end{bmatrix}^T = [S\Phi_s + \varphi_s]\psi_1^0,$$

where

$$\psi = \begin{bmatrix} 0 & 0 & 0 & 0 & \varphi_e\sqrt{V} \end{bmatrix},$$

$$\psi_1 = \begin{bmatrix} \sigma_i & 0 & 0 & 0 & 0 \\ 0 & \sqrt{V} & 0 & 0 & 0 \\ 0 & 0 & \varphi_d\sqrt{V} & 0 & 0 \\ 0 & 0 & 0 & \sigma_m & 0 \end{bmatrix},$$

$$\psi_1^0 = \begin{bmatrix} \sigma_i & 0 & 0 & 0 \\ 0 & \sqrt{V} & 0 & 0 \\ 0 & 0 & \varphi_d\sqrt{V} & 0 \\ 0 & 0 & 0 & \sigma_m \end{bmatrix},$$

$$\Phi_s = [0 \quad 1 \quad \phi \quad 0](\psi_1 \Lambda \psi_1')^{-1},$$

$$\varphi_s = (\psi \Lambda \psi_1')(\psi_1 \Lambda \psi_1')^{-1},$$

$$\Lambda = \frac{1}{dt} E(dwdw'),$$

$$dw = [dw_i \quad dw_c \quad dw_d \quad dw_m \quad dw_x]',$$

and processes w_i^* w_c^* w_m^* *and* w_d^* *are standard Wiener processes under a filtration, which are given by*

$$\begin{bmatrix} dw_i^* \\ dw_c^* \\ dw_d^* \\ dw_m^* \end{bmatrix} = \begin{bmatrix} dw_i \\ \dfrac{x_t - \hat{x}_t}{\sqrt{V}} dt + dw_c \\ \dfrac{\phi(x_t - \hat{x}_t)}{\varphi_d \sqrt{V}} dt + dw_d \\ dw_m \end{bmatrix}.$$

The posterior variance, S_t can be retrieved from the following ordinary differential equation (ODE) which follows from applying Itô's lemma to the definition of the variance in (7) and taking expectations:

$$dS_t = (-\alpha S_t^2 - 2(\rho + \xi) S_t + \Gamma^2) dt,$$

where

$$\alpha = \Phi_s \begin{bmatrix} 0 \\ 1 \\ \phi \\ 0 \end{bmatrix} > 0,$$

$$\xi = tr(\Phi_s \psi_1 \Lambda \psi'),$$

$$\Gamma^2 = \psi \Lambda \psi' - tr\big((\psi \Lambda \psi_1') {\varphi_s}'\big).$$

In the steady state, we have $\frac{dS}{dt} = 0$. Therefore, $S = \frac{-(\rho+\xi)+\sqrt{(\rho+\xi)^2+\alpha\Gamma^2}}{\alpha}$.

With the process of $\hat{x}_t$ given by (6), we can now write down processes of state variables $ln(i)$, g_c, g_d and m under the new measure:

$$dln(i) = \left(m_t - \frac{\sigma_i^2}{2}\right)dt + \sigma_i dw_i^*, \tag{8}$$

$$dg_c = \left(\mu_c + Q_c m_t + \hat{x}_t - \frac{V}{2}\right)dt + \sqrt{V}dw_c^*, \tag{9}$$

$$dg_d = \left(\mu_d + Q_d m_t + \phi\hat{x}_t - \varphi_d^2\frac{V}{2}\right)dt + \varphi_d\sqrt{V}dw_d^*, \tag{10}$$

$$dm = k(m^* - m_t)dt + \sigma_m dw_m^*. \tag{11}$$

3.4 Stochastic Discount Factor

In order to derive the stochastic discount factor (SDF), we assume investors have the stochastic differential utility preferences of Duffie and Epstein (1992).

Assumption 5: *Investors have the stochastic differential utility preferences of Duffie and Epstein (1992):*

$$J(ln(i), g_d, m, \hat{x}, W_t, t) = E_t\int_t^T f(C_s, J_s)ds, \tag{12}$$

with the normalized aggregator

$$f(C_t, J_t) = \frac{\beta}{1-\frac{1}{\psi}}(1-\gamma)J_t\left[\left(\frac{C_t}{\left((1-\gamma)J_t\right)^{\frac{1}{1-\gamma}}}\right)^{1-\frac{1}{\psi}} - 1\right], \tag{13}$$

where T is investors' horizon, C_t is consumption at time t, J_t is the recursive utility at time t, β is the discount rate, γ is the relative risk aversion, and ψ is the elasticity of intertemporal substitution.

For a given consumption process, Duffie and Epstein (1992) prove the Bellman optimality condition and show that it implies that the optimal differential utility satisfies the following PDE:

$$J_t + \hat{\mu}_G^T J_G + \frac{1}{2dt} dG^T J_{GG} dG + f = 0, \tag{14}$$

$$J(T, G) = 0, \tag{15}$$

where $G = (\ln(i), g_c, g_d, m, \hat{x})$' and $\hat{\mu}_G$ is the mean vector given by

$$\hat{\mu}_G = \begin{bmatrix} m_t - \frac{\sigma_i^2}{2} \\ \mu_c + Q_c m_t + \hat{x}_t - \frac{V}{2} \\ \mu_d + Q_d m_t + \phi\hat{x}_t - \varphi_d^2 \frac{V}{2} \\ k(m^* - m_t) \\ -\rho\hat{x}_t \end{bmatrix}.$$

In general, the PDE (14) - (15) does not have a closed-form solution. To get a closed-form solution of the PDE (13) - (14), we use the log-linear approximation suggested by Campbell and Viceira (2003) and Zhu (2006):

$$f \approx \frac{(1-\gamma)}{1 - \frac{1}{\psi}} J[h_0 + hln\beta + hlnF - \beta] \tag{16}$$

$$= h(1-\gamma)J\left[lnC - \frac{1}{(1-\gamma)} lnJ + H\right]$$

where

$$H = \frac{(h_0 + hln\beta - \beta)}{h\left(1 - \frac{1}{\psi}\right)} - \frac{1}{(1-\gamma)} ln(1-\gamma).$$

The detailed derivation is given in chapter two of this thesis. The linear approximation (16) allows us to get closed-form solutions to the model. However, this is not without a cost. With this approximation, the *EIS*, ψ, has no impact on pricing due to the fact that the *EIS* only appears in *H* and has not effect on the aggregator (16) or the value function to the first order of the log-linearization.

Inserting approximation (16) into PDE (14), we look for an exponential-affine solution to the value function $J(t, \ln(i), g_c, g_d, \hat{x}, m)$:

$$J(t, \ln(i), g, d, \hat{x}, m) = \exp\{\xi_{0t} + \xi_{1t}\ln(i) + \xi_{2t}g_c + \xi_{3t}g_d + \xi_{4t}m + \xi_{5t}\hat{x}\}. \quad (17)$$

Since the PDE must hold for any arbitrary values of the state variables, coefficients of the state variables must be all zero, which leads to a system of five ODEs for five ξ's in equation (17). Assuming that investors are infinitely lived, $T \to \infty$, we get the following solutions (see Appendix for details):

$$\xi_1 = \xi_3 = 0, \quad (18)$$

$$\xi_2 = 1 - \gamma, \quad (19)$$

$$\xi_4 = \frac{(1-\gamma)Q_c}{k+h}, \quad (20)$$

$$\xi_5 = \frac{1-\gamma}{h+\rho}. \quad (21)$$

The next step is to find the Stochastic Discount Factor (SDF) in the model. Duffie and Epstein (1992) show that the SDF has the following form:

$$\Pi_t = \exp\left\{\int_0^t f_J dx\right\} f_C. \quad (22)$$

Applying Itô's lemma to (22) and using the normalized aggregator approximation (16) and the state processes (6) and (8) – (11), we get the process of the pricing kernel, π_t (see Appendix for details):

$$\frac{d\Pi}{\Pi} = -[\Omega + \hat{x}_t + Q_c m]dt + \sigma_\Pi^T \begin{bmatrix} dw_i^* \\ dw_c^* \\ dw_d^* \\ dw_m^* \end{bmatrix}, \quad (23)$$

where

$$\Omega = h + \mu_c + (\xi_2 - 1)V + \xi_5 \sqrt{V} \Sigma_x [\rho_{ic} \quad 1 \quad \rho_{cd} \quad \rho_{cm}]' + \xi_4 \sigma_m \sqrt{V} \rho_{mc}, \tag{24}$$

$$\sigma_\Pi = \begin{bmatrix} \xi_5 \Sigma_{1x} \\ (\xi_2 - 1)\sqrt{V} + \xi_5 \Sigma_{2x} \\ \xi_5 \Sigma_{3x} \\ \xi_5 \Sigma_{4x} + \xi_4 \sigma_m \end{bmatrix}.$$

and the process for the nominal pricing kernel is given by

$$\frac{d\Pi^n}{\Pi^n} = \frac{d(\Pi/i)}{\Pi/i} \tag{25}$$

$$= \left[-(\Omega + \hat{x}_t + Q_c m) - m + \sigma_i^2 - \sigma_i \begin{bmatrix} 1 \\ \rho_{ic} \\ \rho_{id} \\ \rho_{im} \end{bmatrix}^T \sigma_\Pi \right] dt$$

$$+ \begin{bmatrix} \xi_5 \Sigma_{1x} - \sigma_i \\ (\xi_2 - 1)\sqrt{V} + \xi_5 \Sigma_{2x} \\ \xi_5 \Sigma_{3x} \\ \xi_5 \Sigma_{4x} + \xi_4 \sigma_m \end{bmatrix}^T \begin{bmatrix} dw_i^* \\ dw_c^* \\ dw_d^* \\ dw_m^* \end{bmatrix}$$

$$= -[\Omega^n + \hat{x}_t + (Q_c + 1)m]dt + \sigma_{\Pi^n}^T \begin{bmatrix} dw_i^* \\ dw_c^* \\ dw_d^* \\ dw_m^* \end{bmatrix}.$$

According to the value function given by equation (17) – (21), investor's utility level is determined by values of four state variables: inflation, consumption growth, mean inflation, and filtered mean consumption growth. This implies four types of priced risks corresponding to shocks in these four state variables. However, the diffusion component of the SDF in equation (25) still includes shocks in the dividend growth dw_d^*. To see why this is the case, we may rewrite $\sigma_{\Pi^n}^T$ as $\sigma_{\Pi^n}^T = \xi_5 \Sigma_x + \begin{bmatrix} -\sigma_i \\ (\xi_2 - 1)\sqrt{V} \\ 0 \\ \xi_4 \sigma_m \end{bmatrix}^T$. The first term on the right-hand side of the equation corresponding to the filtered consumption growth mean, and non-zero elements in the matrix in the second term correspond to inflation, consumption growth, and mean inflation, respectively. The third element of the matrix is zero, which suggests that dividend growth shocks do not enter the diffusion of the SDF directly. Instead, it enters via the shocks in filtered mean consumption growth. In our model, the investor's optimal estimation of mean consumption growth is a linear combination of observed variables including the dividend growth. Therefore, dw_d^* appears in the diffusion of $\hat{x}_t$ (equation (6)) and further the diffusion of the value function and that of the pricing kernel. $\sigma_{\Pi^n}^T$ also provides hints on the sign of the risk premium. Since the SDF is simply the marginal rate of substitution of utility, a large SDF implies a low utility level in the future. Therefore, any return (payoff) that is negatively correlated with the SDF should be positively correlated with utility and considered risky. For a risk averse investor ($\gamma > 1$), $\xi_2 - 1$ and ξ_5 are in general negative. Previous studies (Piazzesi and Schneider (2007), Eraker (2008), and Doh (2012)) suggest that mean consumption growth's loading on mean inflation Q_c is negative, which leads to a positive ξ_4. Therefore, if one asset's payoff is more positively correlated with shocks in inflation, consumption growth, and filtered mean consumption growth, and more negatively correlated with shocks in mean inflation, it should have a higher positive risk premium in the sense that its payoff is more negatively correlated with the SDF. One the other hand, when asset payoffs are negatively correlated with shocks in inflation, consumption growth and its filtered mean, and positively correlated with shocks in mean inflation, the asset provides hedges to priced risks and should have a negative risk premium.

With the nominal SDF process (25), the corresponding nominal risk-free rate r_f^n is given by the drift of the nominal SDF process

$$r_f^n = \Omega^n + \hat{x}_t + (Q_c + 1)m, \tag{26}$$

where $\Omega^n = \Omega - \sigma_i^2 + \sigma_i \begin{bmatrix} 1 \\ \rho_{ic} \\ \rho_{id} \\ \rho_{im} \end{bmatrix}^T \sigma_\pi$. We may rewrite the nominal risk free rate as the sum of five terms:

$$r_f^n = h + (\mu_c + \hat{x}_t + Q_c m) + (\Omega - h - \mu_c) + m + \sigma_i \begin{bmatrix} 1 \\ \rho_{ic} \\ \rho_{id} \\ \rho_{im} \end{bmatrix}^T \sigma_{\Pi^n}, \tag{27}$$

The first and second term imply that the risk free rate increases with the average consumption-wealth ratio and the mean consumption growth. The third term is the risk premium corresponding to the uncertainty in consumption growth (the precautionary saving effect). Intuitively, since the consumption growth is stochastic, the future consumption is uncertain, which leads to a lower expected utility level. Therefore, an investor saves more today. The demand of risk free assets pushes up the asset price and makes the return lower. With a power utility, this term is only a function of relative risk aversion and consumption growth variances and is always negative (see Cochrane (2005)). In our case, since the utility level is jointly determined by consumption and the value function, the precautionary saving term also contains consumption growth's covariance with the value function (see equation (24)). The value of this term is jointly determined by the correlation between state variables $(\rho_{mc}, \rho_{ic}, \rho_{cd})$ and other model parameters. Its magnitude is also determined by the relative risk aversion (γ). A larger γ makes absolute value of all three components larger. The last two terms in nominal short rate capture the impact of inflation. They show that the impact of inflation depends not only on mean inflation but also on inflation's conditional covariance with the SDF. Intuitively, the nominal rate can be considered as return on a mixture of a risk free asset and an asset exposed to inflation risk. To hold this portfolio, an investor requires the risk compensation on exposure to the extra inflation risk on top of the real risk free rate.

3.5 Term Structure

The nominal price of a zero-coupon bond that matures at time s and pays a dollar of face value at its maturity $P(t,s)$ must satisfy the following PDE with risk corrections implied by the nominal pricing kernel (25):

$$E\left(\frac{dP}{P}\right) = r_f^n dt - E\left(\frac{dP}{P}\frac{d\Pi^n}{\Pi^n}\right). \tag{28}$$

Given the affine nature of the short rate, equations (25), (26), and (28) imply that the solution for the bond price has an exponential-affine form as a function of posterior mean of the LRR component $\hat{x}_t$ and the inflation mean m_t (see Appendix for details). Therefore, we look for a time-t price solution in the form of an exponential-affine function of $\hat{x}_t$ and m_t:

$$P = exp\{A(t,s) + B(t,s)\hat{x} + C(t,s)m\}. \tag{29}$$

In our case, the PDE for the bond price (28) takes the following form (see Appendix for details):

$$\frac{P_t}{P} - \rho\hat{x}_t\frac{P_{\hat{x}}}{P} + k(m^* - m_t)\frac{P_m}{P} + \frac{\Sigma_x\Lambda_{1:4,1:4}\Sigma_x{}'}{2}\frac{P_{\hat{x}\hat{x}}}{P} + \frac{\sigma_m^2}{2}\frac{P_{mm}}{P} + \sigma_m\Sigma_x\Sigma_{ym}\frac{P_{\hat{x}m}}{P}$$
$$+ \frac{P_{\hat{x}}}{P}\sigma_{\Pi^n}^T\Lambda_{1:4,1:4}\Sigma_x' + \frac{P_m}{P}\sigma_m\sigma_{\Pi^n}'\Sigma_{ym} - r_{f,t}^n = 0\,. \tag{30}$$

where $\Sigma_{ym} = \begin{bmatrix} \rho_{im} \\ \rho_{cm} \\ \rho_{dm} \\ 1 \end{bmatrix}$, P_t, $P_{\hat{x}}$, and P_m are first order partial derivatives of bond price P,

and P_{mm}, $P_{\hat{x}m}$, and $P_{\hat{x}\hat{x}}$ are second order partial derivatives.

Theorem 1: *The nominal price of a zero-coupon bond paying 1 dollar at its maturity, s, is given by an exponential-affine function of the state variables:*

$$P = exp\{A(t,s) + B(t,s)\hat{x} + C(t,s)m\}, \tag{31}$$

where functions $A(t,s)$, $B(t,s)$, and $C(t,s)$ are given by

$$B(t,s) = -\frac{1}{\rho}(1 - exp\{-\rho(s-t)\}), \tag{32}$$

$$C(t,s) = -\frac{(Q_c+1)}{k}(1 - exp\{-k(s-t)\}), \tag{33}$$

$$\begin{aligned} A = c_0(s-t) &- \frac{c_1}{\rho}(exp\{\rho(t-s)\} - 1) - \frac{c_2}{k}(exp\{k(t-s)\} - 1) \\ &- \frac{c_4}{2k}(exp\{2k(t-s)\} - 1) - \frac{c_3}{2\rho}(exp\{2\rho(t-s)\} - 1) \\ &- \frac{c_5}{k+\rho}(exp\{(k+\rho)(t-s)\} - 1). \end{aligned} \tag{34}$$

Constants c_0, c_1, …, c_5 are defined in Appendix A3. The nominal term premium is defined by the negative covariance of the nominal SDF and the return on the bond, and has the following form:

$$\lambda(t,s) = -E\left(\frac{dP}{P}\frac{d\Pi^n}{\Pi^n}\right) = -B(t,s)\sigma_{\Pi^n}^T\Lambda_{1:4,1:4}\Sigma_x' - C(t,s)\sigma_m\sigma_{\Pi^n}'\Sigma_{ym}. \tag{35}$$

Proof: See Appendix A3.

According to equation (35), the term premium has two components: compensation for exposure to LRR risk ($\lambda_{\hat{x}}$) and compensation for exposure to mean inflation risk (λ_m) .

$$\lambda_{\hat{x}} = -B(t,s)\sigma'_{\Pi^n}\Lambda_{1:4,1:4}\Sigma'_x,$$

$$\lambda_m = -C(t,s)\sigma_m\sigma'_{\Pi^n}\Sigma_{ym}.$$

With positive mean-reversion in latent variables ($\rho > 0$ and $k > 0$), and mean consumption growth's loading on mean inflation $Q_c > -1$, $B(t,s)$ and $C(t,s)$ are negative. Therefore, signs of $\lambda_{\hat{x}}$ and λ_m are determined by conditional covariance of the nominal pricing kernel with the perceived mean consumption growth ($\hat{x}$) and with the mean inflation (m), respectively. For example, if both $\hat{x}$ and m are negatively conditionally correlated with the nominal SDF, the total risk premium should be negative. Intuitively, in this case, both $\hat{x}$ and m are risky in the sense that they are both positively correlated with utility. However, since both $B(t,s)$ and $C(t,s)$ are negative, the bond return has negative loadings on $\hat{x}$ and m. Therefore, the bond return is negatively correlated with utility and serves as insurance.

Given the price function, we can easily compute the yield, $y_{t,s}$, with time to maturity s-t:

$$y_{t,s} = \frac{1}{t-s}(A + B\hat{x}_t + Cm_t), \tag{36}$$

and the conditional variance of the yield

$$var(y_{t,s}|\hat{x}_{t-1}, m_{t-1}) = \frac{B^2 z_1 + 2BCz_2 + C^2\sigma_m^2}{(t-s)^2}. \tag{37}$$

where constants z_1 and z_2 are given in Appendix A3.3.

3.6 Estimation Method

We estimate the Euler-discretized version of the model (1) - (5). The observed variable vector y_t contains log real consumption growth, Δg_c, log real dividend growth Δg_d, and log inflation $\Delta\pi = \ln(i)$.

$$y_t = \begin{bmatrix} \Delta\pi_t \\ \Delta g_c \\ \Delta g_d \end{bmatrix} = \mu_y + F\theta_{t-1} + v_t, \tag{38}$$

where

$$\mu_y = \begin{bmatrix} 0 \\ \mu_c \\ \mu_d \end{bmatrix},$$

$$F = \begin{bmatrix} 1 & 0 \\ Q_c & 1 \\ Q_d & \phi \end{bmatrix},$$

$$\theta_{t-1} = \begin{bmatrix} m_{t-1} \\ x_{t-1} \end{bmatrix},$$

$$v_t \sim N\left(0,\, \Sigma_y\right),$$

$$\Sigma_y = \begin{bmatrix} \sigma_i^2 & \sigma_i\sqrt{V}\rho_{ic} & \sigma_i\varphi_d\sqrt{V}\rho_{id} \\ \sigma_i\sqrt{V}\rho_{ic} & V & \varphi_d V\rho_{cd} \\ \sigma_i\varphi_d\sqrt{V}\rho_{id} & \varphi_d V\rho_{cd} & \varphi_d^2 V \end{bmatrix}.$$

While the latent variable vector θ_t is given by

$$\theta_t = \mu_\theta + G\theta_{t-1} + w_t, \tag{39}$$

where

$$\mu_\theta = \begin{bmatrix} km^* \\ 0 \end{bmatrix},$$

$$G = \begin{bmatrix} 1-k & 0 \\ 0 & 1-\rho \end{bmatrix}$$

$$= \begin{bmatrix} k^* & 0 \\ 0 & \rho^* \end{bmatrix},$$

$$w_t \sim N(0, \Sigma_\theta),$$

$$Ev_t w_t' = \Sigma_{y\theta},$$

$$\Sigma_\theta = \begin{bmatrix} \sigma_m^2 & \sigma_m\varphi_e\sqrt{V}\rho_{mx} \\ \sigma_m\varphi_e\sqrt{V}\rho_{mx} & \varphi_e^2 V \end{bmatrix},$$

$$\Sigma_{y\theta} = \begin{bmatrix} \sigma_i\sigma_m\rho_{im} & \sigma_i\varphi_e\sqrt{V}\rho_{ix} \\ \sigma_m\sqrt{V}\rho_{cm} & \varphi_e V\rho_{cx} \\ \sigma_m\varphi_d\sqrt{V}\rho_{dm} & \varphi_d\varphi_e V\rho_{dx} \end{bmatrix}.$$

We can rewrite the model (38) – (39) as

$$y_t = \mu_y^* + \left(F - \Sigma_{y\theta}\Sigma_\theta^{-1}G\right)\theta_{t-1}^* + \Sigma_{y\theta}\Sigma_\theta^{-1}\theta_t^* + v_t^*, \tag{40}$$

$$\theta_t^* = G\theta_{t-1}^* + w_t, \tag{41}$$

where $v_t^* \sim N(0, R)$, $R = \Sigma_y - \Sigma_{y\theta}\Sigma_\theta^{-1}\Sigma_{y\theta}'$, and $\mu_y^* = \mu_y - F(G - I_2)^{-1}\mu_\theta$. With this model transformation, we get rid of the correlation between observed and latent variables in the original model (i.e. v_t^* is orthogonal to w_t).

If we stack θ_t^* with θ_{t-1}^*, system (40) – (41) becomes

$$y_t = \Lambda \begin{bmatrix} 1 \\ \Theta_t \end{bmatrix} + v_t^* \tag{42}$$

$$\Theta_t = \Phi\Theta_{t-1} + w_t^* \tag{43}$$

where

$$\Lambda = \left[\mu_y^* \,\vdots\, \Lambda_0 \,\vdots\, \Lambda_1\right],$$

$$\Lambda_0 = \Sigma_{y\theta}\Sigma_\theta^{-1} = \begin{bmatrix} \lambda_1 & \lambda_2 \\ \lambda_3 & \lambda_4 \\ \lambda_5 & \lambda_6 \end{bmatrix},$$

$$\Lambda_1 = F - \Sigma_{y\theta}\Sigma_\theta^{-1}G,$$

$$\Theta_t = \begin{bmatrix} \theta_t^* \\ \theta_{t-1}^* \end{bmatrix},$$

$$\Phi = \begin{bmatrix} G & 0_{2\times2} \\ I_2 & 0_{2\times2} \end{bmatrix},$$

$$w_t^* = \begin{bmatrix} w_t \\ 0_{2\times1} \end{bmatrix}.$$

Since the model (42) – (43) is Gaussian, it allows us to use the Gibbs Sampler and FFBS to estimate it. The Gibbs sampler (or Gibbs sampling algorithm) is originally developed by Geman and Geman (1984). It is a special case of the Metropolis-Hastings algorithm. In Gibbs sampler, values of interest are sequentially simulated directly from their full conditional distribution. Let $X = \{x_i, i = 1,2, \dots, n\}$ be the set of values of interest and $X_{-i} = \{x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n\}$. In each iteration t, we draw x_i^t from its full conditional distribution $f_i(x_i | x_1^t, x_2^t, \dots, x_{i-1}^t, x_{i+1}^{t-1}, \dots, x_n^{t-1})$. Since draws are always accepted in the Gibbs sample, the use of the Gibbs sampler limits the choice of full conditional distributions and requires a prior knowledge on f_i (at least we should know how to draw from f_i). FFBS is developed by Carter and Kohn (1994) and Frühwirth-Schnatter (1994) for linear state-space models with Gaussian error (or a mixture of normals) to draw latent factors simultaneously. It is essentially a multi-step Gibbs sampling algorithm, in which Kalman filter results are utilized. Consider a linear Gaussian state space model with observable $Y^T = \{y_t, t = 1,2, \dots, T\}$, latent variable $X^T = \{x_t, t = 1,2, \dots, T\}$ and the model parameter set, θ. To draw X, we first retrieve $p(x_T | Y^T, \theta) = N\left(x_{T|T}, P_{T|T}\right)$ where $x_{T|T}$ and $P_{T|T}$ are the conditional mean and conditional variance of x_T given Y^T and θ, respectively, and both values are from Kalman filter. In the backward-sampling step, we first draw x_T from $p(x_T | Y^T, \theta)$. Then one proceeds by drawing x_t from $p(x_t | x_{t+1}^*, Y^t, \theta)$ for $t = n - 1, \dots, 1$, where x_{t+1}^* is a draw of x_{t+1} (see West and Harrison (1997) and Robert and Casella (2013)).

To make sure our model is identified, we adopt the DFM2 restriction suggested by Bai and Wang (2015). DFM2 contains least restrictions on model parameters that are sufficient to identify the model. For a state space model $y_t = \Lambda_0 x_t + \Lambda_1 x_{t-1} + \cdots + \Lambda_s x_{t-s} + \varepsilon_t$, where y_t is $N \times 1$ state vector and x_t is $q \times 1$ latent vector, DFM2 requires that the first $q \times q$ block of Λ_0 be an identity matrix. Applying DFM2 to our model, we get Λ_0 and Λ_1 as follows:

$$\Lambda_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \lambda_5 & \lambda_6 \end{bmatrix}, \tag{44}$$

$$\begin{aligned} \Lambda_1 &= \begin{bmatrix} 1 - k^* & 0 \\ Q_c & 1 - \rho^* \\ Q_d - \lambda_5 k^* & \phi - \lambda_6 \rho^* \end{bmatrix} \\ &= \begin{bmatrix} 1 - k^* & 0 \\ Q_c & 1 - \rho^* \\ {Q_d}^* & \phi^* \end{bmatrix}. \end{aligned} \tag{45}$$

The model (42) – (43) contains two sets of unknown variables: the latent state factors, $\Theta_t, t = 1,2, \dots, T$, and model parameters $\left\{\mu^*_{y_{3\times 1}}, \lambda_5, \lambda_6, Q_c, {Q_d}^*, \phi^*, k^*, \rho^*, R, Q\right\}$. To estimate the latent factors Θ_t and model parameters simultaneously, we follow Bai and Wang (2015) and use a Bayesian estimation approach containing 3 recursive steps:

Step 1, we sample the latent factors $\Theta_t, t = 1,2, \dots, T,$ conditional on parameters using FFBS;

Step 2, we sample drift parameters $\left\{\mu^*_{y_{3\times 1}}, \lambda_5, \lambda_6, Q_c, {Q_d}^*, \phi^*, k^*, \rho^*\right\}$ conditional on latent factors, $R,$ and Q;

Step 3, we sample R and Q conditional on drift parameters $\left\{\mu^*_{y_{3\times 1}}, \lambda_5, \lambda_6, Q_c, {Q_d}^*, \phi^*, k^*, \rho^*\right\}$ and latent factors Θ_t.

Below we provide details of each of the steps.

Step 1. Sampling the latent factors Θ_t conditional on parameters

We use $Y_t = [y_t, y_{t-1}, \dots, y_1]'$ to denote the history of y up to period t. In this step, we want to sample Θ_t from the joint distribution

$$p(\Theta_T, \dots, \Theta_1|Y_T) = p(\Theta_T|Y_T)\prod_{t=0}^{T-1} p(\Theta_t|\Theta_{t+1}, \dots \Theta_t, Y_t)$$

$$= p(\Theta_T|Y_T)\prod_{t=0}^{T-1} p(\Theta_t|\Theta_{t+1}, Y_t).$$

To do so, we use the backward-sampling algorithm. We first start with sampling Θ_T from

$$p(\Theta_T|Y_T) = N\left(\Theta_{T|T}, P_{T|T}\right),$$

where $\Theta_{T|T}$ and $P_{T|T}$ are the conditional mean and conditional variance of Θ_T given Y_T , respectively. Both $\Theta_{T|T}$ and $P_{T|T}$ are retrieved from the Kalman filter.

The next step is to draw Θ_t from $p(\Theta_t|\Theta_{t+1}, Y_t)$, $t = T-1, \dots, 1$. Given Θ_{t+1}, only the last element of Θ_t , θ^*_{t-1}, is random, which can be drawn from

$$p(\theta^*_{t-1}|\Theta_{t+1}, Y_t) = p(\theta^*_{t-1}|\theta^*_t, \theta^*_{t+1}, Y_t)$$

$$= \frac{p(\theta^*_{t-1}, \theta^*_{t+1}|\theta^*_t, Y_t)}{p(\theta^*_{t+1}|\theta^*_t, Y_t)}$$

$$= \frac{p(\theta^*_{t-1}|\theta^*_t, Y_t)p(\theta^*_{t+1}|\theta^*_t, \theta^*_{t-1}, Y_t)}{p(\theta^*_{t+1}|\theta^*_t, Y_t)}$$

$$= \frac{p(\theta^*_{t-1}|\theta^*_t, Y_t)p(\theta^*_{t+1}|\theta^*_t, Y_t)}{p(\theta^*_{t+1}|\theta^*_t, Y_t)}$$

$$\propto p(\theta^*_{t-1}|\theta^*_t, Y_t).$$

The fourth equality above holds due to the fact that θ^*_{t-1} contains no information about θ^*_{t+1} which is not captured by θ^*_t, and Y_t.

$p(\theta^*_{t-1}|\theta^*_t, Y_t)$ can be derived from $p(\Theta_t|Y_t) = N\left(\Theta_{t|t}, P_{t|t}\right)$, a Kalman filter result.

Let

$$N\left(\Theta_{t|t}, P_{t|t}\right) = N\left(\begin{bmatrix}\mu_{\theta^*_t} \\ \mu_{\theta^*_{t-1}}\end{bmatrix}, \begin{bmatrix} P_{\theta^*_t} & P_{\theta^*_t\theta^*_{t-1}} \\ P_{\theta^*_{t-1}\theta^*_t} & P_{\theta^*_{t-1}}\end{bmatrix}\right),$$

then

$$\theta_{t-1}^{*}|\theta_t^{*}, Y_t \sim N(\mu, P),$$

where $\mu = \mu_{\theta_{t-1}^{*}} + P_{\theta_{t-1}^{*}\theta_t^{*}} {P_{\theta_t^{*}}}^{-1}\left(\theta_t^{*} - \mu_{\theta_t^{*}}\right)$, and $P = P_{\theta_{t-1}^{*}} - P_{\theta_{t-1}^{*}\theta_t^{*}} {P_{\theta_t^{*}}}^{-1} P_{\theta_t^{*}\theta_{t-1}^{*}}$.

Step 2. Sampling model drift parameters $\left\{\mu_{y_{3\times 1}}^{*}, \lambda_5, \lambda_6, Q_c, {Q_d}^{*}, \phi^{*}, k^{*}, \rho^{*}\right\}$ conditional on the factors Θ_t, R, and Q.

Let

$$H = [\theta_0^{*}, \dots, \theta_{T-1}^{*}]',$$

$$X = [\theta_1^{*}, \dots, \theta_T^{*}]',$$

$$Fac = \begin{bmatrix} 1 & \cdots & 1 \\ \Theta_0^{*} & \cdots & \Theta_{T-1}^{*} \end{bmatrix}',$$

$$Y = [y_1, \dots, y_T]',$$

$$Z = [Y \vdots X],$$

In this notation, the matrix representation of the model (42) – (43) is

$$Z = Fac \times \Lambda^{*} + U.$$

We may write the model in a vector form, which is given by the vectorization of the matrix form:

$$vec(Z) = (I_5 \otimes Fac)\lambda + vec(U), \tag{46}$$

where

$$\lambda = vec(\Lambda^{*}),$$

$$vec(U) \sim N(0, \Sigma_Z \otimes I_T),$$

$$\Lambda^{*} = \begin{bmatrix} \mu_y^{*} \vdots 0 \\ \Lambda_0' \vdots 0 \\ \Lambda_1' \vdots G \end{bmatrix},$$

$$\Sigma_Z = \begin{bmatrix} R \vdots 0 \\ 0 \vdots Q \end{bmatrix},$$

and $\otimes$ is the Kronecker product symbol.

Recall that the elements of Λ^* are subject to linear constraints DFM2 (44). To impose the constraints on Λ^* and, thus, on λ, following Bai and Wang (2015) we write λ as

$$\lambda = B\delta + C \tag{47}$$

where $\delta = \left[\mu_y^{*}{'}_{3\times 1}\ \lambda_5\ \lambda_6\ Q_c, {Q_d}^*, \phi^*, k^*, \rho^*\right]'$ is the vector of free parameters. Matrices B and C are determined by both identification restrictions specified by DFM2 (44) and model restrictions (45).

Using equation (47), we rewrite the matrix representation of the model (46) as

$$z = Fac^*\delta + vec(U). \tag{48}$$

where $z = vec(Z) - (I_5 \otimes Fac)C$, $Fac^* = (I_5 \otimes Fac)B$.

To compute the conditional posteriors, we assume a conjugate normal prior distribution for drift parameters:

$$\delta \sim N\left(\delta^0, \Sigma_\delta^0\right).$$

The likelihood function is given by:

$$p(z|Fac^*, \delta, \Sigma_Z) \propto |\Sigma_Z|^{-\frac{T}{2}} exp\left\{-\frac{1}{2}(z - Fac^*\delta)'(\Sigma_Z \otimes I_T)^{-1}(z - Fac^*\delta)\right\}$$

$$= |\Sigma_Z|^{-\frac{T}{2}} exp\left\{-\frac{1}{2}(Z - Fac\Lambda^*)'(\Sigma_Z)^{-1}(Z - Fac\Lambda^*)\right\}$$

$$= |\Sigma_Z|^{-\frac{T}{2}} exp\left\{-\frac{1}{2}tr\left[(Z - Fac\Lambda^*)'(Z - Fac\Lambda^*){\Sigma_Z}^{-1}\right]\right\}.$$

To simplify the likelihood function, note that

$$(Z - Fac\Lambda^*)'(Z - Fac\Lambda^*)$$

$$= \left(Z - Fac\widehat{\Lambda}^*\right)'\left(Z - Fac\widehat{\Lambda}^*\right) + \left(\Lambda^* - \widehat{\Lambda}^*\right)Fac'Fac\left(\Lambda^* - \widehat{\Lambda}^*\right)'$$

$$= S + \left(\Lambda^* - \widehat{\Lambda}^*\right)Fac'Fac\left(\Lambda^* - \widehat{\Lambda}^*\right)',$$

where

$$S = \left(Z - Fac\widehat{\Lambda}^*\right)'\left(Z - Fac\widehat{\Lambda}^*\right),$$

and $\widehat{\Lambda}^* = \left[\widehat{\Lambda}_\pi \vdots \widehat{\Lambda}_c \vdots \widehat{\Lambda}_d \vdots \begin{matrix} 0_{3\times 2} \\ \widehat{G} \end{matrix}\right]$ is a 5 by 5 matrix. The first, second, third column vector ($\widehat{\Lambda}_\pi$, $\widehat{\Lambda}_c$, and $\widehat{\Lambda}_d$), and the 2 by 2 block in the lower right corner ($\widehat{G}$) of matrix $\widehat{\Lambda}^*$ are given by

$$\widehat{G} = (H'H)^{-1}H'X,$$

$$\widehat{\Lambda}_\pi = \left(F_{(1,4)}{}'F_{(1,4)}\right)^{-1}F_{(1,4)}{}'\left(Y - F_{(2)}\right),$$

$$\widehat{\Lambda}_c = \left(F_{(1,4,5)}{}'F_{(1,4,5)}\right)^{-1}F_{(1,4,5)}{}'\left(Y - F_{(3)}\right),$$

$$\widehat{\Lambda}_c = \left(F_{(1,2,3,4)}{}'F_{(1,2,3,4)}\right)^{-1}F_{(1,2,3,4)}{}'Y,$$

and $F_{(i,j,n,m)}$ is the matrix formed by the *ith*, *jth*, *nth*, and *kth* column of F.

This implies that the likelihood function can be written as

$$p(z|Fac^*, \delta, \Sigma_Z)$$

$$\propto |\Sigma_Z|^{-\frac{T}{2}} exp\left\{-\frac{1}{2}tr\left[S\Sigma_Z{}^{-1}\right] - \frac{1}{2}tr\left[\left(\Lambda^* - \widehat{\Lambda}^*\right)Fac'Fac\left(\Lambda^* - \widehat{\Lambda}^*\right)'\Sigma_Z{}^{-1}\right]\right\}$$

The conditional posterior for δ is given by:

$$p(\delta|Z, Fac, \Sigma_Z) \propto p(z|Fac^*, \delta, \Sigma_Z)p(\delta)$$

$$\propto exp\left\{-\frac{1}{2}tr\left[\left(\Lambda^* - \widehat{\Lambda}^*\right)Fac'Fac\left(\Lambda^* - \widehat{\Lambda}^*\right)'\Sigma_Z{}^{-1}\right]\right.$$

$$\left.-\frac{1}{2}(\delta - \delta^0)'\Sigma_\delta^{0^{-1}}(\delta - \delta^0)\right\}$$

$$\propto exp\left\{-\frac{1}{2}tr\left[\left(\lambda - \hat{\lambda}\right)'\left[\Sigma_Z{}^{-1} \otimes Fac'Fac\right]\left(\lambda - \hat{\lambda}\right)\right]\right.$$

$$\left.-\frac{1}{2}(\delta - \delta^0)'\Sigma_\delta^{0^{-1}}(\delta - \delta^0)\right\}.$$

Let $\mu_\delta = (B'B)^{-1}B'\left(\hat{\lambda} - C\right)$, and $\Sigma_\delta^{-1} = B'\left(\Sigma_Z{}^{-1} \otimes Fac'Fac\right)B$. Given equation (47), we have

$$p(\delta|Z, Fac, \Sigma_Z) \propto exp\left\{-\frac{1}{2}tr[(\delta - \mu_\delta)'\Sigma_\delta^{-1}(\delta - \mu_\delta)] - \frac{1}{2}(\delta - \delta^0)'\Sigma_\delta^{0^{-1}}(\delta - \delta^0)\right\}$$

$$\propto exp\left\{-\frac{1}{2}(\delta - \mu_\delta^*)'\Sigma_\delta^{*^{-1}}(\delta - \mu_\delta^*)\right\},$$

which is $N(\mu_\delta^*, \Sigma_\delta^*)$ where

$$\mu_\delta^* = \Sigma_\delta^*\left(\Sigma_\delta^{0^{-1}}\delta^0 + \Sigma_\delta^{-1}\mu_\delta\right),$$

$$\Sigma_\delta^* = \left(\Sigma_\delta^{0^{-1}} + \Sigma_\delta^{-1}\right)^{-1}.$$

Step 3. Sampling model diffusion parameters R and Q conditional on drift parameters and latent factors.

Since v_t^* in equation (40) and w_t in equation (41) are uncorrelated, we can sample R and Q separately. In particular, we sample R from Y conditional on $\Lambda_R = \begin{bmatrix} \mu_y^* \\ \Lambda_0' \\ \Lambda_1' \end{bmatrix}$ and latent variables, and Q conditional on G and latent variables.

We assume an inverse Wishart prior distribution for both R and Q:

$$R \sim W^{-1}(b_R^0, B_R^0),$$

$$Q \sim W^{-1}\left(b_Q^0, B_Q^0\right),$$

The likelihood functions are

$$p\left(Y|Fac, \mu_y^*, \Lambda_0, \Lambda_1, R\right) \propto |R|^{-\frac{T}{2}} exp\left\{-\frac{1}{2}(Y - Fac\Lambda_R)' R^{-1}(Y - Fac\Lambda_R)\right\}$$

$$\propto |R|^{-\frac{T}{2}} exp\left\{-\frac{1}{2} tr(S_R R^{-1})\right\},$$

$$p(X|H, G, Q) \propto |Q|^{-\frac{T}{2}} exp\left\{-\frac{1}{2}(X - HG)' Q^{-1}(X - HG)\right\}$$

$$\propto |Q|^{-\frac{T}{2}} exp\left\{-\frac{1}{2} tr\left(S_Q R^{-1}\right)\right\},$$

where $S_R = (Y - Fac\Lambda_R)'(Y - Fac\Lambda_R)$, and $S_Q = (X - HG)'(X - HG)$.

The conditional posteriors are given by:

$$p(R|Y, Fac, \Lambda_R) \propto p(Y|Fac, \Lambda_R)p(R)$$

$$\propto |R|^{-\frac{T + b_R^0 - col_{Fac}}{2}} exp\left\{-\frac{1}{2} tr[R^{-1}(S_R + B_R^0)]\right\},$$

$$p(Q|X, H, G) \propto p(X|H, G)p(G)$$

$$\propto |G|^{-\frac{T + b_G^0 - col_G}{2}} exp\left\{-\frac{1}{2} tr\left[R^{-1}\left(S_Q + B_Q^0\right)\right]\right\},$$

which implies that

$$p(R|Y, Fac, \Lambda_R) \sim W^{-1}(b_R^*, B_R^*),$$

$$p(Q|X, H, G) \sim W^{-1}\left(b_Q^*, B_Q^*\right),$$

where $b_R^* = T + b_R^0 - col_{Fac}, B_R^* = S_R + B_R^0, b_Q^* = T + b_G^0 - col_G,\ B_Q^* = S_Q + B_Q^0$, $col_{Fac} = 5$ is the number of columns in Fac, and $col_G = 2$ is the number of columns in G.

Step 4. Mapping parameter and latent variable estimates back to the original model (38) – (39).

Once we get parameter and latent variable estimates $\left\{\hat{\mu}^*_{y_{3\times1}}, \hat{\lambda}_5, \hat{\lambda}_6, \hat{Q}_c, \hat{Q}_d{}^*, \hat{\phi}^*, \hat{k}^*, \hat{\rho}^*\right\}$, $\hat{\theta}^*_t$, $\hat{R}$, and $\hat{Q}$ of model (42) – (43), we need to do the following conversions to map parameter and latent variable estimates back to the original model (38) – (39):

$$\hat{\Sigma}_\theta = \hat{Q},$$

$$\hat{\Sigma}_{y\theta} = \hat{\Lambda}_0\hat{Q},$$

$$\hat{\Sigma}_y = \hat{R} + \hat{\Sigma}_{y\theta}\hat{\Sigma}_\theta^{-1}\hat{\Sigma}'_{y\theta},$$

$$\hat{k} = 1 - \hat{k}^*,$$

$$\hat{\rho} = 1 - \hat{\rho}^*$$

$$\hat{Q}_d = \hat{Q}_d^* + \hat{\lambda}_5\hat{k}^*,$$

$$\hat{\phi} = \hat{\phi}^* + \hat{\lambda}_6\hat{\rho}^*,$$

$$\widehat{m}^* = \hat{\mu}^*_{y,1},$$

$$\hat{\mu}_c = \hat{\mu}^*_{y,2} - \hat{Q}_c\widehat{m}^*,$$

$$\hat{\mu}_d = \hat{\mu}^*_{y,3} - \hat{Q}_d\widehat{m}^*,$$

$$\hat{\theta}_t = \hat{\theta}^*_t + \begin{bmatrix}\widehat{m}^*\\ 0\end{bmatrix}.$$

3.7 Empirical Results

3.7.1. Data

The data sample contains log inflation, $\Delta\pi$, log real consumption growth, Δg_c, and log real dividend growth, Δg_d, at monthly frequency. The data span the period from

February 1959 to December 2014. Log inflation $\Delta\pi$ is computed from the Consumer Price Index (CPI)[23]. Log real consumption growth, Δg_c, is computed from monthly personal consumption[24] deflated by its chain-type price index[25]. Monthly dividend data on the S&P 500 are retrieved from Robert Shiller's website[26] and is deflated by CPI to compute the log real dividend growth, Δg_d.

Table 3.1 presents the summary statistics of the monthly observations on the state variables of the model. Both log inflation, $\Delta\pi$, and log real dividend growth, Δg_d, are persistent with first-order autocorrelation coefficients, ρ, 0.6259 and 0.6886, respectively. Log real consumption growth, Δg_c, has a small negative $\rho = -0.1284$. In figure 3.1, we plot the time series of the state variables. According to the plot, there are three interesting sub-periods in our sample period. The first sub-period is from mid 70s to early 80s. During this period, all three state variables are very volatile. While dividend growth and inflation in general move in the same direction, consumption growth seems to be negatively correlated with the other two state variables. The second sub-period of interest is the most current recession period, staring from late 2008. The most dramatic phenomenon over this period is the V-shape pattern of dividend growth. In the years before the recession, the average dividend growth is around 1%, which is more than twice its sample average. In late 2008, it drops sharply to below -2%. Once at the bottom (-2.82%), dividend growth bounces back to positive values. In total, it takes about 2 years for dividend growth to recover from the recession and climb back to its sample average.

[23] The Consumer Price Index for All Urban Consumers: All Items (CPIAUCSL), retrieved from Federal Reserve Economic Data (http://research.stlouisfed.org/fred2/series/CPIAUCSL).
[24] Personal Consumption Expenditures (PCE), retrieved from Federal Reserve Economic Data (https://research.stlouisfed.org/fred2/series/PCE).
[25] Personal Consumption Expenditures: Chain-type Price Index (PCEPI), retrieved from Federal Reserve Economic Data (https://research.stlouisfed.org/fred2/series/PCEPI).
[26] http://www.econ.yale.edu/~shiller/data/ie_data.xls.

We observe similar but less significant declines in inflation and consumption growth. Inflation stays below zero for less than half year. While the consumption growth recovery period is roughly of the same length as that of dividend growth, the decline in consumption growth is milder, with -0.53% at the maximum. Finally, in the 1990s, all three state variables are stable with standard deviations much smaller than that of the whole sample.

In figure 3.1B, we report the normal probability plots of the median residuals[27] of the model (38)-(39). The plots suggest, for all three state variables, the median residuals follow a fat-tailed distribution rather than the normal distribution. One possible explanation is that the conditional shocks on state variables have time-varying variances, which is consistent with the economic uncertainty in the original LRR model of Bansal and Yaron (2004). As an approximation, we still proceed with the constant variance assumption in this chapter in order to retain the closed-form solution for bond prices. Considerations of skewness and kurtosis in the state variables are beyond the scope of current work and are left for future research.

To compute the yield curve, we still need the average monthly consumption-wealth ratio h, which is not directly observable. To estimate the monthly consumption-wealth ratio, we follow Lettau and Ludvigson (2001b). Assuming the aggregate wealth W_t is the sum of the asset holdings, A_t, and human capital, H_t, Lettau and Ludvigson propose that the log aggregate wealth can be approximated as a weighted average of log asset holdings and log human capital:

$$w_t = \omega a_t + (1 - \omega) h_t.$$

[27] Using our Bayesian estimation approach, we obtain the empirical distributions (formed by15,000 draws) of the model residuals at each period in our sample period. However, to draw the normal probability plot, we only use the median for each residual at each period.

Table 3.1 Summary statistics for state variables.

This table reports the summary statistics of monthly state variables. The state variables are log inflation, $\Delta\pi$, log real consumption growth, Δg_c, and log real dividend growth, Δg_d. The sample period is February 1959 to December 2014. Std. error is the sample standard deviation. ρ is the first-order autocorrelation coefficient. ϕ is the first-order partial autocorrelation coefficient. ρ_{sqr} and ϕ_{sqr} are the first-order autocorrelation coefficient and partial autocorrelation coefficient for the squared data series.

Variable	$ln(i_t)$	Δg_c	Δg_d
Mean	0.0031	0.0023	0.0046
Std.error	0.0032	0.0057	0.0063
Min.	-0.0179	-0.0258	-0.0282
Max.	0.0179	0.0301	0.0236
ρ	0.6259	-0.1284	0.6886
ϕ	0.6286	-0.1278	0.6890
ρ_{sqr}	0.5588	0.2674	0.5861
ϕ_{sqr}	0.5589	0.2675	0.5861

Figure 3.1A Time series of state variables.

The state variables are log inflation, $\Delta\pi$, log real consumption growth, Δg_c, and log real dividend growth, Δg_d. The sample period is February 1959 to December 2014.

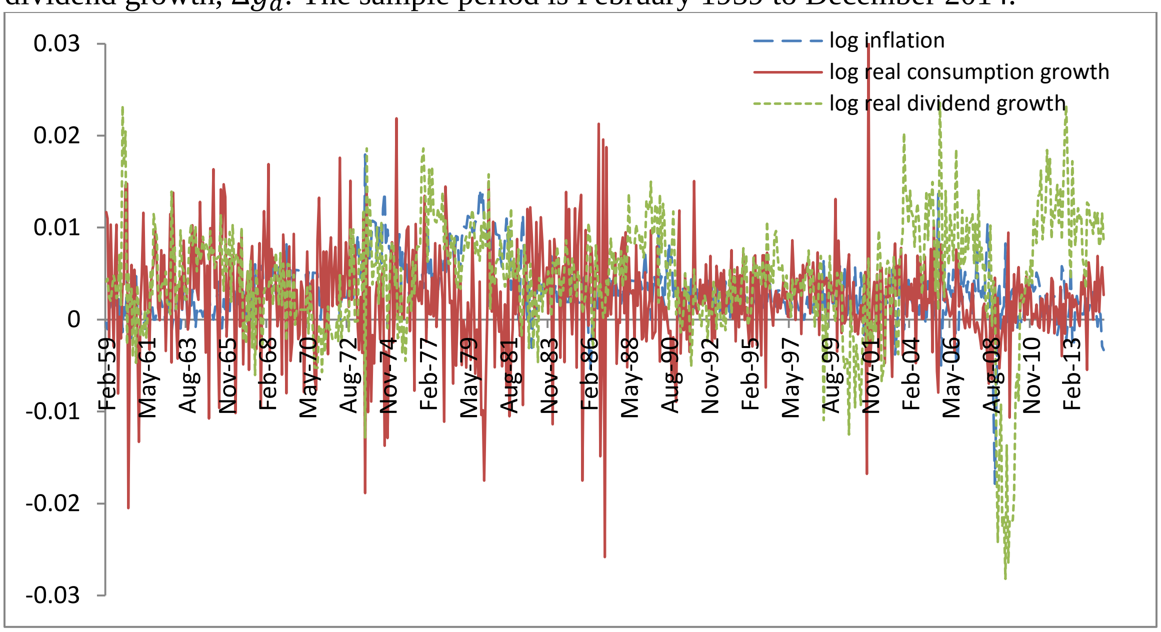
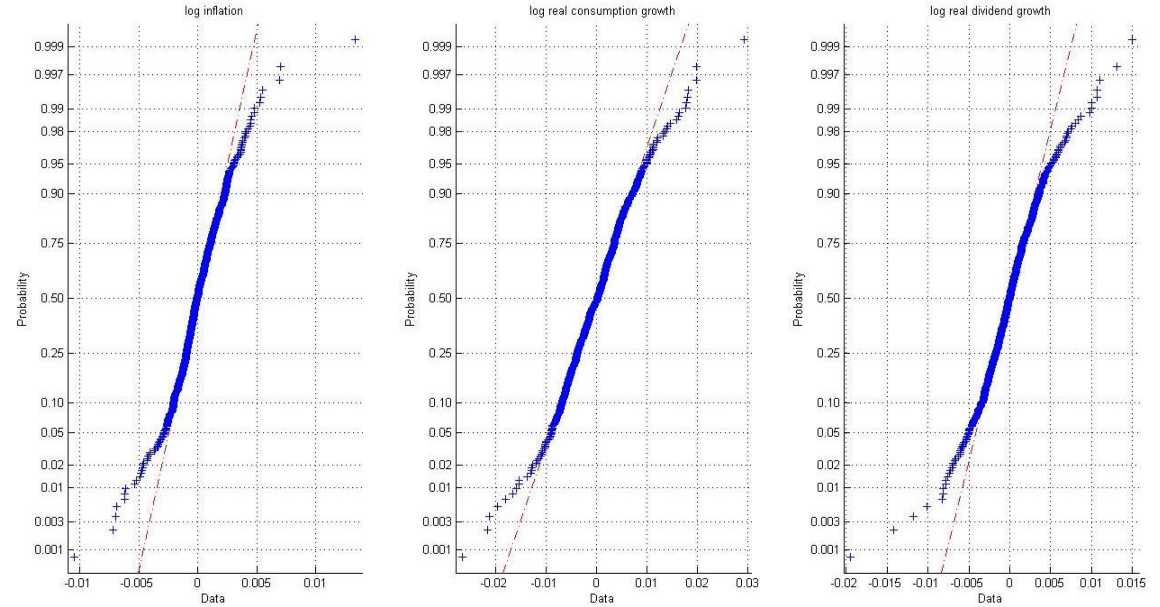


Figure 3.1B Normal probability plot of the median residuals of the model (38)-(39).

The state variables are log inflation, $\Delta\pi$, log real consumption growth, Δg_c, and log real dividend growth, Δg_d. The sample period is February 1959 to December 2014.



They show that the weight ω can be retrieved from the following regression (see Lettau and Ludvigson (2001b) for details):

$$c_t = \alpha + \beta_a a_t + \beta_y y_t + \sum_{i=-k}^{k} b_{a,i}\Delta a_{t-i} + \sum_{i=-k}^{k} b_{y,i}\Delta y_{t-i} + \epsilon_t, \tag{49}$$

where c_t is the personal consumption expenditure (PCE) per capita, y_t is the labor income per capita, β_a and β_y are proportional to ω and $(1-\omega)$, respectively, ($\beta_a = \left(1/_\lambda\right)\omega$, $\beta_y = \left(1/_\lambda\right)(1-\omega)$), and k is chosen to be 8.

According to equation (49), to estimate ω, we need monthly log consumption per capita, log asset holdings, and log labor income. We get both consumption and labor incomes from BEA (NIPA tables 2.8 and 2.6, respectively). The log asset holdings, however, is only available at a quarterly frequency on Martin Lettau's website. To produce monthly asset holdings from the quarterly data series, we follow Lamont (2001) and Vassalou (2003) and form the factor-mimicking portfolio of asset holdings growth using 25 Fama-French size and book-to-market portfolios, the one month T-bill rate, inflation, term spread, and default spread. Returns on 25 Fama-French portfolios and the one month T-bill rate are retrieved from Kenneth R. French's online data library[28]. Inflation is computed using chain-type price index of PCE[29]. Term spread is the difference between 1- and 10-year constant maturity rates, and default spread is the yield gap between Moody's seasoned Aaa corporate bond and Baa bond[30]. It turns out that the factor-mimicking portfolio approach is highly effective and can capture almost 80% of the variation in the asset holdings growth (the R-squared is 79.72%). We take the monthly returns on the factor-mimicking portfolio as a proxy for the monthly asset growth rates. Then, we can compute the monthly asset holding levels in the following way: for a given quarter, log asset holdings for the first month are the sum of log asset holdings at the end of the previous quarter and the first month's growth in asset holdings; log asset holdings of the second month are equal to those at the end of current quarter minus the growth in the third month of the quarter; and, finally, the third month's level is the end of the quarter level.

With the estimated time series of monthly asset holdings, we run regression (49). Using our monthly sample from January 1969 to September 2014, we estimate β_a to be 0.1538 and $\beta_y = 0.9724$[31], which implies a ω of 0.1403 and a sample average monthly consumption-wealth ratio of 0.14%. Assuming both aggregate wealth and consumption

[28] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.
[29] Downloaded from of the Federal Reserve Bank of St. Louis. (https://research.stlouisfed.org/fred2/series/PCECTPI).
[30] Both term spread and default spread data series are retrieved from Federal Reserve Bank of St. Louis.
[31] 95% confidence interval of β_a is (0.1222,0.1855), and that of β_y is (0.8994,0.9860).

are stable within each quarter, our monthly consumption-wealth ratio estimate implies an average log quarterly wealth-consumption ratio of $-ln(0.14\% \times 3) = 5.47$. In Lustig, van Nieuwerburgh, and Verdelhan (2013), the mean log quarterly wealth-consumption ratio is estimated to be 5.81 with the standard deviation of 0.49. Our point estimate is well within the 25% and 75% uncertainty bands of Lustig, van Nieuwerburgh, and Verdelhan (2013)'s estimate.

3.7.2 Implementing the Gibbs Sampler

When computing conditional posterior distributions of model parameters, we use conjugate prior distributions. In particular, we use normal priors for drift parameters, δ, and inverse Wishart priors for diffusion parameters R and Q:

$$\delta \sim N\left(\delta^0, \Sigma_\delta^0\right),$$

$$R \sim W^{-1}(b_R^0, B_R^0),$$

$$Q \sim W^{-1}\left(b_Q^0, B_Q^0\right).$$

Since we expect most of the elements in δ to be close to zero, we choose δ^0 to be equal to $0_{10\times 1}$. To make sure that B_R^0 is of the same scale as true R, we set B_R^0 equal to the sample variance-covariance matrix of log inflation, log consumption growth, and log dividend growth $B_R^0 = 10^{-5} * \begin{bmatrix} 1.00 & -0.59 & -0.15 \\ -0.59 & 3.28 & 0.36 \\ -0.15 & 0.36 & 3.98 \end{bmatrix}$. We expect Q to be of the same scale as R, so we set $B_Q^0 = 3 * 10^{-5} * I_2$. To make our priors uninformative, we set $\Sigma_\delta^0 = 5 * I_{10}$. For the same reason, b_R^0 and b_Q^0 are chosen to be 5 and 4, respectively.

To facilitate convergence diagnostics, we run three chains starting from 3 different sets of initial values.[32,33] The convergence diagnostics results are reported in Appendix A3.6. For each chain, we run 100,000 iterations, due to large autocorrelations of parameter draws such as those for μ_y^*. We discard 50,000 burn-in iterations to eliminate potential dependence on starting conditions. That leaves us with a total of $50{,}000 \times 3 = 150{,}000$ iterations. To further eliminate the autocorrelation in draws of μ_y^* and to reduce the computational burden, we thin the chain by keeping only every tenth draw. Thus, in the following analysis, we have a total of 15,000 post burn-in iterations for each model parameter and latent state variable.

3.7.3. Results

We present our results in two stages. First, we report parameter estimates, and examine the model implied nominal term-structure curve. Then, we study the impact of learning about latent state variable x_t on the properties of model-implied nominal yield curve. Evaluation of model performance is given in Appendix A3.7.

Table 3.2 highlights features of the estimated densities of our sixteen model parameters given observations of log inflation, $\Delta\pi$, log real consumption growth, Δg_c, and log real dividend growth on S&P 500, Δg_d. There are several interesting findings worth noting. Our estimated mean inflation, m_t, is not persistent. Its mean-reverting speed, k, is estimated to have a monthly median of 0.2315 with a standard deviation of 0.0426. This implies a $(1-k)^{12} = 0.0424$ annual persistence parameter of m_t. More importantly, the LRR variable, x_t, is not persistent either. The median of monthly

[32] Initial values: $\delta_i^{init} = a_i * [0.0031, 0.0023, 0.0046, 0.5, 1, 1{,}2, 3{,}0.9, 0.9]$, $R_i^{init} = b_i * B_0^R$, and $Q_i^{init} = b_i * B_0^Q$, where $a = [0.5{,}1{,}1.5]$, $b = [0.5{,}1{,}2]$, and $i = 1{,}2{,}3$.

[33] Gelman-Rubin Statistic- a multiple-chain convergence diagnostic - for each parameter are reported in table A3.6.1 in Appendix A3.6.

persistence parameter is 1- 0.0703 = 0.9297[34], and the corresponding half-life is less than one year.

Table 3.2 Estimated Densities of Model Parameters.
This table highlights features of the estimated densities of the sixteen model parameters based on the 671 months of log inflation, log real monthly consumption growth, and log real monthly dividend growth. The parameters are as defined in equation (38) – (39). Std. error is the sample standard deviation.

Parameter	2.5%ile	25%ile	Median	Mean	75%ile	97.5%ile	Std. error
m^*	0.0024	0.0029	0.0031	0.0031	0.0033	0.0038	0.0003
μ_c	0.0020	0.0036	0.0045	0.0044	0.0053	0.0068	0.0012
μ_d	0.0069	0.0098	0.0110	0.0110	0.0122	0.0147	0.0019
Q_c	-1.0785	-0.8340	-0.6981	-0.6476	-0.5130	-0.0034	0.2747
Q_d	-2.7847	-2.2152	-2.0089	-1.9878	-1.8016	-0.9552	0.4145
ϕ	1.2238	1.7439	2.0382	2.1977	2.4569	4.3252	0.7307
k	0.1514	0.2040	0.2315	0.2321	0.2586	0.3219	0.0426
ρ	0.0178	0.0474	0.0703	0.1027	0.1174	0.3837	0.0916
σ_i	0.0019	0.0020	0.0021	0.0021	0.0021	0.0023	0.0001
$\sqrt{V}$	0.0057	0.0062	0.0064	0.0064	0.0066	0.0069	0.0003
φ_d	0.4961	0.5316	0.5532	0.5551	0.5762	0.6253	0.0329
σ_m	0.0012	0.0015	0.0016	0.0016	0.0017	0.0019	0.0002
φ_e	0.1702	0.2020	0.2225	0.2253	0.2452	0.2973	0.0325
ρ_{ic}	-0.3184	-0.2476	-0.2095	-0.2128	-0.1765	-0.1174	0.0518
ρ_{id}	-0.9459	-0.9310	-0.9199	-0.9151	-0.9050	-0.8545	0.0235
ρ_{im}	0.6083	0.7159	0.7549	0.7466	0.7867	0.8348	0.0572
ρ_{ix}	-0.7057	-0.3802	-0.2656	-0.2862	-0.1642	0.0139	0.1778
ρ_{cd}	0.0677	0.1307	0.1651	0.1670	0.2011	0.2775	0.0531
ρ_{cm}	-0.2019	-0.1225	-0.0794	-0.0874	-0.0475	0.0040	0.0552
ρ_{cx}	0.1702	0.2020	0.2225	0.2253	0.2452	0.2973	0.0325
ρ_{dm}	-0.8575	-0.8101	-0.7824	-0.7803	-0.7531	-0.6893	0.0427
ρ_{dx}	0.0827	0.2825	0.3820	0.3867	0.4926	0.6878	0.1551
ρ_{mx}	-0.8764	-0.5122	-0.3584	-0.3801	-0.2234	0.0197	0.2264

The low persistence of x_t and m_t suggests that the LRR model may provide little help in amplifying the unconditional variance of consumption growth. To see this, consider a simple 1-dimentional model in which log consumption growth Δc_t has a mean-reverting mean, m_c:

[34] This is consistent with Constantinides and Ghosh (2011) in which the annual persistence of the LRR component is 0.437.

$$\Delta c_t = m_{c,t-1} + \sigma_c \epsilon_t,$$
$$m_{c,t-1} = \rho m_{c,t-2} + \sigma_m \varepsilon_{t-1},$$

Then the unconditional variance of Δc_t, Var_c is given by:

$$Var_c = \sigma_c^2 + var(m_{c,t-1}) + 2\sigma_c cov(m_{c,t-1}, \epsilon_t). \tag{50}$$

When persistence parameter ρ is close to zero, $var(m_{c,t-1})$ is approximately equal to σ_m^2. Assuming $cov(m_{c,t-1}, \epsilon_t) = 0$ and the conditional volatility of the latent variable negligible relative to the conditional volatility of log consumption growth, the latter must have an unconditional variance approximately equal to its conditional variance. In our results, estimated mean of log consumption growth, $Dif_{g_c,t} = \mu_c + x_{t-1} + Q_c m_{t-1}$, has an annual persistence parameter of 0.0483. Therefore, the unconditional variance of consumption growth cannot be much larger than the sum of conditional consumption growth variance and the unconditional variance of $Dif_{g_c,t}$. The variance-amplifying mechanism of LRR model is further weakened by correlations between innovations in observables and latent variables. In equation (50), when $cov(m_{c,t-1}, \epsilon_t) < 0$, it is possible that Var_c is even smaller than σ_c^2. In our sample, estimated $Dif_{g_c,t}$ is negatively correlated (-0.4005) with innovations in Δg_c, which leads to a conditional standard deviation $(\sqrt{V} = 0.0064)$ larger than the sample standard deviation(0.0057).

Further, the feedback effects of mean inflation on real mean growth rates in fundamentals, Q_c and Q_d, are estimated to be negative. This result confirms the negative signal effect of inflation on consumption growth (Piazzesi and Schneider (2007)) and is consistent with empirical results of Piazzesi and Schneider (2007), Eraker (2008), and Doh (2012). In these studies, such effect helps to produce a positively-sloped yield curve.

However, in our estimation, it is not large enough to make the nominal yield curve upward-sloping when correlations among innovations are considered.

To compute the model-implied term structure, we still need values of risk aversion γ and average consumption-wealth ratio h in the value function. We take the relative risk aversion to be equal to 10 as reported by Constantinides and Ghosh (2011).

Table 3.3 reports the estimated sample mean and standard deviation of the model-implied nominal short rate, 1-month yield, 1-year yield, and 5-year yield without incomplete information. The most important result is that the model-implied nominal yield curve is downward-sloping. The sample mean of the nominal short rate has a median of 0.6357% per month, which amounts to 7.6284% per year, while the sample mean of the 5-year yield has an annualized median of 4.062%. This result is at odds with results of previous studies (i.e. Piazzesi and Schneider (2007), Eraker (2008), and Doh (2012)), and is due to the fact that the risk premium (88) (appendix A3.3) is negative. According to equation (88), the risk premium contains two parts corresponding to shocks in mean consumption ($\lambda_x^c = -B^c\lambda_1^c$) and shocks in mean inflation ($\lambda_m^c = -C^c\lambda_2^c$). Given estimated value of Q_c, k and ρ, $-B^c$ and $-C^c$ are positive and $-B^c > -C^c$. As a result, signs of λ_x^c and λ_m^c are determined by λ_1^c and λ_2^c, respectively. Our parameter estimates imply that λ_1^c is negative and has a larger absolute value (with sample median of -0.0315%) than λ_2^c (0.0161%), resulting in negative risk premium. And since $-B^c$ and $-C^c$ are both monotonically increasing functions of time to maturity ($\tau = s - t$), the risk premium is more negative for zero-coupon bond with longer time to maturity, and leads to a smaller yield. The sample standard deviations of mean yields reported in Table 3.3 are large because of a few extreme values in the data. We also report standard deviations computed by using only values within the 2.5% and 97.5% bands. The standard deviation of mean yield is increasing in time to maturity. This is because B^c and C^c in equations (85) and (86) are more sensitive to the change in k and ρ coefficients when time to maturity is larger.

Table 3.3 Model-implied Nominal Term Structure.
This table highlights features of the estimated densities of sample mean and standard deviation of model-implied nominal short rate r_f, 1-month yield y_1, 1-year yield y_{12}, and 5-year yield y_{60} based on the 671 months of log inflation, log real monthly consumption growth, and log real monthly dividend growth spanning period from February 1959 to December 2014. The relative risk aversion $\gamma = 10$. The average monthly consumption-wealth ratio $h = 0.14\%$. Other parameter values are given in table 3.2. Std. error is the sample standard deviation. Std. error 95% is the sample standard deviation computed using only values within 95% bands. All numbers in this table are percentages.

	2.5%ile	25%ile	Median	Mean	75%ile	97.5%ile	Std. error	Std. error 95%
$E(r_f)$	0.5643	0.6150	0.6357	0.6263	0.6551	0.6890	0.6873	0.0263
$\sigma(r_f)$	0.1509	0.3011	0.3244	0.3128	0.3440	0.3799	0.0527	
$E(y_1)$	0.5319	0.6020	0.6269	0.6120	0.6495	0.6888	1.0548	0.0320
$\sigma(y_1)$	0.1305	0.2861	0.3117	0.2985	0.3322	0.3692	0.0560	
$E(y_{12})$	0.1113	0.4416	0.5183	0.4383	0.5833	0.6820	5.1354	0.1083
$\sigma(y_{12})$	0.0415	0.1636	0.2149	0.2008	0.2504	0.3103	0.0707	
$E(y_{60})$	-1.2687	0.0914	0.3385	-0.0488	0.5087	0.6831	23.5020	0.3516
$\sigma(y_{60})$	0.0085	0.0432	0.0743	0.0807	0.1078	0.2056	0.0531	

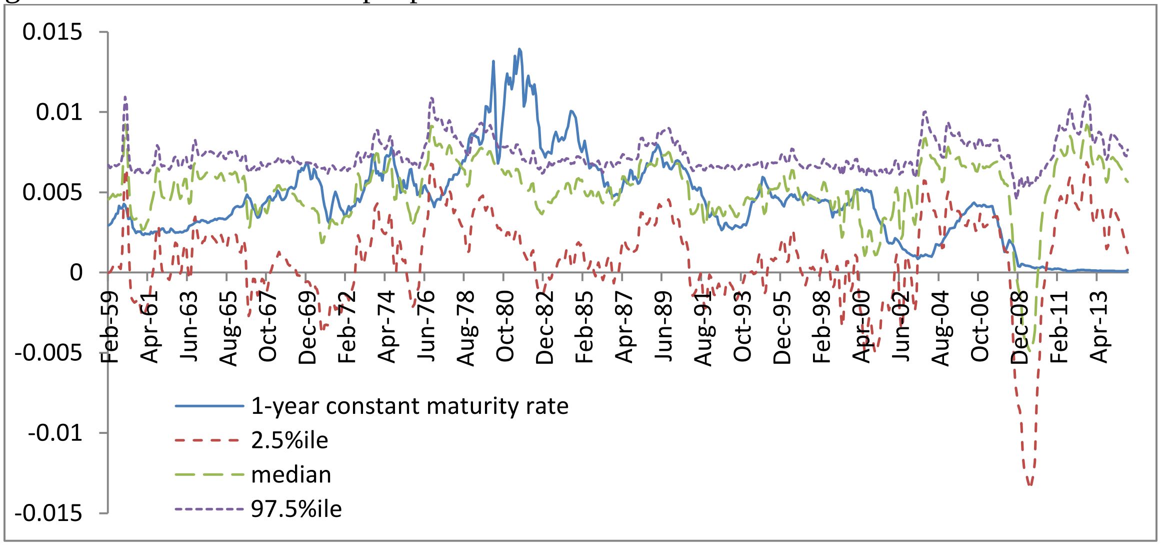
In figure 3.2, we plot the observed 1-year constant maturity rate[35] with the model predicted 95% posterior bands and the median of 1-year yield. In general, the LRR model fails to capture the basic patterns of the observed 1-year yield. There are 25.78% (173 out 671) of observed 1-year yields outside the 95% predictive density bands. Most of them are concentrated in the early 80s, when the observed yield is high, and in the current recession period. In both periods, the model-implied 1-year yield largely replicates the

[35] The data series for 1-year and 5-year constant maturity rates are retrieved from Federal Reserve Bank of St. Louis. Since the original data series are annualized, we divide original values by 12 to get y_{12} in figure 3.2 and y_{60} in figure 3.3.

pattern in the state variables, especially the log dividend growth and log inflation. (This is due to the fact that the nominal yield is driven by latent variables, mean inflation and the LRR component. As we show in Appendix A3.7, these latent variables mainly match the log inflation and log real dividend growth, respectively.)

Figure 3.2 Time series of 1-year nominal yields.

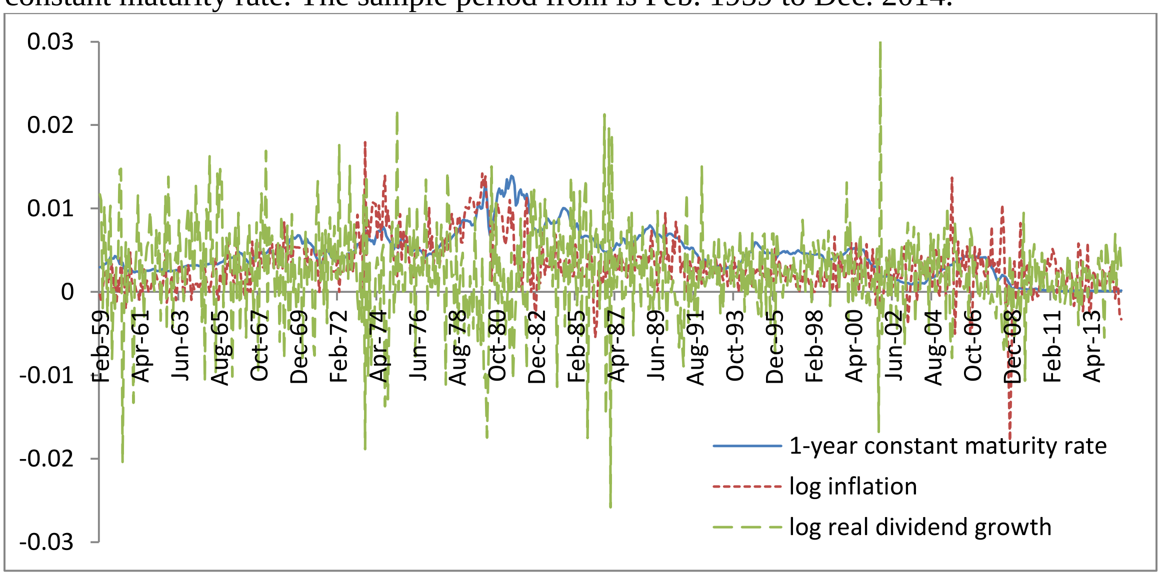
Time series of 1-year yield (1-year constant maturity rate) $y_{12}$with ninety-fifth percentile predictive density bands from the model (1) – (5) based on the 671 months of log inflation, log real monthly consumption growth, and log real monthly dividend growth observations. y_{12} is reported as monthly value (divided by 12). The relative risk aversion $\gamma = 10$. The average monthly consumption to wealth ratio $h = 0.14\%$. Other parameter values are given in Table 3.2. The sample period is from Feb. 1959 to Dec. 2014.



In figure 3.3, we plot 1-year constant maturity rate with log inflation and log real dividend growth. It is apparent that neither state variable matches the fluctuations in 1-year constant maturity rate in the early 80s and the recent recession period. In the early 80s, although initially inflation and dividend growth move in the same direction as the 1-year yield, both state variables reach their turning point before the yield reaches its highest level. The declines in state variables lead to declines in the model-implied yield. As a result the model-implied 1-year yield is below the observed value over this period.

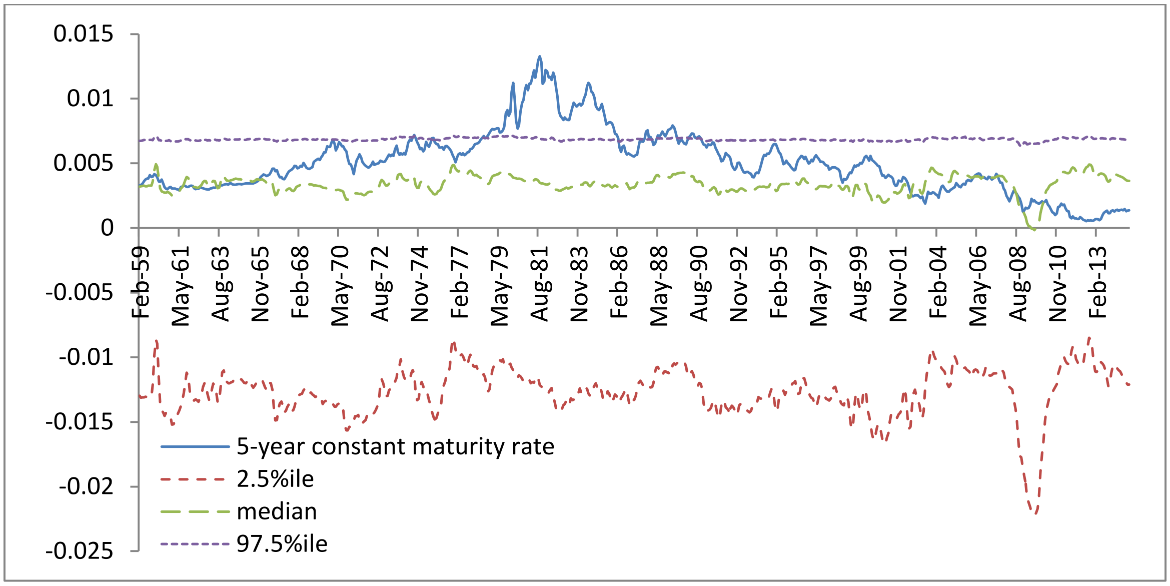
In the recent recession period, the large V-shape pattern in the log dividend growth dominates the model-implied 1-year yield by first pulling it down to negative values[36], and then pushing it back up above the observed value. The observed yield, however, is small and stable. Similar pattern can be seen also in figure 3.4, which is the 5-year constant maturity rate version of figure 3.2. In figure 3.4, the 95% density bands are much wider than those in figure 3.2. This is again because both B^c and C^c in equation (84) are more sensitive to changes in parameters k and ρ when time to maturity is larger, especially, when k and ρ are far away from zero. The wider bands lead to fewer observed values lying outside the bands (18.18% or 122 out of 671). However, this does not mean the model does a better job in reproducing the 5-year yield. In fact, the model largely misses the pattern in observed 5-year yield series with the posterior medians far below the observed values in most of the sample period.

Figure 3.3 Time series of log inflation, log real dividend growth, and 1-year constant maturity rate.
The figure superimposes time series of log inflation, log real dividend growth, and 1-year constant maturity rate. The sample period from is Feb. 1959 to Dec. 2014.



[36] We do not impose any equilibrium lower zero bound constraints on the rates.

Figure 3.4 Time series of 5-year yield y_{60}.
Time series of 5-year yield (5-year constant maturity rate) y_{60} with ninety-fifth percentile predictive density bands from the model (1) – (5) based on the 671 months of log inflation, log real monthly consumption growth, and log real monthly dividend growth observations. y_{60} is reported as monthly value (annualized value divided by 12). The relative risk aversion $\gamma = 10$. The average monthly consumption to wealth ratio $h = 0.14\%$. Other parameter values are given in table 3.2. The sample period is from Feb. 1959 to Dec. 2014.



Overall, the model with complete information fails to reproduce the observed nominal term structure. The model-implied yield curve has the wrong slope. The model-implied constant maturity rates cannot match the time series properties of corresponding observed yields.

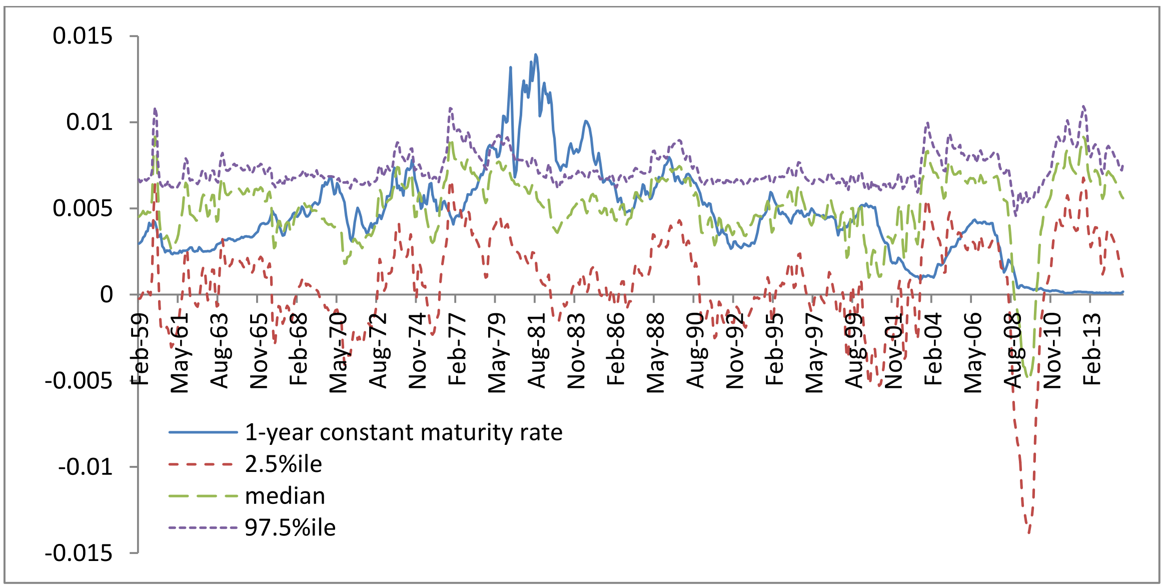
Now we study the model-implied term structure when mean consumption growth is not observed. Table 3.4 contains key features of the estimated sample mean and standard deviation of yields with different time to maturity. First and foremost, incorporating learning into the model does not make the yield curve upward-sloping. The median of mean short rate is 29.11 bps higher than that of 5-year yield. However, learning does shift down the whole yield curve and makes it flatter. Comparing results in Table 3.4 to corresponding values in Table 3.3, we can see that while learning has no observable impact on the yields' standard deviations, it decreases the average yields. The median short rate drops by 0.77 bps with unobserved LRR component. This gap reduces to 0.16 bps when time to maturity goes to 5 years, which leads to a smaller 5-year term spread (reduction of 0.77 – 0.16 = 0.61 bps). The decrease in unconditional mean short rate is a direct consequence of a smaller Ω^n (median = 0.5282%) compared to Ω^{cn} (median = 0.5375%). Meanwhile, the conditional covariance between the LRR component and the SDF reduces from -0.0291% to -0.0286%. This fact leads to a less negative risk premium. With the less negative risk premium, the yield gaps decrease with maturity as we observe across Table 3.3 and Table 3.4.

Table 3.4 Model-implied Term Structure With Incomplete Information.
This table highlights features of the estimated densities of sample mean and standard deviation of model-implied nominal short rate r_f, 1-month yield y_1, 1-year yield y_{12}, and 5-year yield y_{60} based on the 671 months of log inflation, log real monthly consumption growth, and log real monthly dividend growth from February 1959 to December 2014. The relative risk aversion $\gamma = 10$. The average monthly consumption-wealth ratio $h = 0.14\%$. Other parameter values are given in table 3.2. Std. error is the sample standard deviation. Std. error 95% is the sample standard deviation computed using only values within 95% bands. All numbers in this table are percentages.

	2.5%ile	25%ile	Median	Mean	75%ile	97.5%ile	Std. error	Std. error 95%
$E(r_f)$	0.5406	0.6053	0.6280	0.6150	0.6490	0.6854	0.8575	0.0294
$\sigma(r_f)$	0.1509	0.3011	0.3244	0.3128	0.3440	0.3799	0.0527	
$E(y_1)$	0.5070	0.5921	0.6192	0.6009	0.6439	0.6859	1.2255	0.0356
$\sigma(y_1)$	0.1305	0.2861	0.3117	0.2985	0.3322	0.3692	0.0560	
$E(y_{12})$	0.0892	0.4329	0.5130	0.4288	0.5804	0.6818	5.3089	0.1131
$\sigma(y_{12})$	0.0415	0.1636	0.2149	0.2008	0.2504	0.3103	0.0707	
$E(y_{60})$	-1.2928	0.0868	0.3369	-0.0551	0.5078	0.6830	23.687	0.3550
$\sigma(y_{60})$	0.0085	0.0432	0.0743	0.0807	0.1078	0.2056	0.0531	

Even though learning does 'steepen' the yield curve by lowering the short end relative to the long end, the effect is tiny and we conclude that the estimation risk is not enough to produce an upward-sloping yield curve. In figure 3.5, we again plot the observed 1-year yield with its 95% model predicted density bands as in figure 3.2. It is clear that, even with learning, the model still cannot explain the large yield in early 80s and near-zero yields after 2008.

Figure 3.5 Time series of 1-year yield y_{12} with incomplete information.
Time series of 1-year yield (1-year constant maturity rate) $y_{12}$with ninety-fifth percentile predictive density bands from the model (6) – (11) based on the 671 months of log inflation, log real monthly consumption growth, and log real monthly dividend growth observations. y_{12} is reported as monthly value (divided by 12) in this figure. The relative risk aversion $\gamma = 10$. The average monthly consumption-wealth ratio $h = 0.14\%$. Other parameter values are given in table 3.2. The sample period is from Feb. 1959 to Dec. 2014.



3.8 Conclusion

We extend the LRR model of Bansal and Yaron (2004) to include inflation and learning about the LRR component in the mean of state variables. Our work in this chapter makes three contributions to the literature.

Theoretically, the model provides closed form solutions for the term structure of interest rates and the instantaneous risk premium under a new risk-neutral measure corresponding to information available to investors.

Empirically, we develop a Bayesian MCMC method that allows us to estimate not only latent variables and parameters of state processes simultaneously, but also conditional correlations between latent and state variables. Moreover, we estimate our model using only the information for parameter and state variable estimates coming from financial and macro-economic data on aggregate consumption and dividend growth rates and inflation. By doing so, we estimate parameters and latent state variables outside the model, i.e., we use no price information for parameter and state variable inference.

The third contribution is that our estimation results provide empirical evidence contrary to the base case LRR model (without economic uncertainty and estimation risk). First of all, the LRR component is less persistent than it is assumed/calibrated in previous studies. Our estimate of first-order autocorrelation in the LRR component is around 0.417. More importantly, our estimates of conditional correlations between state variables suggest a negative conditional covariance between the pricing kernel and the LRR component and a positive conditional covariance between the pricing kernel and the mean inflation. The result implies the negative risk premium on zero-coupon bonds. In our base case model the risk premium on zero-coupon bonds has two components: a larger contribution from the hedging demand induced by the variation in the LRR component, and a minor contribution from hedging demand due to the mean inflation risk. Since the zero-coupon bond return loads negatively on LRR component and mean inflation, the LRR component risk premium is negative and the mean inflation risk premium is positive. The sum of the two risk premiums leads to a negative total risk premium (-21.3 bps per month for the 1-year zero coupon bond) and a downward sloping nominal yield curve.

Learning reduces the LRR component's conditional covariance with the pricing kernel and has a positive overall impact on the total risk premium. However, the effect is small and yield curves still retain the negative slope even after we incorporate learning into the model. When we turn on the uncertainty (and learning), another result is a reduction in the mean short rate. The monthly short rates with and without complete information are 0.6357% and 0.6280%, respectively. The reduction is mainly due to changes in the precautionary saving effect and adjustments for inflation in both models. The combination of these two effects shifts the yield curve downward and makes it flatter.

In a future study, we can incorporate both the economic uncertainty and jumps into our model and study asset pricing implications of learning in the extended model. However, doing so leads to only numerical solution to the model, which would require the estimation of the extended model using more advanced methods such as Metropolis-Hastings algorithm.

A3.1 Derivation of the Value Function J

With the log-linear approximation (16), the PDE (14) is a parabolic PDE with coefficients affine in states, which implies an exponential affine solution to the PDE:

$$J(t, \ln(i), g, d, \hat{x}, m) = \exp\{\xi_{0t} + \xi_{1t} \ln(i) + \xi_{2t} g_c + \xi_{3t} g_d + \xi_{4t} m + \xi_{5t} \hat{x}\}. \qquad (51)$$

Upon the substitution of (51) into (14) the PDE becomes an identity that must hold for arbitrary values of state variables:

$$\partial_t \xi_{0t} + G^T \partial_t \begin{bmatrix} \xi_{1t} \\ \xi_{2t} \\ \xi_{3t} \\ \xi_{4t} \\ \xi_{5t} \end{bmatrix} + \hat{\mu}_G^T \begin{bmatrix} \xi_{1t} \\ \xi_{2t} \\ \xi_{3t} \\ \xi_{4t} \\ \xi_{5t} \end{bmatrix} + \frac{1}{2dt} tr\left(dGdG^T \frac{J_{GG}}{J} \right) + \frac{f}{J} = 0. \tag{52}$$

Since $\frac{1}{2dt}\ tr(dGdG^T)$ is a function of only diffusion parameters and $\frac{1}{2dt} tr\left(\frac{J_{GG}}{J}\right)$ is a function of ξs, the fourth term on the left-hand side of PDE (52) contains no state variables.

With the log-linear approximation of the normalized aggregator given by equation (16), we also have

$$\frac{f}{J} = h(1-\gamma)\left[H - \frac{\xi_{0t}}{1-\gamma} + G' \begin{pmatrix} -\frac{\xi_{1t}}{1-\gamma} \\ 1 - \frac{\xi_{2t}}{1-\gamma} \\ -\frac{\xi_{3t}}{1-\gamma} \\ -\frac{\xi_{4t}}{1-\gamma} \\ -\frac{\xi_{5t}}{1-\gamma} \end{pmatrix} \right]. \tag{53}$$

Since equation (52) must hold for all values of state variables, coefficients on the state variables must all be zeros. Therefore, after collecting terms containing state variable, we have the following identity:

$$G^T\left[\partial_t\begin{bmatrix}\xi_{1t}\\ \xi_{2t}\\ \xi_{3t}\\ \xi_{4t}\\ \xi_{5t}\end{bmatrix}+\begin{bmatrix}000 & 1 & 0\\ 000 & Q_c & 1\\ 000 & Q_d & \phi\\ 000 & -k & 0\\ 000 & 0 & -\rho\end{bmatrix}^T\begin{bmatrix}\xi_{1t}\\ \xi_{2t}\\ \xi_{3t}\\ \xi_{4t}\\ \xi_{5t}\end{bmatrix}+h(1-\gamma)\begin{pmatrix}-\frac{\xi_{1t}}{1-\gamma}\\ 1-\frac{\xi_{2t}}{1-\gamma}\\ -\frac{\xi_{3t}}{1-\gamma}\\ -\frac{\xi_{4t}}{1-\gamma}\\ -\frac{\xi_{5t}}{1-\gamma}\end{pmatrix}\right]=0. \tag{54}$$

Identity (54) implies the following ODEs

$$\partial_t\xi_{1t}-h\xi_{1t}=0, \tag{55}$$

$$\partial_t\xi_{2t}+h(1-\gamma)-h\xi_{2t}=0, \tag{56}$$

$$\partial_t\xi_{3t}-h\xi_{3t}=0, \tag{57}$$

$$\partial_t\xi_{4t}+\xi_{1t}+Q_c\xi_{2t}+Q_d\xi_{3t}-(k+h)\xi_{4t}=0, \tag{58}$$

$$\partial_t\xi_{5t}+\xi_{2t}+\phi\xi_{3t}-(\rho+h)\xi_{5t}=0. \tag{59}$$

The ODEs are subject to boundary conditions

$$\xi_{1t}(T)=\xi_{2t}(T)=\xi_{3t}(T)=\xi_{4t}(T)=\xi_{5t}(T)=0. \tag{60}$$

Solving the system (55) – (60) for an infinitely lived agent, we obtain solutions (18) – (21).

A3.2 Derivation of the Pricing Kernel.

The normalized Porteus-Kreps aggregator has the following approximate form:

$$f=h(1-\gamma)J\left[lnC-\frac{1}{(1-\gamma)}lnJ+H\right].$$

Applying Ito’s lemma to (22) we have

$$\frac{d\Pi}{\Pi} = f_J dt + \frac{df_C}{f_C} \tag{61}$$

$$= \left(\frac{f}{J} - h\right)dt - \frac{dC}{C} + \frac{dJ}{J} + \frac{(dC)^2}{C^2} - \frac{dC}{C}\frac{dJ}{J}.$$

Further,

$$\frac{dC}{C} = (\mu_c + Q_c m_t + \hat{x}_t)dt + \sqrt{V}dw_c^*, \tag{62}$$

$$dJ = -fdt + J_{ln(i)}\sigma_i dw_i^* + J_C C\sqrt{V}dw_c^* + J_d\varphi_d\sqrt{V}dw_d^* + J_m\sigma_m dw_m^* + J_{\hat{x}}\Sigma_x dw_y^*, \tag{63}$$

$$\frac{J_{ln(i)}}{J} = \frac{J_d}{J} = 0, \frac{J_C}{J} = \frac{\xi_2}{C}, \frac{J_m}{J} = \xi_4, \frac{J_{\hat{x}}}{J} = \xi_5 \tag{64}$$

Combining (62), (63), and (64) with (61), we have the expression for the real pricing kernel:

$$\frac{d\Pi}{\Pi} = -[\Omega + \hat{x}_t + Q_c m]dt + \sigma_\Pi^T \begin{bmatrix} dw_i^* \\ dw_c^* \\ dw_d^* \\ dw_m^* \end{bmatrix},$$

where

$$\Omega = h + \mu_c + (\xi_2 - 1)V + \xi_5\sqrt{V}\Sigma_x[\rho_{ic} \quad 1 \quad \rho_{cd} \quad \rho_{cm}]' + \xi_4\sigma_m\sqrt{V}\rho_{mc},$$

$$\sigma_\Pi = \begin{bmatrix} \xi_5\Sigma_{1x} \\ (\xi_2 - 1)\sqrt{V} + \xi_5\Sigma_{2x} \\ \xi_5\Sigma_{3x} \\ \xi_5\Sigma_{4x} + \xi_4\sigma_m \end{bmatrix}.$$

A3.3 Derivation of the Zero Coupon Bond Price.

By the definition of the state-price process

$$P(t,s) = \frac{E_t \Pi_s^n}{\Pi_t^n} = \frac{E_t(\Pi_s/i_s)}{\Pi_t/i_t} \tag{65}$$

$$= \frac{E_t\left(f_{C_s} exp\left(\int_t^s f_J du\right)/exp\{ln(i_s)\}\right)}{f_{C_t}/exp\{ln(i_t)\}}.$$

From equation (16) – (21) we have

$$f_{C_t} = \frac{h(1-\gamma)}{C_t} \exp\{\xi_{0t} + \xi_{2t} g_c + \xi_{4t} m + \xi_{5t} \hat{x}\} \tag{66}$$

$$f_{J_t} = h(1-\gamma) ln(C_t) - h(\xi_{0t} + \xi_{2t} g_c + \xi_{4t} m + \xi_{5t} \hat{x}) + h(1-\gamma)H - h \tag{67}$$

Inserting (66) and (67) into (65), we are able to separate consumption from other variables in the bond price:

$$P(t,s) = E_t exp\left\{-(lnC_s - lnC_t) + (\xi_{0s} - \xi_{0t}) + (\xi_{2s} lnC_s - \xi_{2t} lnC_t)\right. \tag{68}$$

$$\left. + (\xi_{4s} m_s - \xi_{4t} m_t) + (\xi_{5s} \hat{x}_s - \xi_{5t} \hat{x}_t) - (lni_s - lni_t) + \int_t^s f_J du \right\}$$

$$= E_t Z_C Z_{\hat{x}m},$$

where Z_C is a function of only consumption, and $Z_{\hat{x}m}$ is a function of $\hat{x}$ and m.

Since $\frac{dC}{C} = (\mu_c + Q_c m_t + \hat{x}_t)dt + \sqrt{V} dw_c^*$, and $\frac{di}{i} = m_t dt + \sigma_i dw_i^*$, $F = ln(C_s/C_t)$ is only a function of x, m, and $\tau = s - t$, and $ln(i_s/i_t)$ is a function of m, and $\tau = s - t$.

Therefore,

$$\xi_{2s} lnC_s = \xi_{2s} lnC_t + \xi_{2s} F \tag{69}$$

$$\int_t^s f_J du = \int_t^s [(-h(\xi_{0u} + \xi_{4u}m + \xi_{5u}x) + h(1-\gamma)H - h) + h(1-\gamma)lnC_u \qquad (70)$$
$$- h\xi_{2u}lnC_u]\, du$$
$$= \int_t^s [K_u + h(1-\gamma)(lnC_t + F) - h\xi_{2u}(lnC_t + F)]\, du$$
$$= \widetilde{K} + h(1-\gamma)lnC_t \int_t^s \left(1 - \frac{1}{1-\gamma}\xi_{2u}\right) du$$
$$+ h(1-\gamma) \int_t^s \left(F - \frac{1}{1-\gamma}\xi_{2u}F\right) du$$

where K_u and $\widetilde{K} = \int_t^s \left[K_u + h(1-\gamma)\left(F - \frac{1}{1-\gamma}\xi_{2u}F\right)\right] du$ are functions of $x, m,$ and τ.

Inserting expressions for $\xi_{2s}lnC_s$ and $\int_t^s f_J du$ into equation (68) and collecting all lnC_t terms, we have:

$$Z_C = exp\left\{lnC_t\left[(\xi_{2s} - \xi_{2t}) + h(1-\gamma)\int_t^s \left(1 - \frac{1}{1-\gamma}\xi_{2u}\right) du\right]\right\} = 1.$$

Thus the zero coupon bond price $P(t,s)$ is a function of only $x, m,$ and $\tau = s - t$.

With the nominal SDF given in (25), the zero coupon bond price must satisfy the following PDE:

$$E\left(\frac{dP}{P}\right) = r_f^n dt - E\left(\frac{dP}{P}\frac{d\Pi^n}{\Pi^n}\right).$$

We look for a time-*t* bond price solution in the form of a function of $\hat{x}, m,$ and t:

$$P = P(\hat{x}, m, t).$$

Applying Itô's rule to P leads to:

$$\frac{dP}{P} = \frac{1}{P}\left(P_t + P_{\hat{x}}d\hat{x} + P_m dm + \frac{1}{2}(P_{\hat{x}\hat{x}}d\hat{x}^2 + 2P_{\hat{x}m}d\hat{x}dm + P_m dm^2)\right), \qquad (71)$$

$$\frac{dP}{P}\frac{d\Pi^n}{\Pi^n} = \frac{P_{\hat{x}}}{P}\sigma_{\Pi^n}^T \Lambda_{1:4,1:4}\Sigma_x' + \frac{P_m}{P}\sigma_m \sigma_{\Pi^n}' \Sigma_{ym}. \tag{72}$$

Collecting all the terms in (71) and (72), taking expectations, and dividing throughout by *dt*, we arrive at the PDE for the bond price:

$$\frac{P_t}{P} - \rho\hat{x}_t \frac{P_{\hat{x}}}{P} + k(m^* - m_t)\frac{P_m}{P} + \frac{\Sigma_x \Lambda_{1:4,1:4}\Sigma_x{}'}{2}\frac{P_{\hat{x}\hat{x}}}{P} + \frac{\sigma_m^2}{2}\frac{P_{mm}}{P} + \sigma_m \Sigma_x \Sigma_{ym}\frac{P_{\hat{x}m}}{P} \tag{73}$$

$$+ \frac{P_{\hat{x}}}{P}\sigma_{\Pi^n}^T \Lambda_{1:4,1:4}\Sigma_x' + \frac{P_m}{P}\sigma_m \sigma_{\Pi^n}' \Sigma_{ym} - r_{f,t}^n = 0\,.$$

where $\Sigma_{ym} = \begin{bmatrix} \rho_{im} \\ \rho_{cm} \\ \rho_{dm} \\ 1 \end{bmatrix}$.

The above PDE satisfies Feynman-Kac conditions, and therefore allows us to write the solution as follows:

$$P = exp\{A(t,s) + B(t,s)\hat{x} + C(t,s)m\}. \tag{74}$$

Inserting (74) into the PDE (73), together with the fact that PDE (73) should hold for arbitrary values of $\hat{x}$ and m, we get the following ordinary differential equations for functions $A(t,s)$, $B(t,s)$, and $C(t,s)$:

$$B_t - \rho B - 1 = 0,$$

$$C_t - kC - (Q_c + 1) = 0,$$

$$A_t + Ckm^* + \frac{1}{2}\left[B^2 \Sigma_x \Lambda_{1:4,1:4}\Sigma_x' + 2BC\sigma_m \Sigma_x \Sigma_{ym} + C^2\sigma_m^2\right]$$

$$+ \Omega^n + B\sigma_{\Pi^n}^T \Lambda_{1:4,1:4}\Sigma_x' + C\sigma_m \sigma_{\Pi^n}' \Sigma_{ym} = 0$$

with initial conditions $A(s,s) = B(s,s) = C(s,s) = 0$.

Subject to these initial conditions, the solution of the ODE system is:

$$B(t,s) = -\frac{1}{\rho}(1 - exp\{-\rho(s-t)\}),$$

$$C(t,s) = -\frac{(Q_c+1)}{k}(1 - exp\{-k(s-t)\}),$$

$$A = c_0(s-t) - \frac{c_1}{\rho}(exp\{\rho(t-s)\} - 1) - \frac{c_2}{k}(exp\{k(t-s)\} - 1)$$
$$- \frac{c_4}{2k}(exp\{2k(t-s)\} - 1) - \frac{c_3}{2\rho}(exp\{2\rho(t-s)\} - 1)$$
$$- \frac{c_5}{k+\rho}(exp\{(k+\rho)(t-s)\} - 1).$$

where

$$c_1 = \frac{1}{k\rho^2}(-2kz_1 - (Q_c+1)z_2\rho + k\rho\lambda_1),$$

$$c_2 = \frac{1}{k^2\rho}(Q_c+1)(-\sigma_m^2\rho - kz_2 - Q_c\sigma_m^2\rho + k\rho\lambda_2 + k^2\rho m^*),$$

$$c_3 = \frac{z_1}{\rho^2},$$

$$c_4 = \frac{\sigma_m^2}{2k^2}(Q_c+1)^2,$$

$$c_5 = \frac{z_2(Q_c+1)}{k\rho},$$

$$c_0 = \frac{1}{k^2\rho^2}\Big(-k^2\rho\lambda_1 - (Q_c+1)k\rho^2(km^* + \lambda_2) + k^2 z_1 + \frac{1}{2}(Q_c+1)^2\rho^2\sigma_m^2 + (Q_c$$
$$+ 1)k\rho z_2\Big) - \Omega^n,$$

$$z_1 = \frac{1}{2}\Sigma_x\Lambda_{1:4,1:4}\Sigma_x{}',$$

$$z_2 = \sigma_m\Sigma_x\Sigma_{ym},$$

$$\lambda_1 = \sigma_{\Pi^n}'\Lambda_{1:4,1:4}\Sigma_x{}',$$

$$\lambda_2 = \sigma_m\sigma_{\Pi^n}'\Sigma_{ym},$$

$$\Sigma_{ym} = \begin{bmatrix} \rho_{im} \\ \rho_{cm} \\ \rho_{dm} \\ 1 \end{bmatrix}.$$

Equation (72) implies that the instantaneous term premium is:

$$\lambda = -B(t,s)\sigma_{\Pi^n}^T \Lambda_{1:4,1:4}\Sigma_x' - C(t,s)\sigma_m \sigma_{\Pi^n}' \Sigma_{ym},$$

which has two components: $\lambda_{\hat{x}} = -B(t,s)\sigma_{\Pi^n}^T \Lambda_{1:4,1:4}\Sigma_x'$ is the premium due to shocks to the conditional consumption mean, and $\lambda_m = -C(t,s)\sigma_m \sigma_{\Pi^n}' \Sigma_{ym}$ is the premium due to shocks to the mean inflation.

A3.4 Derivation of the Zero Coupon Bond Price in Complete Information Economy.

The processes of state variables in complete information economy are given by equations (1) – (5). Comparing processes (1) – (5) with (6) – (11), we can see that drifts in two systems have the exact same structure. Equation (54) suggests that only the drift structure matters when solving for state variables' coefficients in the value function. Therefore, the value function J^c should have a similar functional form as J given by equations (17) – (21):

$$J^c(t, \ln(i), g, d, x, m) = \exp\{\xi_{0t} + \xi_{1t}\ln(i) + \xi_{2t}g_c + \xi_{3t}g_d + \xi_{4t}m + \xi_{5t}x\}, \tag{75}$$

where

$$\xi_1 = \xi_3 = 0,$$

$$\xi_2 = 1 - \gamma,$$

$$\xi_4 = \frac{(1-\gamma)Q_c}{k+h},$$

$$\xi_5 = \frac{1-\gamma}{h+\rho}.$$

Given (75) and state variable processes (1) – (5), the process of consumption C and the full differential of J^c are given by:

$$\frac{dC}{C} = (\mu_c + Q_c m_t + x_t)dt + \sqrt{V} dw_c, \tag{76}$$

$$dJ^c = -f dt + J^c_{\ln(i)} \sigma_i dw_i + J^c_C C \sqrt{V} dw_c + J^c_d \varphi_d \sqrt{V} dw_d + J^c_m \sigma_m dw_m \tag{77}$$

$$+ J^c_x \varphi_e \sqrt{V} dw_x,$$

$$\frac{J^c_{\ln(i)}}{J} = \frac{J^c_d}{J} = 0, \frac{J^c_C}{J} = \frac{\xi_2}{C}, \frac{J^c_m}{J} = \xi_4, \frac{J^c_x}{J} = \xi_5. \tag{78}$$

Combining (76) – (78) with (61), we obtain the process of the real SDF:

$$\frac{d\Pi^c}{\Pi^c} = -[\Omega^c + x_t + Q_c m]dt + \sigma_\Pi^{cT} \begin{bmatrix} dw_i \\ dw_c \\ dw_d \\ dw_m \\ dw_x \end{bmatrix},$$

where

$$\Omega^c = h + \mu_c + (\xi_2 - 1)V + +\xi_4 \sigma_m \sqrt{V} \rho_{cm} + \xi_5 \varphi_e \sqrt{V} \rho_{cx},$$

$$\sigma_\Pi^c = \begin{bmatrix} 0 \\ (\xi_2 - 1)\sqrt{V} \\ 0 \\ \xi_4 \sigma_m \\ \xi_5 \varphi_e \sqrt{V} \end{bmatrix}$$

Further, the processes for the nominal pricing kernel and the short rate are given by

$$\frac{d\Pi^{cn}}{\Pi^{cn}} = \frac{d(\Pi^c/i)}{\Pi^c/i} \tag{79}$$

$$= \left[-(\Omega^c + x_t + Q_c m) - m + \sigma_i^2 - \sigma_i \begin{bmatrix} 1 \\ \rho_{ic} \\ \rho_{id} \\ \rho_{im} \\ \rho_{ix} \end{bmatrix}^T \sigma_\Pi^c \right] dt$$

$$+ \begin{bmatrix} -\sigma_i \\ (\xi_2 - 1)\sqrt{V} \\ 0 \\ \xi_4 \sigma_m \\ \xi_5 \varphi_e \sqrt{V} \end{bmatrix}^T \begin{bmatrix} dw_i \\ dw_c \\ dw_d \\ dw_m \\ dw_x \end{bmatrix}$$

$$= -[\Omega^{cn} + x_t + (Q_c + 1)m]dt + \sigma_{\Pi^n}^{cT} \begin{bmatrix} dw_i \\ dw_c \\ dw_d \\ dw_m \\ dw_x \end{bmatrix}.$$

$$r_f^{cn} = \Omega^{cn} + x_t + (Q_c + 1)m_t \tag{80}$$

The zero-coupon bond price $P^c(t,s)$ should satisfy the Euler equation:

$$E\left(\frac{dP^c}{P^c}\right) = r_f^{cn} dt - E\left(\frac{dP^c}{P^c}\frac{d\Pi^{cn}}{\Pi^{cn}}\right).$$

As in A3, it is easy to show that P^c is a function of only x_t, m_t, and t. Applying Itô's lemma to $P^c(x,m,t)$, we have:

$$\frac{dP^c}{P^c} = \frac{1}{P^c}\left(P_t^c + P_x^c dx + P_m^c dm + \frac{1}{2}(P_{xx}^c dx^2 + 2P_{xm}^c dxdm + P_{mm}^c dm^2)\right), \tag{81}$$

$$\frac{dP^c}{P^c}\frac{d\Pi^{cn}}{\Pi^{cn}} = \frac{P_x^c}{P^c}\varphi_e\sqrt{V}\sigma_{\Pi^n}^{cT}\begin{bmatrix} \rho_{ix} \\ \rho_{cx} \\ \rho_{dx} \\ \rho_{mx} \\ 1 \end{bmatrix} + \frac{P_m^c}{P^c}\sigma_m\sigma_{\Pi^n}^{cT}\begin{bmatrix} \rho_{im} \\ \rho_{cm} \\ \rho_{dm} \\ 1 \\ \rho_{mx} \end{bmatrix}. \tag{82}$$

Collecting all the terms in (81) and (82), taking expectations, and dividing throughout by *dt*, we arrive at the PDE for the bond price:

$$\frac{P_t^c}{P^c} - \rho x_t \frac{P_x^c}{P^c} + k(m^* - m_t)\frac{P_m^c}{P^c} + \frac{\varphi_e^2 V}{2}\frac{P_{xx}^c}{P^c} + \frac{\sigma_m^2}{2}\frac{P_{mm}^c}{P^c} + \sigma_m \varphi_e \sqrt{V} \rho_{mx} \frac{P_{xm}^c}{P^c} \tag{83}$$

$$+ \frac{P_x^c}{P^c} \varphi_e \sqrt{V} \sigma_{\Pi^n}^{cT} \begin{bmatrix} \rho_{ix} \\ \rho_{cx} \\ \rho_{dx} \\ \rho_{mx} \\ 1 \end{bmatrix} + \frac{P_m^c}{P^c} \sigma_m \sigma_{\Pi^n}^{cT} \begin{bmatrix} \rho_{im} \\ \rho_{cm} \\ \rho_{dm} \\ 1 \\ \rho_{mx} \end{bmatrix} - r_{f,t}^{cn} = 0 \,.$$

PDE (83) satisfies Feynman-Kac conditions, so we can write the solution as follows:

$$P^c = exp\{A^c(t,s) + B^c(t,s)x + C^c(t,s)m\}. \tag{84}$$

Inserting (84) into the PDE (83), we get the following ordinary differential equations for functions $A^c(t,s)$, $B^c(t,s)$, and $C^c(t,s)$:

$$B_t^c - \rho B^c - 1 = 0,$$

$$C_t^c - kC^c - (Q_c + 1) = 0,$$

$$A_t^c + C^c k m^* + \frac{1}{2}\left[B^{c2}\varphi_e^2 V + 2B^c C^c \sigma_m \varphi_e \sqrt{V} \rho_{mx} + C^2 \sigma_m^2\right]$$

$$+ \Omega^{cn} + B^c \varphi_e \sqrt{V} \sigma_{\Pi^n}^{cT} \begin{bmatrix} \rho_{ix} \\ \rho_{cx} \\ \rho_{dx} \\ \rho_{mx} \\ 1 \end{bmatrix} + C^c \sigma_m \sigma_{\Pi^n}^{cT} \begin{bmatrix} \rho_{im} \\ \rho_{cm} \\ \rho_{dm} \\ 1 \\ \rho_{mx} \end{bmatrix} = 0$$

with initial conditions $A^c(s,s) = B^c(s,s) = C^c(s,s) = 0$.

Subject to these initial conditions, the solution of the ODE system is:

$$B^c(t,s) = -\frac{1}{\rho}(1 - exp\{-\rho(s-t)\}), \tag{85}$$

$$C^c(t,s) = -\frac{(Q_c + 1)}{k}(1 - exp\{-k(s-t)\}), \tag{86}$$

$$A^c = c_0^c(s-t) - \frac{c_1^c}{\rho}(exp\{\rho(t-s)\}-1) - \frac{c_2^c}{k}(exp\{k(t-s)\}-1) \tag{87}$$

$$-\frac{c_4^c}{2k}(exp\{2k(t-s)\}-1) - \frac{c_3^c}{2\rho}(exp\{2\rho(t-s)\}-1)$$

$$-\frac{c_5^c}{k+\rho}(exp\{(k+\rho)(t-s)\}-1).$$

where

$$c_1^c = \frac{1}{k\rho^2}\left(-k\varphi_e^2 V - (Q_c+1)\sigma_m\varphi_e\sqrt{V}\rho_{mx}\rho + k\rho\lambda_1^c\right),$$

$$c_2^c = \frac{1}{k^2\rho}(Q_c+1)\left(-\sigma_m^2\rho - k\sigma_m\varphi_e\sqrt{V}\rho_{mx} - Q_c\sigma_m^2\rho + k\rho\lambda_2^c + k^2\rho m^*\right),$$

$$c_3^c = \frac{\varphi_e^2 V}{2\rho^2},$$

$$c_4^c = \frac{\sigma_m^2}{2k^2}(Q_c+1)^2,$$

$$c_5^c = \frac{\sigma_m\varphi_e\sqrt{V}\rho_{mx}(Q_c+1)}{k\rho},$$

$$c_0^c = \frac{1}{k^2\rho^2}\left(-k^2\rho\lambda_1^c - (Q_c+1)k\rho^2(km^* + \lambda_2^c) + \frac{\varphi_e^2 V k^2}{2} + \frac{1}{2}(Q_c+1)^2\rho^2\sigma_m^2 + (Q_c\right.$$

$$\left. + 1)k\rho\sigma_m\varphi_e\sqrt{V}\rho_{mx}\right) - \Omega^{cn},$$

$$\lambda_1^c = \varphi_e\sqrt{V}\sigma_{\Pi^n}^{cT}\begin{bmatrix}\rho_{ix}\\ \rho_{cx}\\ \rho_{dx}\\ \rho_{mx}\\ 1\end{bmatrix},$$

$$\lambda_2^c = \sigma_m\sigma_{\Pi^n}^{cT}\begin{bmatrix}\rho_{im}\\ \rho_{cm}\\ \rho_{dm}\\ 1\\ \rho_{mx}\end{bmatrix}.$$

The instantaneous risk premium is:

$$\lambda = -B^c \varphi_e \sqrt{V} \sigma_{\Pi^n}^{cT} \begin{bmatrix} \rho_{ix} \\ \rho_{cx} \\ \rho_{dx} \\ \rho_{mx} \\ 1 \end{bmatrix} - C^c \sigma_m \sigma_{\Pi^n}^{cT} \begin{bmatrix} \rho_{im} \\ \rho_{cm} \\ \rho_{dm} \\ 1 \\ \rho_{mx} \end{bmatrix}. \tag{88}$$

Given the price function, the yield, $y_{t,s}^c$, is simply

$$y_{t,s}^c = \frac{1}{t-s}(A^c + B^c x_t + C^c m_t), \tag{89}$$

with a conditional variance of

$$var(y_{t,s}^c | x_{t-1}, m_{t-1}) = \frac{B^{c2} \varphi_e^2 V + 2B^c C^c \sigma_m \varphi_e \sqrt{V} \rho_{mx} + C^2 \sigma_m^2}{(t-s)^2}. \tag{90}$$

A3.5 Simulation Study.

In this section, we use a simulation study to show the effectiveness of our sampler. We consider a state space model with the same structure as model (1) – (6). In order to make the model identifiable, we also adopt the DFM2 restrictions (44) suggested by Bai and Wang (2015).

Our choice of the 'true' values of drift parameters

$\delta = \left[{\mu_y^*}'_{3\times1}\ \lambda_5\ \lambda_6\ Q_c, {Q_d}^*, \phi^*, k^*, \rho^*\right]'$ is as follows:

$$\delta = [0.01, 0,0, 0.5, 1, 1, 1.4, 1, 0.5, 0.9]', \tag{91}$$

which implies a factor loading matrix Λ^* of

$$\Lambda^* = \begin{bmatrix} \mu_y^* \vdots 0 \\ \Lambda_0' \vdots 0 \\ \Lambda_1' \vdots G \end{bmatrix} = \begin{bmatrix} 0.01 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0.5 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0.5 & 1 & 1.4 & 0.5 & 0 \\ 0 & 0.1 & 1 & 0 & 0.9 \end{bmatrix}.$$

'True' diffusion parameters R and Q are assumed to be proportional to identity matrix:

$$R = 0.01 * I_3, \tag{92}$$

$$Q = 0.01 * I_2. \tag{93}$$

We simulate the model 50 times. For each simulation, we estimate model parameters and state variables using our sampler. As prior distributions of δ, R, and Q we choose:

$$\delta \sim N(0_{10\times 1}, 5 * I_5),$$

$$R \sim W^{-1}(0.05 * I_3, 5),$$

$$Q \sim W^{-1}(0.03 * I_2, 4).$$

We run the Gibbs samplers for 50000 iterations and keep the last 10000 draws. Therefore, in total we have 10000×50 draws for each model parameters across all 50 simulations. With these samples, we are able to examine the empirical distribution of each parameter.

Table A3.5.1 reports the 95% bands, 50% bands, median, mean, and standard deviation of the parameter draws. Although for some parameters (such as $\mu^*_{y_2}$, $\mu^*_{y_3}$, Q_c, Q^*_d, and $Q_{2,2}$), estimated median is not right on the true value, for all 19 parameters the true value is well within the 50% bands of the sampling results. Noting the randomness in simulation, we may consider the simulation results as a piece of supportive evidence to the effectiveness of our sampler.

Table A3.5.1 Estimated Densities of Model Parameters.
This table highlights features of the estimated densities of the nineteen model parameters conditioned on the 200 simulated monthly observations of three state variables with true parameter values given by equations (91) – (93). The model is as defined in equation (42) – (43). $R_{i,j}$ $\left(Q_{i,j}\right)$ is element in ith row and jth column of variance matrix R (Q). Std. error is the sample standard deviation.

Parameter	2.5%ile	25%ile	Median	Mean	75%ile	97.5%ile	Std. error
$\mu^*_{y_1}$	-0.0589	-0.0175	0.0046	0.0035	0.0261	0.0723	0.0332
$\mu^*_{y_2}$	-0.5738	-0.0997	-0.0389	-0.0154	0.0625	0.2670	0.1964
$\mu^*_{y_3}$	-1.0390	-0.1771	-0.0663	-0.0238	0.1162	0.4729	0.3466
λ_5	0.0370	0.3579	0.5331	0.5314	0.6978	1.0688	0.2610
λ_6	0.3826	0.8640	1.0684	1.0965	1.2985	1.6199	0.3234
Q_c	0.4046	0.7067	0.8703	0.8624	1.0238	1.3870	0.2464
Q^*_d	0.4833	0.9905	1.2462	1.2416	1.4979	2.0287	0.3900
ϕ^*	0.4275	0.7447	0.9702	0.9456	1.1677	1.6487	0.3181
k^*	0.2062	0.4021	0.4920	0.5009	0.5911	0.7289	0.1359
ρ^*	0.6222	0.8238	0.8640	0.8778	0.9241	0.9891	0.0892
$R_{1,1}$	0.0063	0.0086	0.0102	0.0100	0.0116	0.0155	0.0023
$R_{1,2}$	-0.0038	-0.0012	0.0001	0.0002	0.0014	0.0039	0.0020
$R_{1,3}$	-0.0040	-0.0013	0.0001	0.0001	0.0015	0.0042	0.0021
$R_{2,2}$	0.0063	0.0093	0.0116	0.0113	0.0135	0.0184	0.0031
$R_{2,3}$	-0.0035	-0.0010	0.0010	0.0006	0.0026	0.0078	0.0029
$R_{3,3}$	0.0055	0.0087	0.0115	0.0110	0.0137	0.0204	0.0039
$Q_{1,1}$	0.0062	0.0091	0.0111	0.0109	0.0128	0.0174	0.0028
$Q_{1,2}$	-0.0052	-0.0013	0.0003	0.0004	0.0019	0.0053	0.0026
$Q_{2,2}$	0.0067	0.0098	0.0130	0.0123	0.0149	0.0231	0.0045

A3.6 Convergence Diagnostics.

In this part we report the Gelman-Rubin statistic (Gelman and Rubin (1992)) for each parameter estimate. The Gelman-Rubin statistic shows the ratio between the total variance (sum of within chain variance and between chain variance) and the within chain variance and can be computed easily from multiple chains of parameter draws. For example, for parameter θ, we sample m chains with n iterations in each chain. The within chain variance W is: $W = \frac{1}{m}\sum_{j=1}^{m} s_j^2$, where $s_j^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(\theta_{ij} - \bar{\theta}_j\right)^2$ and $\bar{\theta}_j$ is the mean of parameter draws in chain j. The between chain variance is $B =$

$\frac{n}{m-1}\sum_{j=1}^{m}\left(\bar{\theta}_j - \bar{\bar{\theta}}\right)^2$, where $\bar{\bar{\theta}} = \frac{1}{m}\sum_{j=1}^{m}\bar{\theta}_j$. With both within and between chain variances at hand, the Gelman-Rubin statistic is

$$R = \sqrt{\left(1-\frac{1}{n}\right)+\frac{B}{nW}}.$$

The Gelman-Rubin statistic close to 1 is normally taken as a signal that a sampler has a good convergence to the stationary distribution. Table A3.6.1 lists Gelman-Rubin statistics for each of the 19 model parameters.

Table A3.6.1 Gelman-Rubin Statistic.
This table reports the Gelman-Rubin Statistic for each parameter. For each parameter, we run three chains starting from 3 different sets of initial values. For each chain, we run 100,000 iterations. We took 50,000 burn-in iterations for each chain and compute Gelman-Rubin Statistic using the rest $50{,}000 \times 3 = 150{,}000$ post burn-in iterations. $R_{i,j}$ $(Q_{i,j})$ is element in ith row and jth column of variance matrix R (Q).

$\mu^*_{y_1}$	$\mu^*_{y_2}$	$\mu^*_{y_3}$	λ_5	λ_6	Q_c	Q^*_d	ϕ^*	k^*	ρ^*
1.0001	1.0173	1.0224	1.0004	1.0010	1.0100	1.0099	1.0058	1.0007	1.0091
$R_{1,1}$	$R_{1,2}$	$R_{1,3}$	$R_{2,2}$	$R_{2,3}$	$R_{3,3}$	$Q_{1,1}$	$Q_{1,2}$	$Q_{2,2}$	
1.0001	1.0001	1.0000	1.0045	1.0002	1.0000	1.0003	1.0079	1.0002	

A3.7 Model Assessment.

In this section, we look at the goodness of fit of the model (1) – (6).

Figures A3.7.1, A3.7.2, and A3.7.3 show the time series of state variables, $\Delta\pi$, Δg_c, and Δg_d, along with the corresponding 95th percentile bands of the posterior predictive densities (Rubin (1984), Aitkin (1991)). The posterior predictive density, $f(X|X_{obs})$, is the density of a new set of state variables conditioning on the real ones, which is given by:

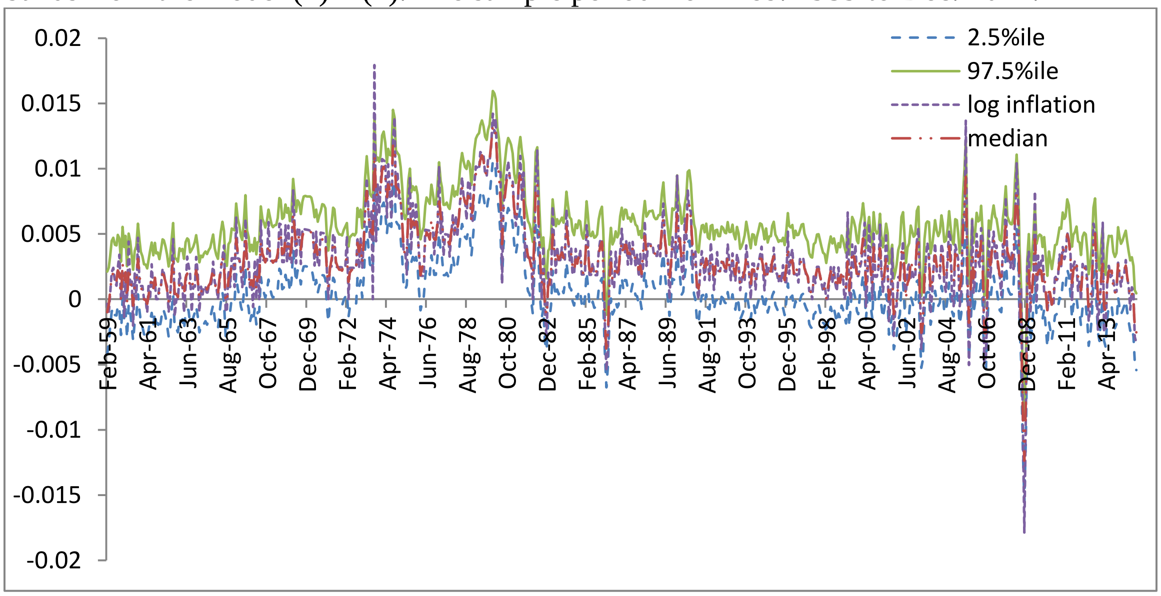
$$f(X|X_{obs}) = \int f(X|\delta) f(\delta|X_{obs})\, d\delta,$$

where δ is the model parameters and latent variables. In a MCMC simulation, $f(X|X_{obs})$, can be estimated from

$$\hat{f}(X|X_{obs}) = \frac{1}{m}\sum_{j=1}^{m} f\left(X|\delta_j^*\right),$$

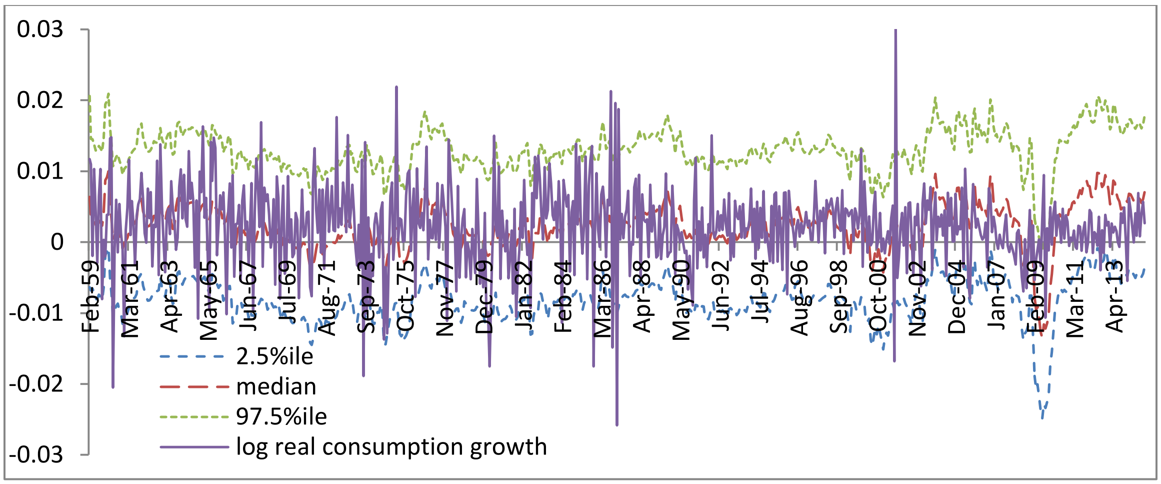
where $\{\delta_j^*; j = 1, \dots, m\}$ is the sample of parameter draws (Gelfand (1996)). When sampling from the posterior predictive density, Gelfand (1996) suggests that draws from $\left(X|\delta_j^*\right)$, X_j^*, are marginally samples from $f(X|X_{obs})$. In our case, since $X_t, t = 1, \dots, T$ are independent given model parameters and latent variables, we are allowed to draw X_t directly from $f\left(X_t|\delta_j^*\right)$ rather than draw the whole time series X simultaneously.

Figure A3.7.1 Time series of log inflation $\Delta\pi$.
Time series of log monthly inflation $\Delta\pi$ with ninety-fifth percentile predictive density bands from the model (1) – (2). The sample period from Feb. 1959 to Dec. 2014.



In general, the model captures the log inflation (Figure A3.7.1) and the log real dividend growth (Figure A3.7.3) very well. There are only 11 log inflation and 9 dividend growth observations outside our 95% percentile bands of the posterior predictive densities. For log real consumption growth (Figure A3.7.2), the number of observations outside the bands is 45. This is 6.7% (45/671) of the whole sample, which is 1.7% higher than the conventional tolerance level. Among these 45 observations, 26 observations are outliers located outside the 95% bands of the log consumption growth ([-1.01%, 1.40%]). The remaining 19 observations are mainly (12 out 19) from the first half of the sample, with the other 6 observations from the last recession period. In the recession period from 2008, we observe a pronounced V-shape pattern in the log real dividend growth. It's apparent that the model is forced to fit the abnormal pattern in the dividend data in this period, which leads to its failure in capturing the consumption data.

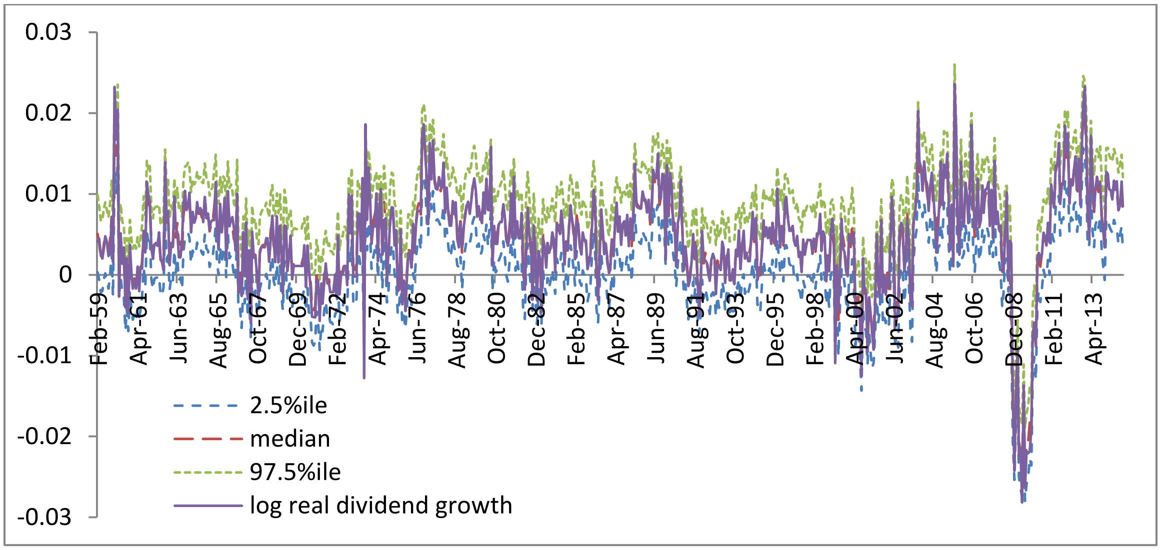
Figure A3.7.2 Time series of log real consumption growth Δg_c.
Time series of log real monthly consumption growth Δg_c with ninety-fifth percentile predictive density bands from the model (1) – (2). The sample period from Feb. 1959 to Dec. 2014.



In the pre-2008 consumption growth data, a quick F-test[37] suggests that there is a structural break in consumption growth variance in the early 90s. The change in consumption growth variance may help explain the failure of the model. Since the variance is assumed to be constant over the whole sample period, the model under-estimates the consumption growth variance in the first half of the sample, which leads to narrow predictive density bands before the 90s.

[37] The sample log consumption growth variance before 1990 is 4.50 e-5, and that from 1990 to 2008 is 1.95 e-5, which implies an F-statistic = 2.3129 with p-value = 1.

Figure A3.7.3 Time series of log real dividend growth Δg_d.
Time series of log real monthly dividend growth Δg_d with ninety-fifth percentile predictive density bands from the model (1) – (2). The sample period from Feb. 1959 to Dec. 2014.



In Figure A3.7.4 (Figure A3.7.5), we plot the median of the LRR component (mean inflation) with log real consumption and dividend growth (log inflation). Figure A3.7.4 shows that the LRR component captures fluctuations in the log dividend growth more than those in the log consumption growth. This is confirmed by the unconditional correlations of the LRR component with the log consumption and dividend growth. While the LRR component is weakly unconditionally correlated with the log consumption growth (0.0512), its unconditional correlation with the log dividend growth has the median of 0.8210. In Figure A3.7.5, the median of mean inflation matches the log inflation almost perfectly, with the unconditional correlation between these two variables of 0.8991.

Figure A3.7.4 Time series of log real consumption and dividend growth, and median of the LRR component.

The figure presents the time series of log real monthly consumption and dividend growth with median of the LRR component. The sample period is from Feb. 1959 to Dec. 2014.

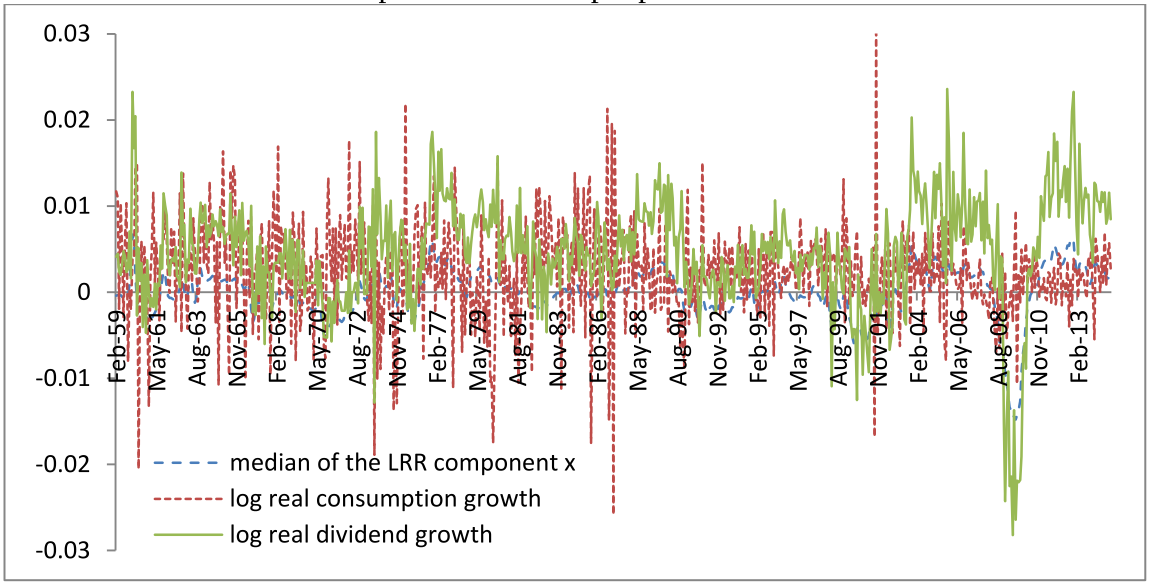
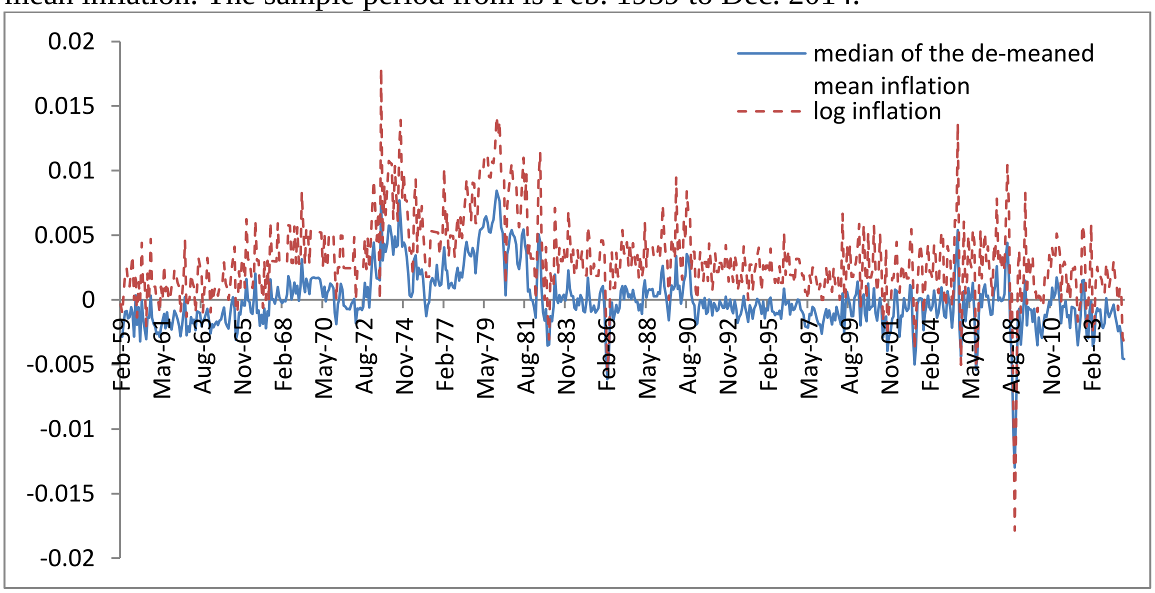


Figure A3.7.5 Time series of log inflation and median of the demeaned mean inflation

The figure presents the time series of log monthly inflation and median of the demeaned mean inflation. The sample period from is Feb. 1959 to Dec. 2014.



In Figure A3.7.6, we report the log conditional predictive ordinate (CPO), d_t, for state variables at each time period t, where

$$d_t = f(X_{t,obs}|X_{(t),obs}),$$

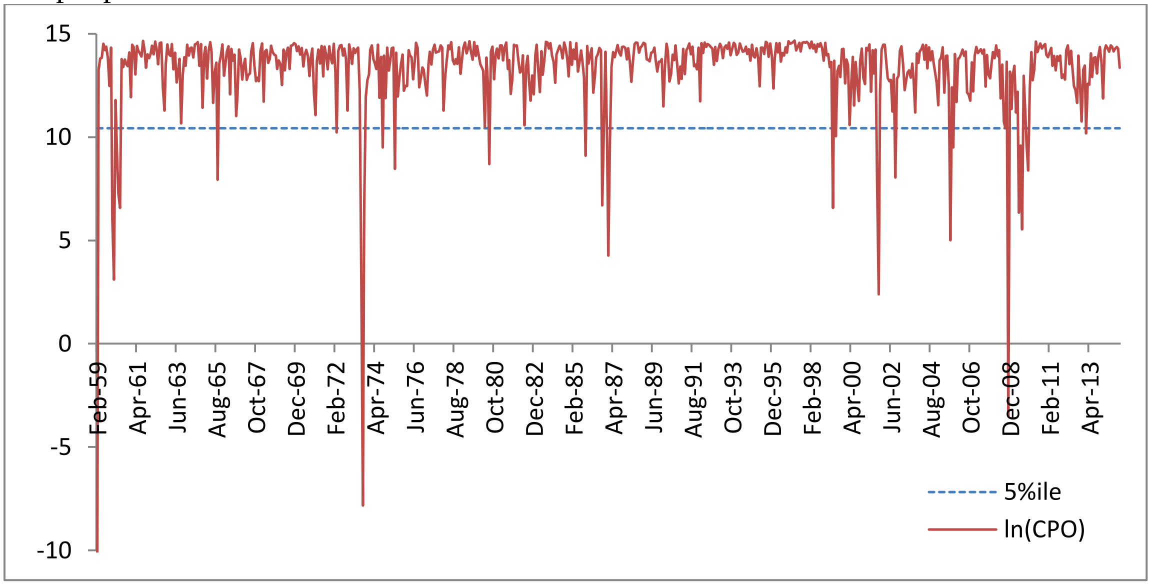
and $X_{(t),obs}$ is X_{obs} without the tth element $X_{t,obs}$.

$f(X_t|X_{(t),obs})$ is the cross-validation predictive density of X_t, which suggests whether X_t is likely given all X_{obs} except $X_{t,obs}$ itself. The estimate of $f(X_t|X_{(t),obs})$ is

$$\hat{f}(X_t|X_{(t),obs}) = \frac{1}{\frac{1}{m}\sum_{j=1}^{m} f(X_t|X_{(t),obs}, \delta_j^*)}.$$

Relation $f(X_t|X_{(t),obs}, \delta_j^*) = f(X_t|\delta_j^*)$ holds when $\{X_t\}$ are independent given parameters and latent variables (Gelfand (1996)).

Figure A3.7.6 Log conditional predictive ordinate (CPO).
Log conditional predictive ordinate (CPO) for vector of state variable, $y_t = [\Delta\pi_t \quad \Delta g_c \quad \Delta g_d]'$ at each month, alone with 5%ile band of log(CPO). $\Delta\pi_t$ is log inflation, Δg_c is log real consumption growth, and Δg_d is log real dividend growth. The sample period is from Feb. 1959 to Dec. 2014.



Comparing figure A3.7.6 with state variable plots in Figure 3.1A, we can see that surprises mainly come from unexpected moves in the log real dividend growth at the end of the sample. Most significantly, the drastic fall in the log dividend growth in late 2008 and early 2009 is matched by six surprising observations with low log CPO. In the other part of the sample, consumption growth plays a more important role in producing surprising observations. For example, four outlier observations in early 1987 may be due to the large fluctuation in the consumption growth around that period.

Overall, the model is able to capture the dynamics of both log inflation and log real dividend growth. However, it has a relatively inferior performance in explaining the log real consumption growth. The model failure is possibly due to a structural break and abnormally large short-time fluctuations in the data. Potential remedies include incorporating stochastic variance and fast-moving components (jumps) into state variable dynamics.

Chapter Four: Unconditional Tests of Linear Asset pricing Models with Time-Varying Betas

4.1 Introduction

Previous studies show that, although unconditional factor models have generally failed in explaining the cross section of average stock returns, dynamic asset pricing models may succeed in the task. By allowing factor loadings in a stochastic discount factor (SDF) or factor betas to vary over time, conditional models (such as Jagannathan and Wang (1996)'s (hereafter JW) conditional CAPM, Lettau and Lidvigson (2001b)'s (hereafter LL01) scaled C-CAPM, and Santos and Veronesi (2006)'s (hereafter SV) labor income ratio model) are able to explain a large proportion of variation in average stock returns. These models suggest that identifying proper instruments to describe time variation in an investment opportunity set in conditional linear factor models is a fruitful avenue for future research in asset pricing. However, Lewellen and Nagel (2006) (hereafter LN) warn that the success of the conditional models may be illusory and should be treated with caution. LN argue that conditional models mentioned above may not be able to produce statistically small unconditional alphas. Their analysis implies that good performance of those conditional models in cross-sectional tests may be due to the fact that some important unconditional constraints are omitted in the estimation procedure.

LN start with beta representation $E(R_{t+1}|\mathcal{F}_t) = \beta_t \lambda_t$[38]. If we assume that the conditional factor beta β_t is an affine function of a (zero-mean) instrument z_t (i.e.,

[38] For simplicity, we use a one-factor model here as an example.

$\beta_t = \beta_0 + \delta_z z_t$, the 'affine beta' constraint), we can derive unconditional moments by taking unconditional expectation on both sides of the conditional model:

$$E(R_{t+1}) = \beta_0 E(\lambda_t) + \delta_z cov(z_t \lambda_t).$$

Based on the unconditional moment above, LN argue that the slope on β_0 should equal $E(\lambda_t)$ and slopes on δ_z should equal $cov(z_t \lambda_t)$in cross-sectional regressions. LN claim that neglecting these constraints on unconditional risk prices in model estimation may be responsible for the surprisingly good reported performance of conditional affine models. Ludvigson (2011) takes a closer look at LN's critique. She shows that, LN's critique cannot be applied to CCAPM.[39] LN's critique is a direct consequence of writing time-varying betas as a function of instruments and unconditional betas. Such relation does not exist in the conditional CCAPM. More specifically, in the conditional CCAPM, pricing kernel's loadings on factors, rather than factor betas, are time-varying functions of instruments. Therefore LN's critique is not valid for the model. We return to the technical details in the later part of the paper.

While Ludvigson (2011) shows that CCAPM is immune to LN's critique, it is still not clear if models with time-varying beta assumptions are subject to LN's critique. In this Chapter, we follow LN's work and examine the unconditional moment implications of the conditional factor models where factor betas are assumed to be time-varying or, more specifically, affine functions of instruments. Note that the last assumption is not implied in any way by linear SDF models. Instead, it must be considered as an additional constraint on the model. Theoretically, the assumption that betas are affine functions of

[39] We conduct a simple empirical test and confirm her results in Appendix A4.4.

instruments is likely invalid. Many linear asset pricing models, including the one we study here, assume an SDF affine in factors. However, the 'affine beta' constraint does not necessarily correctly map back into the SDF affine in factors. Nevertheless, even if the affine beta assumption is not valid theoretically, it may still be a valid empirical approximation. In this paper, we examine the empirical validity of this assumption. Once the conditional model is conditioned down to its unconditional version, the affine restriction on conditional factor betas leads to unconditional constraints on unconditional betas and risk prices (LN's critique). This feature allows us to assess the empirical validity of the affine beta assumption by comparing estimation results with or without imposing these unconditional constraints. To do so, we develop an empirical testing procedure which allows us to incorporate the constraints on factor betas and risk prices.

To impose the 'affine beta' constraint in empirical tests, we develop a three-stage regression procedure. We first estimate time-varying betas using rolling-window regressions. In the second-stage, we run a regression of the estimated time-varying betas on instruments to retrieve unconditional betas. The last step is to run the cross-sectional regression to estimate factor risk premiums. Our methodology produces the estimates of unconditional betas from estimated time series of time-varying betas based on the functional relationship between them. Our empirical results of the time-varying beta model (SV's labor income ratio model is used as an example in our tests) show that imposing the constraint changes estimates of unconditional betas and those of factor risk premiums significantly and has a significant negative impact on the model performance measured by the root mean squared error (RMSE). As an alternative estimation method, we could use the Kalman filter to estimate the time-varying betas. However, this requires us to model the processes of the time-varying betas and instruments, which is usually not assumed in the conditional models (such as that of SV). To be true to the original conditional model, we use the rolling-window regressions. In the same spirit, we take the time-varying assumption on factor betas as given. In this chapter, we focus on the estimation procedure of the conditional models developed for the time-varying beta models. While testing the time-varying beta assumption is of special interest, it is not the focus of this chapter.

In this chapter we extend the work of LN by identifying and testing unconditional moment implications of conditional models with time-varying betas. Our work is related to Nagel and Singleton (2011), who expand the set of conditional restrictions and devise an optimal GMM procedure to test this expanded set. Our work is also related to Miao and Santa-Clara (2012) where ICAPM-based theoretical restrictions on multifactor models are examined. While Nagel and Singleton (2011) and Miao and Santa-Clara (2012) focus on the conditional constraints on factor models, we test the unconditional constraints arising in the conditioning-down stage of testing conditional models.

This chapter is organized as follows. In Section 2, we briefly review related literature. In section 3, we present LN’s critique of cross-sectional tests of conditional models and our discussion of the critique. Section 4 describes our testing approach and presents main empirical results. We conduct several robustness checks in Section 5 and conclude in Section 6.

4.2 Related literature.

It is well known that the static CAPM or the consumption CAPM (CCAPM) cannot explain the cross-sectional variation of asset prices (e.g., Fama and French (1992, 1996, 2008, 2012), Campbell and Cochrane (1999), Acharya and Pedersen (2005), Lundblad (2007), Asness, Moskowitz, and Pedersen (2013)). However, since both models impose constraints on conditional (as opposed to unconditional) moments of returns, with conditioning on investor's information, it is possible that a conditional CAPM or CCAPM may succeed where unconditional models fail. Intuitively, by allowing risk prices and/or asset risk exposure to be time-varying, conditional models allow for extra flexibility in fitting the data. Earlier work on this topic includes Jensen (1968), Dybvig and Ross (1985), Ferson, Kandel, and Stambaugh (1987), Hansen and Richard (1987), Bollerslev, Engle, and Wooldridge (1988), and Harvey (1989) among others. More recently, studies such as Jagannathan and Wang (1996), Lettau and Ludvigson (2001b), Lustig and Van Nieuwerburgh (2005) and Santos and Veronesi (2006), examine the performance of the conditional CCAPM. Wang (2003), Petkova and Zhang (2005), Ang and Chen (2006), and Adrian and Franzoni (2009) among others study the conditional CAPM. While Cochrane (1996) does not adopt a CAPM framework, he incorporates the conditional SDF feature into a production-based pricing model. The working assumption in all of the above papers is that the SDF or factor betas are either time-varying or have a certain functional form.

One special case we consider is a model with the affine SDF loading or factor betas, in which SDF's factor loading or factor betas are affine functions of some state variables (instruments). The benefit of this kind of model is that the model remains affine even after being conditioned down. Among others, Jagannathan and Wang (1996), Cochrane (1996), Lettau and Ludvigson (2001b), Lustig and Van Nieuwerburgh (2005) and Santos and Veronesi (2006) study this special type of conditional model.

Jagannathan and Wang (1996) decompose the asset's conditional beta into the asset's unconditional beta, conditional beta's projection on the market risk premium, and a zero-mean residual. With this linear decomposition of conditional beta, the authors are able to condition down the conditional CAPM to a 2-factor model for unconditional expected returns. Since the human capital part in the wealth portfolio is not traded, Jagannathan and Wang also include the human capital return (approximated by labor income growth) in their model, which leads to a three-factor linear model. They test the unconditional version of model using excess returns on 100 size and book-to-market portfolios of NYSE and AMEX stocks from July 1963 to December 1990. Empirical results show that, together with human capital factor, their conditional model is able to explain more than 55% of the variation in portfolio returns in the cross-sectional regression and is not driven out by the Fama-French SMB and HML factors.

Cochrane (1996) studies a conditional model in which the pricing kernel is an affine function of investment returns. Although not directly observed, two investment return factors (non-residential and residential) are constructed from the production data. Cochrane examines the performance of both static and conditional version of the model. For the conditional model, Cochrane uses yield spread and market dividend-to-price ratio as instruments, which leads to a six-factor unconditional model. Estimated using GMM over the data sample of 10 size portfolios of NYSE stocks, the static model performs as well as the CAPM. As expected, the conditional versions of the model perform substantially better than the static one.

Lettau and Ludvigson (2001b) construct an instrument to capture information in total wealth. They assume that the log total wealth is approximately a weighted average of log financial wealth and log human capital. They further assume that the log human capital is an affine function of only labour income. With these two assumptions, together with the log-linear approximation of the consumption to wealth ratio, they show that the dynamics of consumption to wealth can be described by cay, the residual of regressing log consumption on log household asset holdings, log labor income, and their lags. With the new instrument, cay, the authors condition-down the conditional CCAPM to a three-factor model with consumption growth, cay, and their product. They test the model on the sample of 25 size and book-to-market portfolios using Fama-MacBeth regressions. The empirical results show that the conditional CCAPM with instrument cay performs about as well as the Fama-French 3-factor model in explaining the cross-sectional variation in average returns of 25 Fama-French size and book-to-market portfolios.

Santos and Veronesi (2006) propose that the labor income to consumption ratio may be a good instrument in the sense that it captures changes in the overall covariance between the consumption growth and financial asset returns. Intuitively, when labor income to consumption ratio is large, consumption is mainly funded by labor income and has smaller loading on cash flow from financial assets. Therefore, consumption growth is less correlated with return on financial assets. The authors assume that the asset excess return is determined by its beta with consumption and labor income growth, and factor betas are affine functions of the labor income to consumption ratio. Santos and Veronesi test the model using Fama-MacBeth regressions and excess returns on 25 Fama-French size and book-to-market portfolios. Although the pricing errors are significantly different from zero, the model is able to explain a large fraction (with an adjusted R^2 of 48%) of cross-sectional variation of test asset returns.

Lustig and Van Nieuwerburgh (2005) suggest that the housing collateral ratio, the ratio of housing wealth to human wealth, may help identify the conditional distribution of consumption beta in consumption CAPM. The authors argue that the shocks in consumption growth have a larger impact on the distribution of consumption growth across households with a low collateral ratio. In empirical tests on 25 size and book-to-market stock portfolios, the collateral ratio model outperforms the Fama-French 3-factor model. The value premium results from a larger correlation between returns on value stocks and consumption growth when the collateral ratio is low.

While studies cited above report an improvement in the model performance of conditional model relative to the static model, some recent work warns that one should be critical about the good fit of these models. E.g., Lewellen and Nagel (2006) suggest that the good performance of conditional models may be simply due to neglecting certain economic restrictions on regression intercepts and slopes. More specifically, the authors argue that the unconditional alpha produced by conditional CAPM is too large. To illustrate their concern, the authors run short-window CAPM regressions to retrieve estimates of conditional alphas and betas. They show that the unconditional alpha should be roughly in line with the covariance between conditional betas and the market risk premium. With their estimated parameter values, the conditional betas are just too stable to explain the observed unconditional alphas. Although the authors test conditional CAPM in this paper, they claim that other conditional models may also be subject to similar problems. For instance, when Lettau and Ludvigson (2001b) estimate their consumption CAPM model, they fail, according to Lewellen and Nagel (2006), to impose restrictions on return's loading on factor betas arising from the conditioning down procedure.

As an answer to Lewellen and Nagel (2006)'s critique, Ludvigson (2011) examines the conditioning-down and estimation procedure of Lettau and Ludvigson (2001b). She shows that Lewellen and Nagel's critique on the scaled consumption CAPM is not valid in the sense that they fail to separate the implications of the SDF model from those of the beta models. Assumptions on the distribution of conditional factor beta or the SDF's loading on factors make these two types of model distinctively different, and Lewellen and Nagel's concern can only be applied to conditional models in which betas are assumed to be time-varying.

Ang and Kristensen (2012) extend Lewellen and Nagel (2006)'s work by developing a nonparametric approach to estimate and test conditional alphas and betas and their long-run counterparts. They propose a weighted least mean square estimate for conditional alphas and betas where the weights on observations in different periods are given by a Gaussian density. The authors show that the proposed estimates are consistent with a large sample size. In their empirical study, Ang and Kristensen consider the conditional CAPM and the conditional version of the Fama-French three-factor model. When the authors test these two models on book-to-market and momentum decile portfolios, they find that both models fail to capture the value and momentum premiums.

Nagel and Singleton (2011) develop an optimal GMM estimator for conditional models with an affine SDF. The proposed moment condition set contains restrictions implied not only by the conditional mean of risky asset returns but also by that of the risk free asset. Based on the optimal GMM estimator, the authors give suggestions on the choice of instruments to minimize the asymptotic covariance matrix of parameters. Moreover, the authors derive the optimal managed portfolios which can maximize the testing power of Wald tests and LM tests. In the application of the optimal estimator to models of Jagannathan and Wang (1996), Lettau and Ludvigson (2001b), Santos and Veronesi (2006), the authors find evidence of these models' failure to match variation in conditional moments of returns.

4.3 Conditional factor models and LN's critique.

4.3.1 Two ways of conditioning-down

We start with the fundamental pricing equation

$$E[q_{t+1}^{*} R_{t+1} | \mathcal{F}_t] = 0, \qquad (1)$$

where R_{t+1} is the asset excess return, $\mathcal{F}_\mathrm{t}$ is the representative agent's information set in period *t*. In linear factor models the stochastic discount factor (SDF), q_{t+1}^{*}, is an affine function of *K* factors. For simplicity, we let $K = 2$, but one may easily add more factors to the model:

$$\boldsymbol{q_{t+1}^{*} = \phi_{t,0} + \phi_{t,1} f_{t+1,1} + \phi_{t,2} f_{t+1,2}}, \qquad (2)$$

where $\phi_{t,0}$, $\phi_{t,1}$, and $\phi_{t,2}$ are SDF's loading on 1, $f_{t+1,1}$, and $f_{t+1,2}$, respectively.

Despite the linear model's simple structure, testing the model is not a straightforward task. The difficulties arise due to the fact that $\phi_{\mathrm{t},0}$, $\phi_{\mathrm{t},1}$, and $\phi_{\mathrm{t},2}$ are functions of agents' information set, $\mathcal{F}_t$, which is unobserved. Without the knowledge of $\mathcal{F}_t$, it's impossible to condition down equation (1) without losing any information and test the model empirically. A partial solution to this issue is scaling factors by instruments (leading to a so-called scaled-factor model) z_t[40] and conditioning the model down on z_t. The implied assumption is that z_t captures most relevant information in $\mathcal{F}_t$. A similar procedure of scaling factors has been used in previous research on linear models of conditional betas such as Ferson, Kandel, and Stambaugh (1987), Harvey (1989), and Shanken (1990).

Since coefficients $\phi_{t,0}$, $\phi_{t,1}$, and $\phi_{t,2}$ are closely related to the Sharpe ratio of a mean-variance efficient return, then so long as the conditional moments of that return are time-varying so should be $\phi_{t,0}$, $\phi_{t,1}$, and $\phi_{t,2}$. If one starts with pricing equations (1)-(2), then a natural way to condition the model down is to assume that $\phi_{t,0}$, $\phi_{t,1}$, and $\phi_{t,2}$ are functions of z_t, an instrument useful for forecasting expected returns. For example, as described in Cochrane (1996), in the simplest case of a linear SDF model, factor loadings are affine functions of the instruments, $\phi_t = \phi_t^0 + \phi_t^z\, z_t$. One can then write the pricing kernel as

$$q_{t+1}^* = \phi_{t,0}^0 + \phi_{t,0}^z\, z_t + \phi_{t,1}^0 f_{t+1,1} + \phi_{t,1}^z\, z_t f_{t+1,1} + \phi_{t,2}^0 f_{t+1,2} + \phi_{t,2}^z\, z_t f_{t+1,2} \quad (3)$$

The conditioning-down procedure just described leads to a proliferation of priced factors. More precisely, even though only two factors are priced by agents, an econometrician testing the model has to assume a five-factor model with fixed coefficients. As Singleton (2006) suggests, the extra priced factors arise for every state variable z_t that is useful for predicting factor risk premiums. Adopting specification (3), we can then use the law of iterated expectations and write the unconditional version of model (1) as

[40] For illustration, we use one instrument here. Again, it's easy to extend the model to include multiple instruments.

$$E((\phi_{t,0}^{0} + \phi_{t,0}^{z}\, z_t + \phi_{t,1}^{0} f_{t+1,1} + \phi_{t,1}^{z}\, z_t f_{t+1,1} + \phi_{t,2}^{0} f_{t+1,2} + \phi_{t,2}^{z}\, z_t f_{t+1,2})R_{t+1}) = 0. \quad (4)$$

The corresponding beta representation is

$$E(R_{t+1}) = \beta'\lambda, \quad (5)$$

where $\beta = cov(f_{t+1}^{\#}, f_{t+1}^{\#})^{-1} cov(f_{t+1}^{\#}, R_{t+1})$,

$\lambda = E(f_{t+1}^{\#}) - \frac{1}{E(q_{t+1}^{*}|\mathcal{F}_t)} cov(q_{t+1}^{*}, f_{t+1}^{\#})$,

and $f_{t+1}^{\#} = [\, z_t, f_{t+1,1},\, z_t f_{t+1,1}, f_{t+1,2},\, z_t f_{t+1,2}]$.

Model (5) can be estimated by Fama-MacBeth regressions. This is the approach taken by LL01, who estimate β's for each test asset from a single time series regression over the whole data sample. The risk premia, λ, follow from the second stage cross sectional regressions of returns on betas from the first stage.

Equivalently, one may work with the beta-representation:

$$E(R_{t+1}|\mathcal{F}_t) = \beta_t{}'\lambda_t, \quad (6)$$

where β_t is a vector of expected factor betas given by $\beta_{t,i} = \frac{cov(R_{t+1}, f_{t+1,i}|\mathcal{F}_t)}{var(f_{t+1,i}|\mathcal{F}_t)}, i = 1{,}2$, λ_t is a vector of expected risk premiums on factors $f_{t+1,1}$ and $f_{t+1,2}$ conditional on $\mathcal{F}_t$.

If we assume that $\beta_{t,1}$ and $\beta_{t,2}$ are functions of z_t

$$\beta_{t,i} = \beta_{0,i} + \delta_{z,i} z_t, i = 1{,}2, \quad (7)$$

equation (6) becomes

$$E(R_{t+1} \mid \mathcal{F}_t) = (\beta_{0,1} + \delta_{z,1} z_t)\lambda_{t+1,1} + (\beta_{0,2} + \delta_{z,2} z_t)\lambda_{t+1,2}, \quad (8)$$

Taking unconditional expectation on both sides gives corresponding unconditional moments:

$$
\begin{aligned}
E(R_{t+1}) &= \beta_{0,1}E(\lambda_{t+1,1}) + \delta_{z,1}E(z_t\lambda_{t+1,1}) + \beta_{0,2}E(\lambda_{t+1,2}) + \delta_{z,2}E(z_t\lambda_{t+1,2}) \\
&= \beta_{0,1}E(\lambda_{t+1,1}) + \beta_{0,2}E(\lambda_{t+1,2}) + \delta_{z,1}[E(z_t)E(\lambda_{t+1,1})] \\
&\quad + \delta_{z,1}cov(z_t,\lambda_{t+1,1}) + \delta_{z,2}[E(z_t)E(\lambda_{t+1,2})] \\
&\quad + \delta_{z,2}cov(z_t,\lambda_{t+1,2}). \qquad (9)
\end{aligned}
$$

While both approaches start with the same conditional pricing model (1), the additional affine assumption (on ϕ's in the SDF model (2) and $\boldsymbol{\beta_{t,i}}$ in the beta model (7)) distinguishes them from each other. SDF representation with affine SDF parameters does not in any way imply betas affine in instruments. In fact, generally, betas in this case are (nonlinear) functions of past returns and factors as well as instruments. Similarly, the beta representation of the model does not imply a linear SDF with affine SDF parameters in the instruments. In fact, linear betas will likely map into a pricing kernel that is nonlinear in factors, instruments, and returns. The last result, in general, contradicts the assumption that the SDF is linear in factors, which allows one to write down the beta representation of the pricing model.

4.3.2 LN's critique of cross-sectional tests of conditional models

LN's critique of cross-sectional tests of conditional models is directly based on equation (9). LN argue that, for conditional models where factor betas follow equation (7), the slopes on $\beta_{0,1}$ and $\beta_{0,2}$ should equal $E(\lambda_{t+1,1})$ and $E(\lambda_{t+1,2})$, respectively, and slopes on $\delta_{z,1}$ and $\delta_{z,2}$ should equal $cov(z_t,\lambda_{t+1,1})$ and $cov(z_t,\lambda_{t+1,2})$, respectively, in cross-sectional regressions when instruments are demeaned. LN claim that, in studies such as JW, LL, SV, and Lustig and Van Nieuwerburgh (2005) (hereafter LVN), the constraints imposed by equation (9) are neglected. Instead, the slope on $\beta_{0,i}$ and the slope on $\delta_{z,i}$ are estimated as free parameters. The failure to impose model-implied constraints, as pointed out by LN, may explain why conditioning can seem to improve the performance of the models significantly in those studies. However, examining equations (1) to (9) carefully, one may see that LN's conclusion above is valid in some cases and is not applicable in others.

On the one hand, LN's critique is specific to conditional models where factor betas are time-varying or, more specifically, affine functions of instruments. The cross-term $\delta_{z,i}E(z_t\lambda_{t+1,i})$ emerges when we condition equation (8) down to equation (9). The critique is not applicable to the conditional SDF model specified by equations (1) to (3) due to the fact that the unconditional moments, equation (4), are derived using law of iterated expectations. Therefore no extra term emerges in the conditioning down process. Readers may also see section 5 of Ludvigson (2011) for more discussions.

On the other hand, LN's critique does not address the following issue. In equation (9), not only slopes on $\beta_{0,i}$ and $\delta_{z,i}$, but $\beta_{0,i}$ and $\delta_{z,i}$ themselves cannot be estimated as free parameters either. $\beta_{0,i}$ and $\delta_{z,i}$ are not factor betas but rather the unscaled time-varying factor betas' ($\beta_{t,i}$ in the unscaled model) loadings on 1 and the instrument z_t, which are specified by equations (7). If one estimates the model using Fama-MacBeth regressions (e.g. SV), without constraint (7) being imposed, the first pass of Fama-MacBeth regressions amounts to the estimation of an equation (3) type model, in which cross terms, $E(z_t\lambda_{t+1,i})$ are treated as extra factors in the unconditional model. The estimates of $\beta_{0,i}$ and $\delta_{z,i}$ from the first pass Fama-MacBeth regressions are correct only when all factors are excess returns. When factors are macroeconomic variables instead of excess returns on traded assets, risk premiums on factors are not equal to factors themselves. The first row of equation (9) is not equivalent to $E(R_{t+1}) = \beta_{0,i}E(f_{t,1}) + \delta_{z,1}E(z_t f_{t,1}) + \beta_{0,2}E(f_{t,2}) + \delta_{z,2}E(z_t f_{t,2})$, which are unconditional moments used in first-pass Fama-MacBeth regressions. In this circumstance, the estimates of $\beta_{0,i}$ and $\delta_{z,i}$ from the first pass Fama-MacBeth regressions can be different from the values of those beta loadings. This issue may be circumvented by forming factor-mimicking portfolios for non-return risk factors and substituting non-return factors with excess returns on corresponding factor-mimicking portfolios. However, doing so requires one to choose optimal set of assets forming the factor-mimicking portfolios so that pricing information of factors is fully captured by factor-mimicking portfolios' excess returns. Another problem with Fama-MacBeth regressions is that the model's key assumption (7) cannot be tested separately. The reason is quite straightforward. Directly checking the validity of assumption (7) requires estimation of the unscaled time-varying factor betas, which cannot be retrieved from Fama-MacBeth regressions. Given all these constraints, the most straightforward strategy to impose constraints (7) seems to be estimating the conditional time-varying betas first, and then using equation (7) to find corresponding unconditional betas.

4.3.3 Research Methodology

In their time-series analysis, LN conduct time-series intercept tests of conditional CAPM. The basic idea is fairly simple: if the conditional CAPM holds, a stock's conditional alpha should be close to zero. Also, the unconditional alpha, as shown by LN, should be approximately equal to the covariance between stock's conditional market beta and the market risk premium $cov(\beta_{it}, \lambda_{Mt})$ (see Appendix A, LN, p.311). In order to test these conditional/unconditional restrictions of the CAPM, one faces a difficult task of choosing the optimal set of instruments to condition the models down to an unconditional form amenable to empirical testing.[41] To circumvent the problem of choosing instruments, LN test the CAPM using short-window regressions $R_{it} = \alpha_i + \beta_i R_{Mt} + \varepsilon_{it}$ on size, B/M, and Momentum portfolios from July 1964 to June 2001. Instead of estimating the scaled CAPM using the entire sample of returns on test assets, LN estimate the unscaled CAPM over short periods (a month and a quarter) using high frequency data (daily and weekly) to get the estimates of conditional alpha and conditional beta over the estimation window. In doing so, as long as the market beta is stable in the estimation window, LN can get reliable estimates of conditional betas and conditional alphas. The empirical results turn out to be discouraging. The conditional alphas are large, especially for B/M and Momentum portfolios. Moreover, although conditional betas are quite volatile over time, the estimated values of $cov(\beta_{it}, \lambda_{Mt})$ are nowhere near the large estimated unconditional alphas. In short, they conclude that the CAPM is unlikely to be salvaged by time-varying betas.

However, the short-window regression approach as described in LN cannot be applied to conditional factor models in which factors are not returns. The reason is straightforward. When factors are not excess returns on traded assets, the corresponding risk premiums cannot be written as moments of factors themselves. Therefore, the covariance between asset returns and a factor (see equation A.1 of LN) is related not to the variance of the factor but covariance between risk premium and the factor. To see this, let's consider a one-factor case $R_{it} = \beta_t \lambda_t^f + \varepsilon_t$, in which factor f_t is not an excess return. Assuming $\beta_t = \beta_0 + \delta_t$, equation A.1 of LN becomes

[41] Nagel and Singleton (2011) is a recent attempt to formalize tests of conditional moment restrictions of asset pricing models.

$$
\begin{aligned}
cov(R_{it}, f_t) &= cov\left[(\beta_0 + \delta_t)\lambda_t^f, f_t\right] \\
&= \beta_0 cov\left(\lambda_t^f, f_t\right) + cov\left(\delta_t \lambda_t^f, f_t\right). \qquad (10)
\end{aligned}
$$

And the value of λ_t^f depends on the covariance between the SDF and factor f_t. Without specifying the functional form of the SDF and using instruments to capture the variability of factor loadings, there is no simple way to expand the left-hand side of equation (10) for a conditional non-return factor model. One may get around this problem by using a factor-mimicking portfolios approach. But choosing optimal set of assets forming the factor-mimicking portfolios is as difficult a task as choosing the optimal set of instruments.

Because of the difficulty in conducting LN's time-series analysis on conditional models with non-return factors, we proceed by addressing LN's critique of cross-sectional tests of conditional models and incorporating linear approximation constraint (7) into empirical tests.

First, we assess the importance of imposing constraint (7) in testing the time-varying beta models. We use a two-stage regression to estimate $\beta_{0,i}$ and $\delta_{z,i}$ in equation (9) for a conditional-beta model. In the first stage, we estimate the factor beta of the proposed model by running the first-pass Fama-MacBeth rolling-window regression:

$$R_{j,t} = \beta_{j,t,1} f_{t,1} + \beta_{j,t,2} f_{t,2} + e_t, j = 1,2, \dots, n,$$

where $R_{j,t}$ is the excess return on test asset j, $f_{t,1}$ and $f_{t,2}$ are factors. In the rolling window regression, we estimate $\beta_{j,t,1}$ and $\beta_{j,t,2}$ using data on factors and excess returns on each test asset j from period t-n to period t-1, assuming that $\beta_{j,t,1}$ and $\beta_{j,t,2}$ can be reliably estimated by using data points over time period t-n to t. The first stage results in a time series of estimates of factor betas, $\beta_{j,t,1}$ and $\beta_{j,t,2}, t = 1,2, \dots, T - n,$ for each test asset j. To choose the length of beta estimation window, we follow Fama and MacBeth (1973), and set $n = 60$ or five years of monthly data.

In the second stage, we run time series OLS regressions $\beta_{j,t,i} = \beta_{j,0,i} + \delta_{j,z,i} z_t$, $i = 1,2$, to retrieve the cross section of time-series estimates of $\beta_{j,0,i}$ and $\delta_{j,z,i}$.

To address LN's critique of cross-sectional regressions, we use the second stage estimates of $\beta_{j,0,i}$ and $\delta_{z,i}$ in Fama-MacBeth cross-sectional regressions to estimate the slopes on $\beta_{j,0,i}$ and $\delta_{z,i}$. In the rest of the paper we call this procedure of estimating $\beta_{j,0,i}$ and $\delta_{z,i}$, together with the second pass Fama-MacBeth cross-sectional regressions mentioned above, the rolling-beta regressions.

To examine the impact of imposing constraint (7), we also estimate the model using the Fama-MacBeth regressions with constraint (7) omitted. Specifically, we estimate $\beta_{j,0,i}$ and $\delta_{j,z,i}$ in the first-pass regression $R_{j,t+1} = \beta_{j,0,1} f_{t,1} + \delta_{j,z,1} z_t f_{t,1} + \beta_{j,0,2} f_{t,2} + \delta_{j,z,2} z_t f_{t,2} + e_{it}, j = 1,2, \ldots, n$, and then feed the estimates of $\beta_{j,0,i}$ and $\delta_{j,z,i}$, $i = 1,2$, into the second-pass cross-sectional regression to estimate slopes on $\beta_{j,0,i}$ and $\delta_{j,z,i}$.

By comparing the rolling-beta results, which impose constraint (7), with corresponding values retrieved from the unconstrained Fama-MacBeth regressions, we should be able to judge if constraint (7) has a significant effect on the model performance. Moreover, if constraint (7) is critical, the Fama-MacBeth procedure may not be an appropriate way to test conditional models with conditioned-down beta representation given by equation (9).

Whether constraint (7) is critical or not, there still remains an issue of generated regressors. This problem occurs in multi-stage regressions when estimates from previous stages are used as inputs in later stages. Since estimates from early stages are imprecise, treating them as true parameter values ignores their sampling variation and leads to biased estimates of parameters in later stages. Unfortunately, our rolling-beta regression is a three-stage procedure. In both second and third stage regressions, we use estimates from previous stages as inputs. It's likely that our estimates from second and third stage regressions are subject to the problem. One feasible way to fix this problem is to estimate all parameters simultaneously using GMM. However, doing so leaves us with no closed-form solution for GMM estimates. Moreover, there is a much more serious problem with using GMM in this case. Since we need to estimate time-varying factor betas, the number of moment conditions needed is large compared to the length of the data sample, which leads to the dimensionality problem and makes GMM infeasible. We return to this problem in robustness checks in the Appendix.

Our next task is to address LN's critique of conditional models. LN suggest that the slope on $\beta_{j,0,i}$ should be the average risk premium on factor f_i, $\hat{\lambda}_{t+1,i}$, while that on $\delta_{j,z,i}$ should be $E\left(z_t \lambda_{t+1,i}\right)$. Two natural hypothesises to test are

$$H_0: \hat{\gamma}_i = \hat{\lambda}_{t+1,i}, \tag{H.1}$$

$$\text{and} \quad H_0: \hat{\gamma}_{z,i} = \hat{E}\left(z_t \lambda_{t+1,i}\right), \tag{H.2}$$

where $i = 1,2$, $\hat{\gamma}_i$ and $\hat{\gamma}_{z,i}$ are estimated slopes on $\beta_{j,0,i}$ and $\delta_{j,z,i}$ from the second pass Fama-MacBeth regressions on both scaled and unscaled factors, $\hat{\lambda}_{t+1,i}$ is the mean of a factor risk premium estimated by running rolling-window regressions of the unscaled model, and $\hat{E}\left(z_t \lambda_{t+1,i}\right)$ is the vector of the estimated mean of $z_t \lambda_{t+1,i}$.

First, to retrieve $\hat{\lambda}_{t+1,i}$, we estimate the non-scaled model specified by equation (1) using the rolling-window Fama-MacBeth regressions, where we first estimate time-varying betas $\beta_{j,t,i}$ using data from period t-60 to period t-1 and then feed our estimates of $\beta_{j,t,i}$ into the cross-sectional regressions as independent variables to estimate factor risk premiums. With the series of estimated risk premiums on the original factors f_i in hand, it's easy to compute both $\hat{\lambda}_{t+1,i}$ and $\hat{E}\left(z_t\lambda_{t+1,i}\right)$. Since both $\hat{\lambda}_{t+1,i}$ and $\hat{E}\left(z_t\lambda_{t+1,i}\right)$ are estimates of risk prices following LN's suggestion, we call them $\hat{\gamma}_{LN}$ hereafter.

Second, we estimate β_{i0} and δ_i for each asset in three steps. Using the rolling-beta regressions, we first estimate time-varying betas of the unscaled model (1). Then we regress time series of the estimated time-varying betas on $[1, z_t]$ to retrieve estimates of $\beta_{j,0,i}$ and $\delta_{j,z,i}$ for each test asset. Finally, with the estimates of $\beta_{j,0,i}$ and $\delta_{j,z,i}$, we run the second pass of Fama-MacBeth regressions to estimate $\hat{\gamma}_i$and $\hat{\gamma}_{z,i}$.

With all four ingredients in hand, we test (H.1) and (H.2) by conducting the sample mean tests.

4.4 Empirical Results

In this section we assess the validity and impact of LN's critique as well as that of the linear approximation constraint (7) on the estimation of conditional affine factor models. We use the labor income ratio model proposed by SV as an example in our analysis. We conduct two empirical tests of the labor income ratio model, which are described in detail below. In the first test, we show that the Fama-MacBeth procedure may not be an appropriate way to test conditional models with conditioned-down beta representation as in equation (9). The second test is a hypothesis test designed to address LN's critique of conditional models.

4.4.1 The labor income ratio model.

The labor income ratio model is a typical time-varying beta model in which the expected excess asset return has the following beta-style representation

$$E_t\left(R_t^i\right) = \beta^{w,i}(s_t)E_t(R_t^w) + \beta^{M,i}(s_t)E_t(R_t^M), \tag{11}$$

where R_t^i is the excess return on individual assets, R_t^w is the excess return on human capital, and R_t^M is the excess market return. To condition down the model, SV assume that betas are affine in the instrument, s_t^w:

$$\beta^{w,i}(s_t) \approx \beta_1^{w,i} + \beta_2^{w,i} s_t^w, \tag{12}$$

$$\beta^{M,i}(s_t) \approx \beta_1^{M,i} + \beta_2^{M,i} s_t^w, \tag{13}$$

where the conditioning variable, s_t^w, is the labor income to consumption ratio. s_t^w, suggested by SV, is a good predictor of stock returns. Results of the predictability regression reported in table 2 of SV show that s_t^w is able to explain a large proportion of variation in future market excess return in the four-year horizon case. The predictability of s_t^w survives when log dividend-to-price ratio is included in the model.

The rationale behind this result is straightforward. Assume that the representative agent only has two income sources, labor income and dividends. When s_t^w is large, dividends from financial assets only account for a small portion of consumption and, therefore, they are less correlated with consumption. In the opposite case, since consumption is mainly funded by dividends, the covariance between them will be high, which leads to a high premium on financial assets.

4.4.2 Data.

We test the model at a monthly frequency. As a robustness check, we also test the model at a quarterly frequency to see if data frequency affects test results. Since the data series on monthly log real consumption c_t and log labor income y_t are not reported by BEA, we estimate both series using corresponding nominal data series from BEA. To estimate monthly real c_t, we first of all estimate the Fisher price index by treating personal consumption expenditure (PCE) on nondurable goods and that on services as two indivisible components of c_t. Then, we use our estimates of the Fisher price index to deflate nominal c_t to obtain real values. Following Lettau and Ludvigson (2001a), we construct monthly real y_t based on data on personal income reported in NIPA table 2.6[42] from BEA's website. The details of constructing monthly real c_t and y_t are given in the Appendix. Both c_t and y_t are real in all our tests. Δc_t is computed by subtracting c_{t-1} from c_t. Following SV, we calculate s_t^w and R_t^w using log real wage y_t as the proxy for log real labor income. With the series of c_t and y_t in hand, we compute s_t^w as $s_t^w = \exp\{y_t - c_t\}$ and R_t^w as $R_t^w = \exp\{y_t - y_{t-1}\}$. The series of monthly nominal market returns are retrieved through WRDS and that of the risk-free rate from Kenneth R. French's online data library[43]. The proxy for the risk-free rate, R_t^f, is the one-month Treasury bill rate. Since the monthly consumption data series starts in January 1959 and WRDS' monthly market return ends in December 2013, we restrict our analysis of the labor income ratio model to monthly frequency in this period. We deflate both data series of nominal market return and risk-free rate by Consumer Price Index (CPI)[44] inflation rate to obtain corresponding real data series. For robustness checks, we use quarterly

[42] http://www.bea.gov/iTable/iTable.cfm?ReqID=9&step=1#reqid=9&step=3&isuri=1&904=2013&903=76&906=q&905=2014&910=x&911=1

[43] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ftp/F-F_Research_Data_Factors.zip

[44] Consumer Price Index for 'All Urban Consumers: All Items', retrieved from Federal Reserve Economic Data (http://research.stlouisfed.org/fred2/series/CPIAUCSL). For quarterly data, we use the average of three monthly price levels in a quarter as the price level of the quarter.

market excess returns, which we obtain by compounding both real market returns and the real risk-free rate over each quarter, and computing the difference.

When examining the labor income ratio model's ability to fit the cross section of Fama-French 25 portfolios sorted on size and book to market ratio (B/M) (SBM25) (table 6 of SV, p 29), instead of using the product of instrument and original factor ($s^w R^M$, and $s^w R^w$), SV use the component of product that is orthogonal to the original factor to circumvent the multicollinearity problems. Following SV, we orthogonalize the cross-terms by regressing $s^w R^M$, and $s^w R^w$ on R^M and R^w, respectively, and use the residuals of the regressions as new cross-terms. This leads to new conditional variables z^M and z^w given by

$$z^M = R^{M^{*-1}} M^M (s^w \circ R^M), \tag{14}$$

$$z^w = R^{w^{*-1}} M^w (s^w \circ R^w), \tag{15}$$

where $M^M = I_T - R^{M^*} \left(R^{M^{*\prime}} R^{M^*} \right)^{-1} R^{M^{*\prime}}$ and $M^w = I_T - R^{w^*} \left(R^{w^{*\prime}} R^{w^*} \right)^{-1} R^{w^{*\prime}}$ are annihilator matrices, $R^{M^*} = [\mathbf{1}, R^M]$, $R^{w^*} = [\mathbf{1}, R^w]$, $\mathbf{1}$ is the column vector of 1's, and T is the length of our data sample.

Our test assets are real excess returns on Fama-French 25 value-weighted portfolios sorted on size and book to market ratio (B/M) (SBM25). Monthly nominal returns on SBM 25 are retrieved from Kenneth R. French's online data library[45]. Like all other returns mentioned above, returns on SBM25 are deflated to the real returns using CPI inflation rate. Corresponding real excess returns are obtained as the difference between the real portfolio returns and R_t^f.

[45] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ftp/25_Portfolios_5x5.zip

We also need to deal with the fact that returns on SBM25 exhibit strong factor structure. As suggested by Lewellen, Nagel, and Shanken (2010) (LNS hereafter), only very mild restrictions on a group of factors are required for the factors to have an arbitrarily good explanatory power for the cross–sectional variation of expected returns on these portfolios. One way to circumvent this limitation of SBM25 is to conduct an out-of-sample test. Following LNS, we perform the cross-sectional version of an out-of-sample test. Specifically, we use excess returns on 30 Fama-French value-weighted industry portfolios (IND30) as out-of-sample test assets. As a robustness check, in a later section we also run our test on 10 value-weighted momentum portfolios (M10) and 6 value-weighted earnings to price ratio portfolios (EP6)[46] . Returns on IND30, M10, and EP6 are from Kenneth R. French's online data library and converted to both monthly and quarterly real excess returns in the same fashion as returns on SBM25.

Table 4.1A Summary statistics for monthly state variables.
This table reports the summary statistics of monthly state variables. The state variables are real market excess return (R_t^M), labor income to consumption ratio (s_t^w), real excess return on human capital (R_t^w). The sample period is February 1959 to December 2013. Std.error is the sample standard deviation. ρ is the first-order autocorrelation coefficient.

Variable	R_t^M	R_t^w	s_t^w
Mean	0.0039	-0.0017	0.9672
Std.error	0.0448	0.0077	0.0471
Min.	-0.2621	-0.0407	0.8803
Max.	0.1475	0.0484	1.0824
ρ	0.0978	-0.0207	0.9856

Table 4.1A presents the summary statistics of the monthly state variables we use in our analysis of the labor income ratio model. The instrument s_t^w is highly persistent

[46] The first portfolio in this sample is the portfolio of stocks with negative earnings to price ratio. The remaining five are quintile-portfolios sorted on E/P ratio.

with first-order autocorrelation coefficient $\rho = 0.9856$ in our monthly data sample. R_t^M is less persistent while R_t^W has a small negative ρ.

Table 4.1B Summary statistics for test assets (monthly).
Panel A reports the summary statistics of monthly real excess returns on 25 size and book to market portfolios (SBM25). Panel B presents the summary statistics of monthly real excess returns on 30 industry portfolios (IND30). The sample period is February 1959 to December 2013. Std.error is the sample standard deviation.

Panel A: real excess returns on SBM25						
		Low	2	3	4	High
Mean	Small	0.26%	0.77%	0.79%	0.99%	1.11%
	2	0.45%	0.70%	0.89%	0.92%	1.01%
	3	0.49%	0.78%	0.75%	0.87%	1.01%
	4	0.58%	0.58%	0.71%	0.83%	0.81%
	Big	0.46%	0.50%	0.51%	0.54%	0.63%
		Low	2	3	4	High
Std.error	Small	0.079	0.068	0.059	0.056	0.060
	2	0.071	0.059	0.054	0.053	0.059
	3	0.066	0.054	0.050	0.049	0.054
	4	0.059	0.051	0.050	0.048	0.055
	Big	0.047	0.044	0.043	0.043	0.050

Panel B: real excess returns on IND30						
	Food	Beer	Smoke	Games	Books	Hshld
Mean	0.69%	0.69%	0.96%	0.75%	0.56%	0.58%
Std.error	0.044	0.051	0.061	0.072	0.058	0.048
	Clths	Hlth	Chems	Txtls	Cnstr	Steel
Mean	0.69%	0.68%	0.53%	0.67%	0.53%	0.29%
Std.error	0.064	0.050	0.055	0.071	0.060	0.072
	FabPr	ElcEq	Autos	Carry	Mines	Coal
Mean	0.57%	0.76%	0.48%	0.72%	0.52%	0.90%
Std.error	0.061	0.063	0.067	0.064	0.074	0.097
	Oil	Util	Telcm	Servs	BusEq	Paper
Mean	0.66%	0.46%	0.51%	0.68%	0.59%	0.51%
Std.error	0.053	0.040	0.046	0.066	0.068	0.051
	Trans	Whlsl	Rtail	Meals	Fin	Other
Mean	0.55%	0.65%	0.67%	0.71%	0.58%	0.38%
Std.error	0.057	0.056	0.054	0.062	0.054	0.059

Table 4.1B reports summary statistics of test assets at the monthly frequency. In panel A of table 4.1B, higher B/M is in general associated with higher excess returns and lower standard deviation. Big portfolios tend to have lower but more stable excess returns compared to small portfolios. Unlike those of SBM 25, real excess returns on IND30 display no obvious pattern.

4.4.3 Test Results.

In figure 4.1, we report adjusted R-squares of OLS regressions of the second stage of the rolling-beta regressions (rolling window betas regressed on instruments) for both 25 size and book to market portfolios and 30 industry portfolios. In 25 size and book to market portfolios (SBM25) case in Figure 4.1.1, the largest adjusted R-square (22.1%) occurs in the regression of the first-stage estimated market beta of the medium-size value portfolio on SV's instrument (labor income to consumption ratio). The average adjusted R-square is 8.07% for regressions of market beta and 1.52% for those of labor income beta. The fact that all $2 * 25$ adjusted R-squares (corresponding to 25 market betas and 25 human capital betas) are all below 25% with close-to-zero mean suggests that the major part of the variation in first-stage estimates of time-varying factor betas is not well explained by the linear approximation (7). Further, these adjusted R-squares display patterns that we should not see if eq. (7) is a good approximation of time varying betas. Portfolios of smaller-size stocks tend to have higher adjusted R-squares for both R^M and R^w. Portfolios with moderate B/M ratios have in general the lowest adjusted R-square for factor beta of R^w. In the case of IND30 reported in Figure 4.1.2, the SV's instrument has a slightly better performance in capturing the variation in the unscaled factor betas than in the case of SMB25. For the portfolio of Meals industry (portfolio 28 in figure 4.1.2), the

SV's instrument is able to explain 39.48% of the variation in the market factor beta. However, the average adjusted R-square is still low for both market factor beta (13.87%) and labor income factor beta (1.24%). Again, these results demonstrate that the linear constraint on betas (7) is likely problematic in that it has a poor fit to rolling beta estimates from the first stage.

Figure 4.1 Adjusted R-square of Regressions $\beta_t^i = \beta_0^i + \delta^i s_t^w + e_t^i$.

This figure reports the adjusted R-squares of regressions $\beta_t^i = \beta_0^i + \delta^i s_t^w + e_t^i$, the second stage of the rolling-beta regressions. Test asset returns are monthly real excess returns on 25 size and book to market portfolios (SBM25 and 30 industry portfolios (IND30) given in figures 4.1.1 and 4.1.2, respectively. The sample period is February 1959 to December 2013. The solid line is for the adjusted R-squares of regressions of market betas, β_t^M, and the dashed line is for those of regressions of the human capital betas, β_t^w.

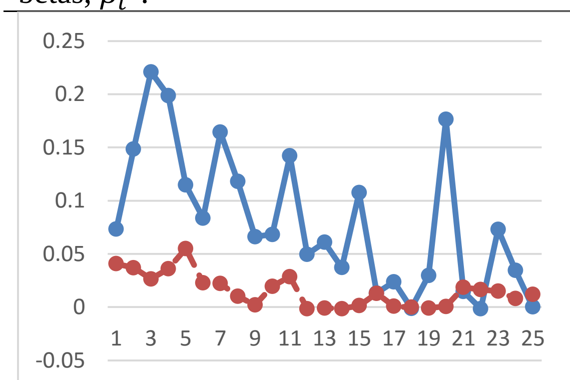


Figure 4.1.1 SBM25

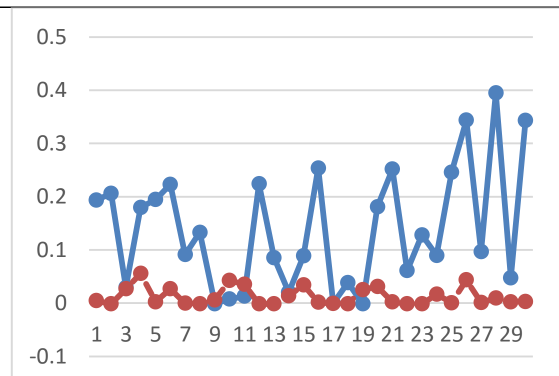


Figure 4.1.2 IND30

Table 4.2 reports the number of portfolios in which Fama-MacBeth estimates of factor betas are outside the 95% confidence interval of the corresponding estimates from rolling-beta regressions across both samples. In general, Fama-MacBeth estimates of factor betas are inconsistent with the corresponding rolling-beta estimates. Only for the unscaled market factor for IND30 portfolios, the number of portfolios with inconsistent beta estimates in two different approaches is less than 50% of the total number of portfolios. Moreover, betas of the scaled factors, β^{zM} and β^{zw}, are more inconsistent in the two approaches than their unscaled counterparts.

Table 4.2. Empirical results for the labor income ratio model.
This table reports the numbers of portfolios for which the Fama-MacBeth estimated factor betas are outside the 95% confidence interval of the corresponding estimates from rolling-beta regressions. The tested model is the labor income ratio model. Test asset returns are monthly real excess returns on 25 size and book to market portfolios (SBM25 and 30 industry portfolio (IND30). The sample period is February 1959 to December 2013.

	β^{M}	β^{zM}	β^{w}	β^{zw}
SBM25	19	20	16	25
IND30	10	26	23	30

Factor risk premium estimates are displayed in Table 4.3. In panel A, we compare the Fama-MacBeth estimates of risk prices with our rolling-beta regressions results. For real excess returns on SBM25, without constraint (7) imposed, the slope associated with R^{M} is positive and statistically significant. Estimates of risk price on R^{w} (-0.016% per month) is significantly negative, its absolute value is roughly 20 times larger than that on $z^{w}R^{w}$. In the case of rolling-beta regressions (with constraint (7) imposed), estimated slopes associated with R^{M} are similar to their counterparts from the Fama-MacBeth regressions and have similar significance levels. However, substantially different estimates are obtained for scaled market return, $z^{M}R^{M}$, and return on wealth, R^{w}. While point estimate of risk price on $z^{M}R^{M}$ is statistically significant and 4 times larger than its counterpart, the price of risk on R^{M} is 50% smaller in scale than the corresponding value from the Fama-MacBeth regressions.

Table 4.3 Empirical results of labor income ratio model.

This table reports the main results of empirical tests of the labor income model. Test returns are monthly real excess returns on 25 size and book to market portfolios (SBM25) and 30 industry portfolios (IND30). The sample period is February 1959 to December 2013. Panel A presents the estimation of risk prices on the unscaled and scaled market factor, R^M and $z^M R^M$,and the unscaled and scaled labor income factor, R^w and $z^w R^w$, from both Fama-MacBeth regressions, $\hat{\gamma}_{FM}$, and rolling-beta regressions, $\hat{\gamma}_{RB}$. *t*-stats with Newey-West adjustment for autocorrelation (*lag* = 3) are reported in parentheses below the point estimates. $RMSE$ is the root mean squared error of the model. Paired difference test in the third section of panel A tests the hypothesis $\boldsymbol{H_0}: \hat{\gamma}_{RB} - \hat{\gamma}_{FM}$, t_{RB-FM} is computed as $t_{RB-FM} = \sqrt{T}\frac{\hat{\gamma}_{RB}-\hat{\gamma}_{FM}}{\sqrt{\hat{\sigma}^2_{\gamma_{RB,t}}+\hat{\sigma}^2_{\gamma_{FM,t}}-2\widehat{cov}(\gamma_{RB,t},\gamma_{FM,t})}}$, where $\hat{\sigma}^2_{\gamma_{RB,t}}$, $\hat{\sigma}^2_{\gamma_{FM,t}}$, and $\widehat{cov}(\gamma_{RB,t},\gamma_{FM,t})$ are estimated from the time series of $\hat{\gamma}_{RB,t}$ and $\hat{\gamma}_{FM,t}$ with Newey-West adjustment. Panel B presents the results of testing hypothesis (H.1) and (H.2). $\hat{\gamma}_{LN}$ is the estimate of slopes on β_{i0} and δ_i suggested by equation (9). $t^{LN}_{RB} = \frac{\hat{\gamma}_{RB}-\hat{\gamma}_{LN}}{\sqrt{\hat{\sigma}^2_{\gamma_{RB,t}}/T}}$. $RMSE_{LN}$ is the root mean squared error of the model after imposing constraint (9).

Panel A						
Fama-MacBeth procedure						
		R^M	$z^M R^M$	R^w	$z^w R^w$	*RMSE*
$\hat{\gamma}_{FM}$	SBM25	**0.0043****	3.2 E-04	**-0.016*****	-4.3423 E-05	0.0018
		(2.2008)	(0.6246)	(-3.3060)	(-0.2129)	
	IND30	**0.0056*****	2.3 E-04	-6.5 E-04	**2.0 E-04***	0.0013
		(2.8635)	(1.1678)	(-0.3168)	(1.7873)	
Rolling-beta regressions						
$\hat{\gamma}_{RB}$	SBM25	**0.004****	**0.0016*****	**-0.0107*****	-5.7 E-04	0.0015
		(2.0663)	(3.0824)	(-3.3105)	(-1.3033)	
	IND30	**0.0052*****	2.8 E-04	-0.0021	1.6 E-04	0.0013
		(2.6096)	(0.7817)	(-1.3506)	(0.5345)	
Paired Difference Test $(\boldsymbol{H_0}: \hat{\gamma}_{RB} - \hat{\gamma}_{FM} = 0)$						
t_{RB-FM}	SBM25	-1.3305	**3.7051*****	**2.0303****	**-1.8014***	
	IND30	**-2.0187****	0.1935	-0.7028	-0.1819	
Panel B: LN's critique						
SMB25	$\hat{\gamma}_{LN}$	0.0055	**-1.0 E-05*****	**-0.0015*****	8.0 E-06	0.0067
	t^{LN}_{RB}	-0.7787	3.1018	-2.8611	-1.3218	
IND30	$\hat{\gamma}_{LN}$	0.0050	-3.4 E-05	-0.0012	-5.2 E-05	0.0016
	t^{LN}_{RB}	0.1030	0.8769	-0.6013	0.7127	

Results of paired difference test of $\boldsymbol{H_0}: \hat{\gamma}_{RB} - \hat{\gamma}_{FM} = 0$, where $\hat{\gamma}_{RB}$ and $\hat{\gamma}_{FM}$ are risk price estimates in rolling-beta regressions and Fama-MacBeth regressions, respectively, confirms that the differences between estimates associated with both $z^M R^M$ and R^w in the two estimation approaches are statistically significant. To measure the overall model performance, we calculate the root mean squared error (*RMSE*) of pricing errors for both rolling-beta and Fama-MacBeth regressions: $RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \hat{\alpha}_i^2}$, where $\hat{\alpha}_i, i = 1, \dots, N$ is the average pricing error on asset i. $RMSE$ in the rolling-beta regressions (0.0015) is more than 15% smaller than that in the Fama-MacBeth regressions (0.0018) for SBM25 sample in column 7 of Table 4.3.

Test results based on real excess returns on IND30 in the first and second section of Panel A in Table 4.3 are relatively weaker than those in the SBM25 sample. The estimated slope associated with R^w is no longer significant regardless of whether constraint (7) is imposed or not. The t-statistics, t_{RB-FM}, for the difference between prices of risk from the two estimation approaches reported in the third section of Panel A of Table 4.3 show that only the risk price of R^M changes significantly from one estimation approach to the other. Regarding the overall model performance, $RMSE$ in the rolling-beta regressions is roughly the same as that in the Fama-MacBeth regressions.

Overall, our results above show that imposing constraint (7) changes estimation results for factor betas as well as corresponding risk premiums. To sum up, when some model factors are not excess returns, estimating the model using the Fama-MacBeth procedure omits the unconditional constraint (7) on β_{i0} and δ_i. Due to this problem, the Fama-MacBeth results strongly differ from our 3-stage regression estimation results, and, thus the Fama-MacBeth procedure may not be an appropriate way to estimate β_{i0} and δ_i of a time-varying beta model.

In Panel B of Table 4.3, we report our estimates of $\lambda^M, \lambda^w, E(z^M\lambda^M)$, and $E(z^w\lambda^w)$ following LN's suggestion, $\hat{\gamma}_{LN}$, as well as t-statistics of sample mean tests of (H.1) and (H.2), t_{RB}^{LN}, for all four test asset groups. t_{RB}^{LN} is computed as $t_{RB}^{LN} = \frac{\hat{\gamma}_{RB}-\hat{\gamma}_{LN}}{\sqrt{\hat{\sigma}^2_{\gamma_{RB,t}}/T}}$. For SBM25, the rolling-beta regressions overestimate the risk price on the scaled market factor while producing much larger negative risk price on the human capital factor. When running tests on real excess returns on IND30, none of the four estimates of risk prices are significantly different from their model implied values. The only deficiency is that $\hat{\gamma}_{RB}$'s on $z^M R^M$ and $z^w R^w$ are of the opposite sign to that of the corresponding $\hat{\gamma}_{LN}$'s. Imposing constraints suggested by LN's critique also worsens model performance. In the case of SBM25, $RMSE$ increases from 0.0015 to 0.0067 when $\hat{\gamma}_{LN}$, instead of $\hat{\gamma}_{RB}$, are used when computing pricing errors. Similarly, $RMSE$ increases drastically in IND30 sample.

4.4.4 Out-of-sample Tests.

In this section, we conduct out-of-sample tests on 10 momentum portfolios (M10) and 6 earnings to price portfolios (EP6). Test results are reported in Figure 4.2, table 4.4, and table 4.5.

Figure 4.2 reports the adjusted R-squares of the second stage OLS regressions in our three-stage regressions in both M10 and EP6 samples. Similar to what we have in the SBM25 and IND30 samples, the average adjusted R-squares, especially those for the human capital beta, are low in both samples.

Figure 4.2 Adjusted R-square of Regressions $\beta_t^i = \beta_0^i + \delta^i s_t^w + e_t^i$.
This figure reports the adjusted R-squares of regressions $\beta_t^i = \beta_0^i + \delta^i s_t^w + e_t^i$, the second stage of the rolling-beta regressions. Test asset returns are monthly real excess returns on 10 momentum portfolios (M10), and six earnings-to-price portfolios (EP6) given in figures 4.2.1 and 4.2.2, respectively. The sample period is February 1959 to December 2013. The solid line is for the adjusted R-squares of regressions of market betas, β_t^M, and the dashed line is for those of regressions of the human capital betas, β_t^w.

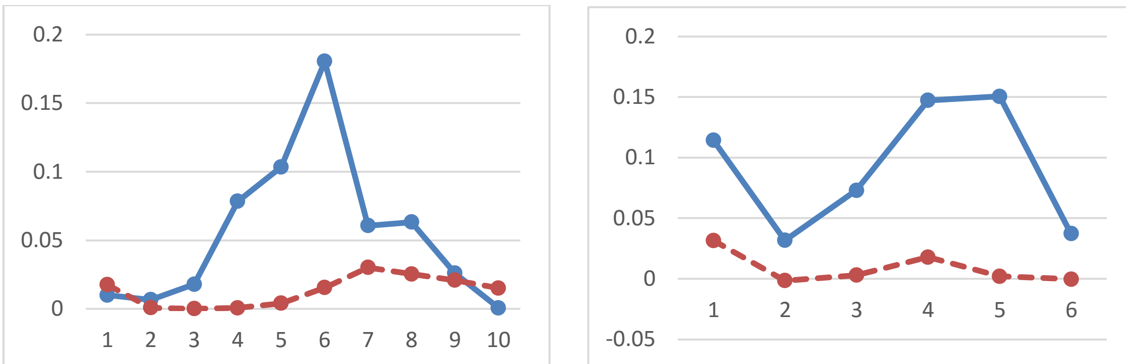


Figure 4.2.1 M10 Figure 4.2.2 EP6

Table 4.4 is the M10 and EP6 version of table 4.2. Results in table 4.4 are similar to their counterparts in table 4.2. For all four betas, more than 50% of the estimates are significantly different in two estimation approaches in both M10 and EP6 samples.

Table 4.4 Empirical results of labor income ratio model.
This table reports the numbers of portfolios for which the Fama-MacBeth estimated factor betas are outside the 95% confidence interval of the corresponding estimates from rolling-beta regressions. The tested model is the labor income ratio model. Test asset returns are monthly real excess returns on 10 momentum portfolios (M10) and six earnings-to-price portfolios (EP6). The sample period is February 1959 to December 2013.

	β^M	β^{zM}	β^w	β^{zw}
M10	7	9	7	7
EP6	3	4	5	6

In panel A of table 4.5, we compare the risk premium estimates retrieved from two different approaches.

Table 4.5 Empirical results of labor income ratio model.

This table reports the main results of empirical tests of the labor income model. Test returns are monthly real excess returns on 10 momentum portfolios (M10), and six earnings-to-price portfolios (EP6). The sample period is February 1959 to December 2013. Panel A presents the estimation of risk prices on the unscaled and scaled market factor, R^M and $z^M R^M$, and the unscaled and scaled labor income factor, R^w and $z^w R^w$, from both Fama-MacBeth regressions, $\hat{\gamma}_{FM}$, and rolling-beta regressions, $\hat{\gamma}_{RB}$. t-stats with Newey-West adjustment for autocorrelation (*lag* = 3) are reported in parentheses below the point estimates. $RMSE$ is the root mean squared error of the model. Paired difference test in the third section of panel A tests the hypothesis $\boldsymbol{H_0}: \hat{\gamma}_{RB} - \hat{\gamma}_{FM} = 0$. t_{RB-FM} is computed as $t_{RB-FM} = \sqrt{T}\frac{\hat{\gamma}_{RB}-\hat{\gamma}_{FM}}{\sqrt{\hat{\sigma}^2_{\gamma_{RB,t}}+\hat{\sigma}^2_{\gamma_{FM,t}}-2\widehat{cov}(\gamma_{RB,t},\gamma_{FM,t})}}$, where $\hat{\sigma}^2_{\gamma_{RB,t}}, \hat{\sigma}^2_{\gamma_{FM,t}}$, and $\widehat{cov}(\gamma_{RB,t}, \gamma_{FM,t})$ are estimated from the time series of $\hat{\gamma}_{RB,t}$ and $\hat{\gamma}_{FM,t}$ with Newey-West adjustment. Panel B presents the results of testing hypothesis (H.1) and (H.2). $\hat{\gamma}_{LN}$ is the estimate of slopes on β_{i0} and δ_i suggested by equation (9). $t^{LN}_{RB} = \frac{\hat{\gamma}_{RB}-\hat{\gamma}_{LN}}{\sqrt{\hat{\sigma}^2_{\gamma_{RB,t}}/T}}$. $RMSE_{LN}$ is the root mean squared error of the model after imposing constraint (9).

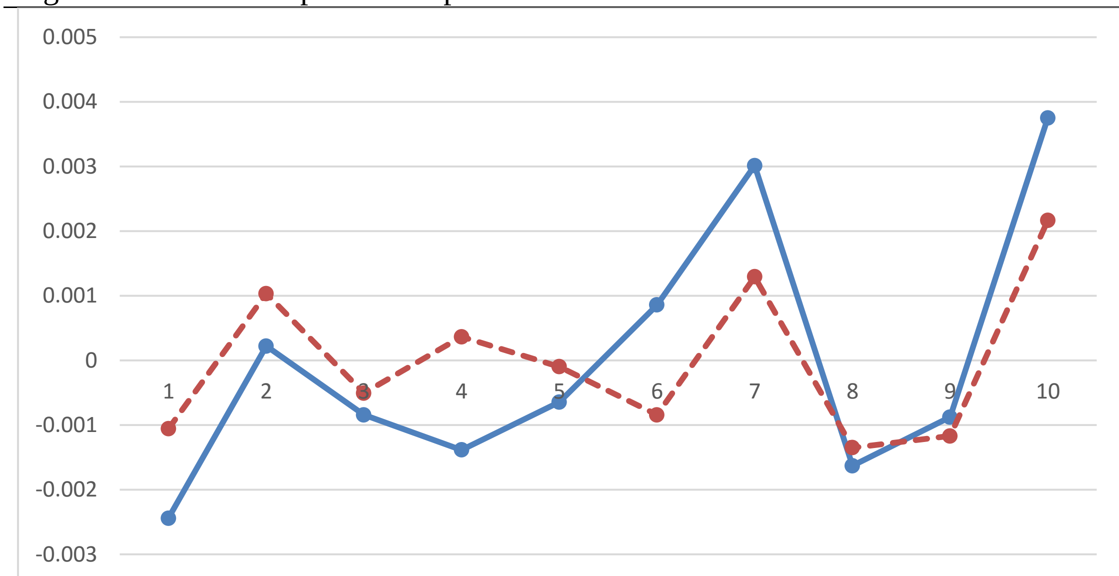
Panel A						
Fama-MacBeth procedure						
		R^M	$z^M R^M$	R^w	$z^w R^w$	*RMSE*
$\hat{\gamma}_{FM}$	M10	**0.0045****	3.3 E-04	**-0.0176*****	6.8 E-06	0.0011
		(2.0957)	(0.5653)	(-4.1800)	(0.0280)	
	EP6	**0.0068*****	**-0.0051*****	**-0.0404*****	**7.3 E-04***	0.0005
		(3.3957)	(-3.6304)	(-3.7465)	(1.9136)	
Rolling-beta regressions						
$\hat{\gamma}_{RB}$	M10	0.0033	0.0011	**-0.0229*****	**-0.0023*****	0.0019
		(1.6251)	(0.6686)	(-4.7455)	(-3.0593)	
	EP6	**0.0056*****	**-0.0035*****	**-0.0200*****	**-0.0023****	0.0005
		(2.8989)	(-2.6001)	(-3.3359)	(-2.2162)	
Paired Difference Test $(\boldsymbol{H_0}: \hat{\gamma}_{RB} - \hat{\gamma}_{FM} = 0)$						
t_{RB-FM}	M10	-1.5390	0.04320	**-3.0498*****	**-2.6386*****	
	EP6	**-3.3881*****	**2.3905****	**3.3862*****	**-2.9137*****	
Panel B: LN's critique						
M10	$\hat{\gamma}_{LN}$	0.0048	-2.7 E-05	**-0.0033*****	**1.1 E-04*****	0.0030
	t^{LN}_{RB}	-0.7627	0.6849	-4.0656	-3.2050	
EP6	$\hat{\gamma}_{LN}$	0.0049	**-1.5 E-05****	**-0.0013*****	**-3.0 E-04****	0.0017
	t^{LN}_{RB}	0.3346	-2.5888	-3.1125	-2.2132	

In the momentum portfolios case, the Fama-MacBeth results are similar to those in SBM25 sample. Both unscaled terms are associated with significant risk price, while scaled market and scaled human capital factors are not priced. The inconsistency between Fama-MacBeth and rolling-beta regressions continues to be evident here as well. In the rolling-beta regressions on momentum portfolios, only risk prices on terms related to the human capital factor, both scaled and unscaled, are significant with much larger scale than their Fama-MacBeth counterparts. The market factor is no longer significant in determining asset returns although the paired difference test result shows that estimates in two approaches are not significantly different.

The difference between the two estimation approaches is more dramatic in the EP6 sample. Despite the fact that only the risk price on the scaled human capital factor in Fama-MacBeth regressions is not significant, the paired difference test results show that none of the four estimated risk prices are the same in the two approaches, with the slope on $z^w R^w$ having opposite signs in the two procedures.

The impact on the overall model performance measured by $RMSE$ is mixed. While the two approaches give similar $RMSE$ in the EP6 sample, $RMSE$ from the rolling-beta regressions is 70% larger in the M10 case. Figure 4.3 reports the average alphas for 10 individual momentum portfolios for both Fama-MacBeth regressions and our three-stage regressions. In both estimation approaches, portfolios with larger prior (2-12 months) return generally have larger average alpha. The difference in $RMSE$ in two approaches mainly comes from portfolios in the first, fourth, seventh, and last prior return deciles. For these portfolios, the 3-stage regression alphas are more than double the magnitude of their counterparts from Fama-MacBeth regressions. We have not been able to identify reasons for the three-stage regression model underperformance.

Figure 4.3 Portfolio alphas of 10 momentum portfolios (M10).
This figure reports individual portfolio alphas of 10 momentum portfolios (M10) for both Fama-MacBeth regressions and the 3-stage regressions. The sample period is February 1959 to December 2013. The solid line is for individual portfolio alphas produced by the 3-stage regressions, and the dash line is for Fama-MacBeth regressions estimated portfolio alphas.



In panel B of table 4.5, we further impose LN's constraint on risk prices. In both M10 and EP6, $\hat{\gamma}_{LN}$'s on R^w and $z^w R^w$ are statistically larger than corresponding $\hat{\gamma}_{RB}$'s, and the same holds in the case of the slope associated with the scaled market factor for EP6. RMSE also increases drastically (50% for M10 and 240% for EP6) after imposing LN's constraint. Again, these results are in-line with results in SBM25 and IND30 samples.

Given our results in SBM25 and IND30 samples as well as those for M10 and EP6 portfolios, we conclude that the good performance of the labor income ratio model in describing the cross section of average stock returns is at least partially due to neglecting certain unconditional constraints, which is along the lines of LN's critique of cross-sectional tests. Moreover, Fama-MacBeth procedure may not be an appropriate way to test conditional models with conditioned-down beta representation represented by equation (9).

Although not closely related to the unconditional constraints (7) as well as those on risk prices suggested by LN, the inconsistency in risk price estimates across different samples cannot be ignored. Intuitively, these estimates are only related to either the factor risk premium or the product of instrument and factor risk premium. If we use the same instruments, the estimates of risk price should be independent of the test assets used assuming the excess return spaces we use in the tests are complete. If the model is correct, we should expect the risk price estimates to be stable across different test assets. In our case, only the market factor has a stable estimated risk premium. Estimates of risk price on $z^M R^M$,and $z^w R^w$ vary a lot from sample to sample in both approaches. This phenomenon may be a signal of potential model misspecification.

4.5 Conclusion.

We study the validity and the impact of unconditional constraints on factor risk prices suggested by LN with respect to the estimation and evaluation of conditional linear factor models. After reviewing the conditioning down process, we argue that LN's critique only applies to conditional models with time-varying betas. Moreover, we identify an extra constraint on unconditional betas in the time-varying beta models specified by a functional relation between conditional and unconditional betas and develop a three-stage test procedure to incorporate this unconditional constraint into model estimation.

With unconditional constraints on risk prices and unconditional betas imposed, we test the labor income ratio model of Santos and Veronesi (2006) on both 25 Fama-French size and B/M portfolios (SBM25) and 30 Fama-French value-weighted industry portfolios (IND30). Our results suggest that, for labor income ratio model, the constraint on unconditional betas has significant effect on both estimates of unconditional betas and risk prices. Further, imposing the unconditional constraints on risk prices for SBM25 portfolios has significant impact on estimated risk premiums.

A4.1 Estimation of monthly log real personal consumption expenditure, c_t.

We estimate the monthly log real personal consumption expenditure (PCE) on nondurable goods (NG) and services (SV) using data series of nominal PCE on NG and SV and corresponding chain-type price index (CF_t^{NG} for NG and CF_t^{SV} for SV). The nominal PCE data series are reported in NIPA table 2.8.5[47] and the data series of chain-type indexes are reported in NIPA table 2.8.4[48] on Bureau of Economic Analysis (BEA)'s website. Since the chain-type price index is a recursive product of the Fisher price index:

$$CF_t = CF_{t-1} \times F_t,$$

where F_t is the Fisher price index at period t, F_t can be obtained by dividing CF_t by CF_{t-1}. With the data series of Fisher price index for both NG and SV, we further assume that $\frac{p_t^{ND}}{p_{t-1}^{ND}} \approx F_t^{ND}, \frac{p_t^{SV}}{p_{t-1}^{SV}} \approx F_t^{SV}$, and $p_0^{ND}: p_0^{SV} = 1: n$, where p_t is the price level at period t, n is an arbitrary number. Together with the fact that $\frac{p_t^{ND} q_t^{ND}}{p_{t-1}^{ND} q_{t-1}^{ND}} = \frac{ND_t}{ND_{t-1}}$ and $\frac{p_t^{SV} q_t^{SV}}{p_{t-1}^{SV} q_{t-1}^{SV}} = \frac{SV_t}{SV_{t-1}}$, we get the Fisher price index for PCE on NG plus SV

[47] http://www.bea.gov/iTable/iTable.cfm?ReqID=9&step=1#reqid=9&step=3&isuri=1&904=1959&903=82&906=q&905=2014&910=x&911=0
[48] http://www.bea.gov/iTable/iTable.cfm?ReqID=9&step=1#reqid=9&step=3&isuri=1&904=2013&903=81&906=q&905=2014&910=x&911=1

$$F_t^T = \sqrt{F_t^{SV} F_t^{ND} \frac{(F_t^{ND} ND_{t-1} + F_t^{SV} SV_{t-1})}{(F_t^{ND} SV_t + F_t^{SV} ND_t)} \frac{ND_t + SV_t}{ND_{t-1} + SV_{t-1}}}.$$

Then, the log real PCE on NG plus SV, c_t, and its growth rate, Δc_t, are

$$c_t = \ln\left\{(NG_t + SV_t)\frac{CF_t^T}{CF_{July,09}^T}\right\},$$

$$\Delta c_t = ln\left(\frac{ND_t + SV_t}{ND_{t-1} + SV_{t-1}} / F_t^T\right),$$

where $CF_{July,09}^T$ is the chain-type price index of July, 2009, which is assumed to be equal to 100.

A4.2 Estimation of monthly log labor income, y_t.

Following Lettau and Ludvigson (2001a), we construct monthly real y_t based on data on personal income reported in NIPA table 2.6 from BEA website. The nominal labor income is the sum of wages and salaries, supplements to wages and salaries, personal current transfer receipts minus contributions to 'government social insurance domestic' and taxes. Taxes are computed as personal current taxes multiplied by the ratio between wages and salaries and the sum of all personal incomes including wages and salaries, proprietors' income with inventory valuation and capital consumption adjustment, rental income of persons with capital consumption adjustment, and personal income receipts on assets.

To get the log real labor income, y_t, we deflate the nominal labor income with CF_t^T described in A4.1 and take the natural logarithm of the result.

A4.3 Robustness Checks.

A4.3.1 Quarterly Results.

As a further robustness check, we estimate the labor income ratio model at a quarterly data frequency to see whether it affects test results. We use quarterly consumption data and quarterly labor income from Martin Lettau's website. For this exercise, the data sample spans the period from 1952:2 to 2012:3. Since quarterly data results are in general in line with monthly data results, to save space we only report estimates for SBM25 and IND30 portfolios[49].

Results for the adjusted R-square reported in Figure A4.3.1 demonstrate that, at quarterly frequency, the instrument, labor income to consumption ratio, again has little power in explaining the variation of factor betas. With both SBM25 and IND30, most of the adjusted R-squares are below 10% for both factor betas.

Figure A4.3.1 Adjusted R-square of Regressions $\beta_t^i = \beta_0^i + \delta^i s_t^w + e_t^i$.
This figure reports the adjusted R-squares of regressions $\beta_t^i = \beta_0^i + \delta^i s_t^w + e_t^i$ of the second stage of the rolling-beta regressions. Test returns are quarterly real excess returns on 25 size and book to market portfolios (SBM25), and 30 industry portfolios (IND30) for figure 4.3.1.1, and 4.3.1.2, respectively. The sample period is 1952:2 to 2012:3. The solid line is for the adjusted R-squares of regressions of market betas, β_t^M, and the dash line is for those of regressions of the human capital betas, β_t^w.

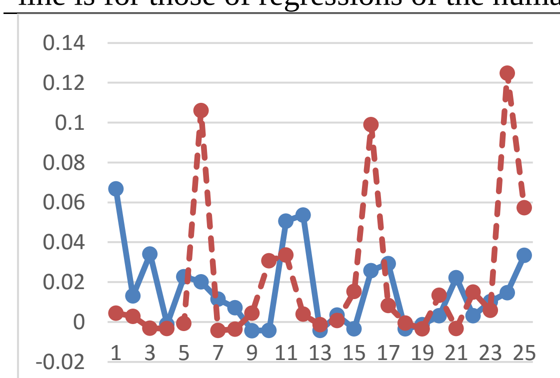


Figure 4.3.1.1 SBM25

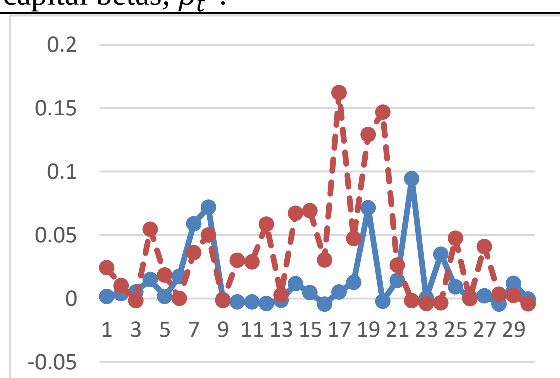


Figure 4.3.1.2 IND30

[49] Estimation results for M10 and EP6 portfolios are available upon request.

In Table A4.3.1, Fama-MacBeth estimates of factor betas remain inconsistent with rolling-beta estimates for more than 50% of the portfolios in SBM25 group at quarterly frequency as well. The results for IND30 group are slightly weaker, but still, for betas of scaled factors, Fama-MacBeth estimates are different from rolling-beta estimates for most of the portfolios in the group.

Table A4.3.1 Empirical results of labor income ratio model.
This table reports the number of portfolios of which the Fama-MacBeth estimated factor beta is outside the 95% confidence interval of the corresponding estimate from rolling-beta regressions. The tested model is the labor income ratio model. Test returns are quarterly real excess returns on 25 size and book to market portfolios (SBM25) and 30 industry portfolios (IND30). The sample period is 1952:2 to 2012:3.

	β^{M}	β^{zM}	β^{w}	β^{zw}
SBM25	13	20	15	25
IND30	7	25	12	30

Regarding estimated risk prices reported in table A4.3.2, the differences between Fama-MacBeth estimates and rolling-beta estimates for SBM25 portfolios are less significant in the quarterly sample than those in the monthly sample. However, the rolling-beta estimates of slopes associated with $z^{M}R^{M}$ and R^{w} are still much smaller than their counterparts in the Fama-MacBeth regressions. If we compare $\hat{\gamma}_{RB}$'s with $\hat{\gamma}_{LN}$'s, we find that the rolling-beta regression estimate for the risk price on $z^{w}R^{w}$ is significantly lower than the corresponding value suggested by LN. Also, according to *RMSE* estimates reported in the last column of table A4.3.2, *RMSE* increases by roughly 60% after we impose the constraint on risk prices suggested by LN, thus supporting LN's conclusions.

Table A4.3.2 Empirical results of labor income ratio model.

This table reports the main results of our empirical tests on labor income model. The test returns are quarterly real excess returns on 25 size and book to market portfolios (SBM25) and 30 industry portfolios (IND30). The sample period is 1952:2 to 2012:3. Panel A presents the estimation of risk prices on the unscaled and scaled market factor, R^M and $z^M R^M$, and the unscaled and scaled labor income factor, R^w and $z^w R^w$, from both Fama-MacBeth regressions, $\hat{\gamma}_{FM}$, and rolling-beta regressions, $\hat{\gamma}_{RB}$. *t*-stats with Newey-West adjustment for autocorrelation (*lag* = 3) are reported in parentheses below the point estimates. $RMSE$ is the root mean squared error of the model. Paired difference test in the third section of panel A tests the hypothesis $\boldsymbol{H_0}: \hat{\gamma}_{RB} - \hat{\gamma}_{FM} = 0$. t_{RB-FM} is computed as $t_{RB-FM} = \sqrt{T}\frac{\hat{\gamma}_{RB}-\hat{\gamma}_{FM}}{\sqrt{\hat{\sigma}^2_{\gamma_{RB,t}}+\hat{\sigma}^2_{\gamma_{FM,t}}-2\widehat{cov}(\gamma_{RB,t},\gamma_{FM,t})}}$, where $\hat{\sigma}^2_{\gamma_{RB,t}}$, $\hat{\sigma}^2_{\gamma_{FM,t}}$, and $\widehat{cov}(\gamma_{RB,t},\gamma_{FM,t})$ are estimated from the time series of $\hat{\gamma}_{RB,t}$ and $\hat{\gamma}_{FM,t}$ with Newey-West adjustment. Panel B presents the results of testing hypothesis (H.1) and (H.2). $\hat{\gamma}_{LN}$ is the estimate of slopes on β_{i0} and δ_i suggested by equation (9). $t^{LN}_{RB} = \frac{\hat{\gamma}_{RB}-\hat{\gamma}_{LN}}{\sqrt{\hat{\sigma}^2_{\gamma_{RB,t}}/T}}$. $RMSE_{LN}$ is the root mean squared error of the model after imposing constraint (9).

Panel A						
Fama-MacBeth procedure						
		R^M	$z^M R^M$	R^w	$z^w R^w$	*RMSE*
$\hat{\gamma}_{FM}$	SBM25	**0.0203*****	**0.0022****	-0.0076	**-5.3 E-04*****	0.0054
		(3.8393)	(2.0086)	(-1.5632)	(-3.8020)	
	IND30	**0.0182*****	-6.6448 E-04	-8.7486 E-04	**2.5714 E-04****	0.0048
		(3.3609)	(-1.06)	(-0.4051)	(2.5853)	
Rolling-beta regressions						
$\hat{\gamma}_{RB}$	SBM25	**0.0207*****	8.0241 E-04	-0.0027	**-5.4 E-04*****	0.0048
		(3.8808)	(0.3469)	(-0.9895)	(-3.4513)	
	IND30	**0.0182*****	-0.0011	-0.0018	1.0134 E-04	0.0054
		(3.3587)	(-0.6778)	(-1.0466)	(1.2050)	
Paired Difference Test ($\boldsymbol{H_0}: \hat{\gamma}_{RB} - \hat{\gamma}_{FM} = 0$)						
t_{RB-FM}	SBM25	0.6211	-0.5968	**1.8714***	-0.0701	
	IND30	0.0779	-0.2850	-0.9981	-2.3821	
Panel B: LN's critique						
SMB25	$\hat{\gamma}_{LN}$	0.019	2.5013 E-04	-1.3639 E-04	**9.3 E-05*****	0.0074
	t^{LN}_{RB}	0.3223	0.2388	-0.9387	-4.0460	
IND30	$\hat{\gamma}_{LN}$	0.0162	1.413 E-04	3.2477 E-04	2.8044 E-05	0.0062
	t^{LN}_{RB}	0.3654	-0.7622	-1.2315	0.8715	

In the case of IND30, Fama-MacBeth estimate of risk price on $z^w R^w$ is statistically smaller than the corresponding rolling-beta estimate. While further imposing LN constraints (H.1) and (H.2) on risk prices does not change the risk price estimates significantly, it deteriorates the overall model performance illustrated by the 15% increase in *RMSE*.

Overall, the quarterly sample results are weaker but still consistent with monthly sample results. The discrepancy in results at different frequencies may be due to second moments being estimated more precisely from higher frequency data.

A4.3.2 Bootstrap Execrise.

In both empirical tests we conduct, we use multi-stage regressions, which are potentially subject to the generated regressor problem. As shown by Vassalou (2003), treating results of previous stage regressions as given in the following stages can potentially bias not only estimated standard errors but also point estimates of variables of interest. There are a few ways to mitigate the generated regressor problem. One approach is to estimate the model using a single-stage method such as a single-stage GMM. Applying GMM requires the ratio between the number of moment conditions and the length of the data series to be small. When the ratio is large (>10% as a rule of thumb), the estimate of the moment covariance matrix deteriorates and GMM estimates become unreliable. In our case, we need to estimate time-varying factor betas in the first-stage using OLS regressions. Each of these OLS time-varying factor beta estimates requires two moment conditions if estimated using GMM. Together with the fact that we have $(T - N)$ betas to estimate, it leads to $2(T - N)$ moment conditions for each factor, where N is the length of the estimation window. Since N cannot be too large, it's virtually impossible to satisfy the GMM requirement even for the first-stage regressions. A way to check the severity of the generated regressor problem is to run a Monte-Carlo simulation. However, to generate artificial data series in the simulation exercise, thorough knowledge of the underlying model, including the functional form of the stochastic discount factor (SDF), is needed. Unfortunately, the SDF cannot be derived from the beta representation, for factor betas contain no information on the actual structure of the SDF. Another way is using a bootstrap method. The difficulty of applying bootstrap to our case lies in the first-stage regressions since, in our rolling-window regressions, the sequence of the observations is crucial. To bootstrap factor betas, we can either resample the whole data sample or do it for betas in each period. The problem with the first approach is that, when estimating betas at period t, we may mistakenly use observations which are not originally in the estimation period $t-N$ to $t-1$. If we bootstrap factor betas in each period separately, in the third stage of rolling-beta regressions, we may run into a problem of forming an inconsistent data sample of portfolio returns.[50]

Given the discussion above, there may be no easy approach to check the generated regressor problem for all three stages. However, it's still possible to bootstrap the second and the third stage regressions. In our bootstrap exercise, we resample the estimated factor betas from the first stage with replacement, and then run the second and third stage regressions on the bootstrap sample. With the total number of paired-bootstrap samples of 5000, we get empirical distributions of factor risk premiums estimated from

[50] For instance, if we bootstrap estimation period of β_t and that of β_{t-1}, it is possible that the y_{t-N} to y_{t-2} we used to estimate β_t are different from those for β_{t-1}. The problem here is which y_{t-N} to y_{t-2} should be used in the final stage when estimating risk prices. We cannot find an easy solution to this problem at this time.

all four test asset groups. Bootstrap results for SBM25 and IND30 portfolios are reported in Figure A4.3.2 and Figure A4.3.3, respectively.

Comparing bootstrap results in figure A4.2 with the original estimates reported in panel A of table 4.3, we can see that our rolling-beta regression estimates are in general consistent with the bootstrap results. Moreover, if we compare FM estimates with our bootstrap estimates, we get similar results to those reported in the first section of table 4.3, panel A. Estimated risk premiums on market factor and $z^w R^w$ in two different approaches are not statistically different from each other. FM estimate of risk premium on $z^M R^M$ and that on R^w are outside the 95% confidence band of the corresponding bootstrap estimates, which is consistent with our results in table 4.3.

In the case of IND30 reported in figure A4.3.3, again the bootstrap results are in line with estimated results of the rolling-beta regressions in panel A of table 4.3. Unlike in the case of SBM25, FM estimates reported in table 4.3 are also all within the 95% confidence band of the corresponding bootstrap estimate.

Likewise, for M10 and EP6 portfolios (bootstrap results not reported here[51]) rolling-beta regression estimates are all consistent with their counterparts obtained from the bootstrap analysis.

Overall, our bootstrap results above suggest that our rolling-window regression is robust to the generated regressor problem in the second and third stages.

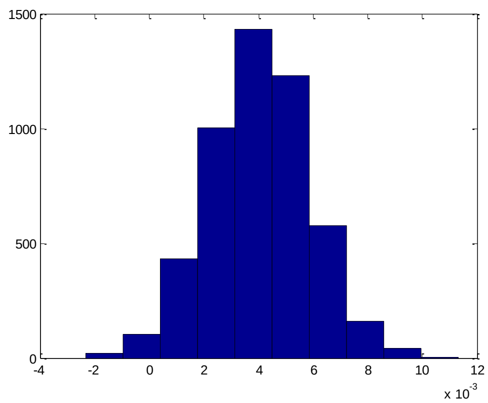
[51] Available upon request.

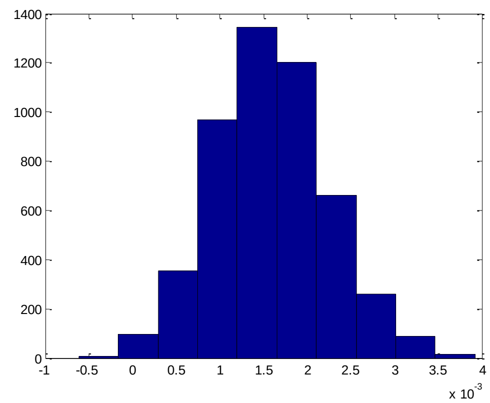
Figure A4.3.2 Empirical distribution of estimated factor risk premiums (SBM25).

The figure reports the empirical distribution of estimated factor risk premiums, $\hat{\gamma}_{R^M}, \hat{\gamma}_{z^M R^M}, \hat{\gamma}_{R^w}, \hat{\gamma}_{z^w R^w}$, based on paired bootstrap samples. Test returns are monthly real excess returns on 25 size and book to market portfolios (SBM25). The sample period is February 1959 to December 2013. The number of bootstrap samples is 5000. Panel B contains 3-stage regression estimates of factor risk premiums $\hat{\gamma}_{RB}$ originally reported in the second section of Panel A in Table 3.

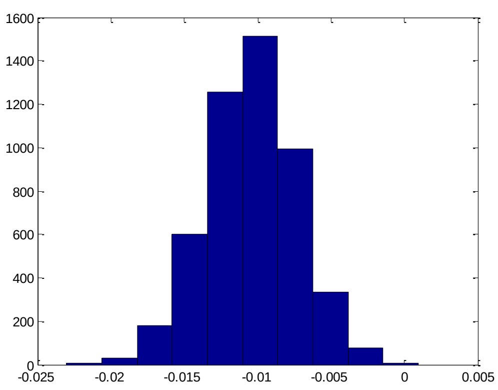
Panel A

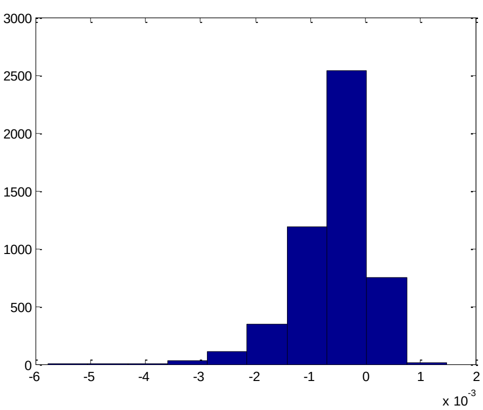
$\hat{\gamma}_{R^M}$	*Median*	*Mean*	*Std. Dev.*	$\hat{\gamma}_{z^M R^M}$	*Median*	*Mean*	*Std. Dev.*
	0.0042	0.004	0.0018		0.0016	0.0016	0.0007
	95% Confidence Band				*95% Confidence Band*		
	[0.001028, 0.007038]				[0.000558, 0.002688]		





$\hat{\gamma}_{R^w}$	*Median*	*Mean*	*Std. Dev.*	$\hat{\gamma}_{z^w R^w}$	*Median*	*Mean*	*Std. Dev.*
	-0.0104	-0.0104	0.0031		-0.0004	-0.0006	0.0007
	95% Confidence Band				*95% Confidence Band*		
	[-0.0156, -0.0054]				[-0.00183, 0.000219]		





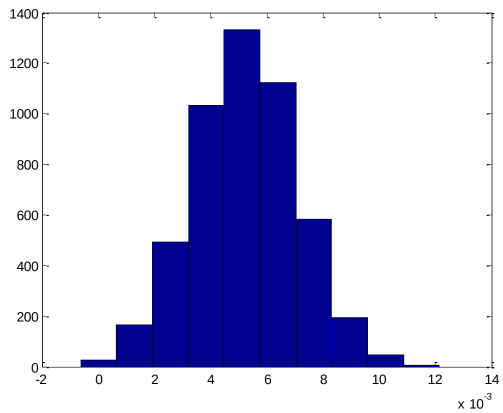
Panel B

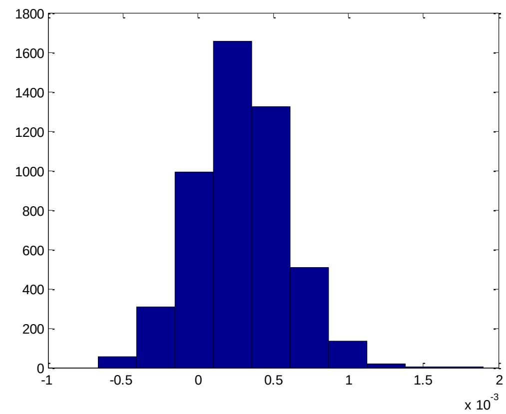
	R^M	$z^M R^M$	R^w	$z^w R^w$
$\hat{\gamma}_{RB}$	**0.004****	**0.0016*****	**-0.0107*****	-5.7 E-04

Figure A4.3.3 Empirical distribution of estimated factor risk premiums (IND30).
The figure reports the empirical distribution of estimated factor risk premiums, $\hat{\gamma}_{R^M}, \hat{\gamma}_{z^M R^M}, \hat{\gamma}_{R^w}, \hat{\gamma}_{z^w R^w}$, based on paired bootstrap samples. Test assets are real excess returns on 30 industry portfolios (IND30). The sample period is February 1959 to December 2013. The number of bootstrap samples is 5000. Panel B contains 3-stage regression estimates of factor risk premiums $\hat{\gamma}_{RB}$ originally reported in the second section of Panel A in Table 3.

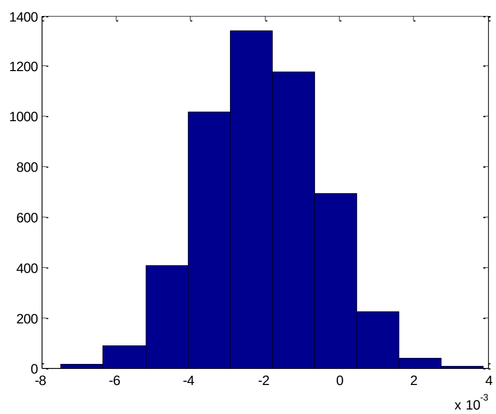
Panel A

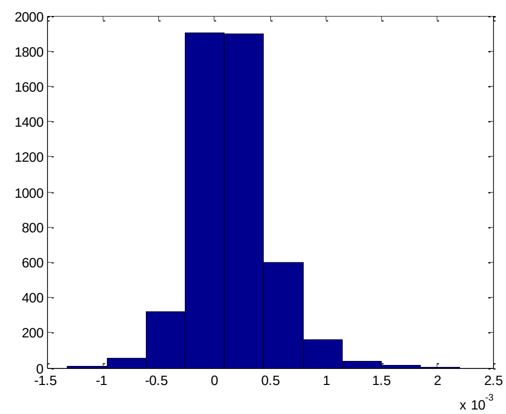
$\hat{\gamma}_{R^M}$	*Median*	*Mean*	*Std. Dev.*	$\hat{\gamma}_{z^M R^M}$	*Median*	*Mean*	*Std. Dev.*
	0.0052	0.0052	0.0019		0.0003	0.0003	0.0003
	95% Confidence Band				*95% Confidence Band*		
	[0.00217, 0.00830]				[-0.00021, 0.000787]		





$\hat{\gamma}_{R^w}$	*Median*	*Mean*	*Std. Dev.*	$\hat{\gamma}_{z^w R^w}$	*Median*	*Mean*	*Std. Dev.*
	-0.0021	-0.0021	0.0016		0.0001	0.0002	0.0003
	95% Confidence Band				*95% Confidence Band*		
	[-0.00466, 0.000531]				[-0.00033, 0.000754]		





Panel B

	R^M	$z^M R^M$	R^w	$z^w R^w$
$\hat{\gamma}_{RB}$	**0.0052*****	2.8 E-04	-0.0021	1..6 E-04

A4.4 The scaled CCAPM.

Given the discussion on LN's critique in the introduction section above, we argue that LN's critique is not applicable to conditional SDF models. LN suggest that, when estimating risk premiums, the unconditional constraints on the risk premiums are neglected. A natural way to check the validity of LN's critique is to estimate the model both with and without the unconditional constraints and see if two sets of estimated risk premiums are consistent.

To incorporate the unconditional constraints, we estimate the CCAPM model using the first-stage GMM procedure, with equation (5) providing a set of equally weighted moment conditions. The benefit of the GMM procedure is that we can impose all possible constraints implied by the fundamental pricing equation (1). Also, it allows us to estimate ϕ_0^f, ϕ^f in equation (4) and factor risk premiums λ in equation (6) simultaneously. When test assets are excess returns, as Cochrane (2005) points out, the first-stage GMM estimation is equivalent to running a cross-sectional regression of average excess returns on the average product of the excess returns and risk factors. In order to get unique estimates, we assume that the first element of ϕ_0^f, $\phi_{0,1}^f$, is 1, then we have

$$\hat{\phi} = \left[\widehat{\phi_{0,-1}^f}', \ \widehat{\phi}^{f\prime}\ \right]^T = -(d'd)^{-1}d'E(R_t), \tag{11}$$

$$d = E(R_t f_t'), \tag{12}$$

$$\lambda = E(f_t^{\#}) - \frac{[1, \hat{\phi}']}{E([1, \hat{\phi}'] f_t^{\#})} var(f_t^{\#}), \tag{13}$$

$$\beta^i = (f^{\#\prime} f^{\#})^{-1}(f^{\#\prime} R^i), \tag{14}$$

where $\phi_{0,-1}^f$ is ϕ_0^f without the first element, and β^i is a vector of factor betas for asset *i*.

Here too, a side effect of using the CCAPM as an example is the generated regressor problem in parameter estimates. The problem is due to the instrument in the CCAPM, *cay*, being an estimated value (see Lettau and Ludvigson (2001a) for details).

Table A4.4.1 presents both Fama-MacBeth estimates and first-stage GMM estimates of risk premium on cay_{t-1}, log consumption growth, Δc_t, and the cross-term, $\Delta c_t * cay_{t-1}$, for SBM25. In case of Fama-MacBeth regressions (Panel A), all three estimates of risk premium are positive and statistically significant. The estimated risk premium on $\Delta c_t * cay_{t-1}$ is considerably smaller in scale compared to the other two estimates. This is because the cross-term itself is relatively small. The results for the first-stage GMM procedure are displayed in Panel B. SDF has negative loadings on all three factors. Again, because the cross-term is small in magnitude, loading on it is large relative to those on Δc_t and cay_{t-1}. But *t*-statistics (reported below the point estimates) show that none of these loadings are statistically significant. Regarding risk premiums, the GMM point estimates are the same as the corresponding estimates from Fama-MacBeth regressions. Unlike their counterparts in Panel A, all three estimated risk premiums are insignificant. This discrepancy is likely due to the Fama-MacBeth estimates being subject to the generated regressor problem, which is exacerbated by the fact that the instrument is not a return. Conditioning on point estimates from the first stage leads to understated sampling variation of parameter estimates in the second stage ultimately inflating t-statistics. Moreover, $RMSEs$ in the two estimation approaches are also very similar. Since we estimate the betas using the whole sample in Fama-MacBeth regressions and the estimated risk prices are the same in the two approaches as reported in Table A4.4.1, the unconditional pricing errors (alphas) should be the same in the two approaches as well.

Table A4.4.1 Empirical results of the scaled CCAPM (SBM25).

This table reports the main results of our empirical tests of CCAPM. Test returns are quarterly real excess returns on SBM25. The sample period is 1952:2 to 2012:3. Panel A reports the estimations of factor risk premiums, $\hat{\lambda}$, and overall model performance measure, $RMSE$, from the Fama-MacBeth regressions. Panel B presents the estimation of SDF's factor loadings $\hat{b}$, factor risk premiums, $\hat{\lambda}$, and overall model performance measure, $RMSE$, from the first-stage GMM procedure. *t*-stat's reported in the parentheses below the corresponding point estimates are not corrected for autocorrelation. $RMSE$ is computed as $RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\hat{\alpha}_i^2}$, where $\hat{\alpha}_i, i = 1, \dots, N$ is average pricing error for asset i.

Panel A: Fama-MacBeth procedure			
	cay_{t-1}	Δc_t	$\Delta c_t * cay_{t-1}$
$\hat{\lambda}_{FM}$	**0.0156****	**0.0045****	**0.000091****
	(2.5794)	(2.0655)	(2.1177)
$RMSE$	0.0059		
Panel B: first-stage GMM			
$\hat{b}_{GMM}$	-15.6024	-226.8431	-11960.4448
	(-0.2971)	(-1.5979)	(-0.9053)
$\hat{\lambda}_{GMM}$	0.0156	0.0045	0.000091
	(1.6049)	(1.3289)	(1.2420)
$RMSE$	0.0059		

In the test with IND30 (Table A4.4.2), we observe the same consistency in the estimates of factor risk premiums across the two estimation approaches. Again, none of the GMM estimates of risk premiums are significant as reported in Panel B while their counterparts in Panel A are all statistically different from zero. $RMSE$ in Fama-MacBeth procedure (0.006) equals the corresponding value in the first-stage GMM procedure.

Table A4.4.2 Empirical results of the scaled CCAPM (IND30).

This table reports the main results of our empirical tests of CCAPM. Test returns are quarterly real excess returns on IND30. The sample period is 1952:2 to 2012:3. Panel A reports estimates of factor risk premiums, $\hat{\lambda}$, and overall model performance measure, $RMSE$, from Fama-MacBeth regressions. Panel B presents estimates of SDF's factor loadings, $\hat{b}$, factor risk premiums, $\hat{\lambda}$, and overall model performance measure, $RMSE$, from the first-stage GMM procedure. t-stat's reported in the parentheses below the corresponding point estimates are not corrected for autocorrelation. $RMSE$ is computed as $RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\hat{\alpha}_i^2}$, where $\hat{\alpha}_i, i = 1, \dots, N$ is average pricing error for asset *i*.

Panel A: Fama-MacBeth procedure			
	cay_{t-1}	Δc_t	$\Delta c_t * cay_{t-1}$
$\hat{\lambda}_{FM}$	**0.0097****	**0.0045*****	**-0.000051****
	(2.1812)	(2.7884)	(-2.2596)
$RMSE$	0.0060		
Panel B: First-stage GMM			
$\hat{b}_{GMM}$	-73.54	-162.49	14355.82
	(-1.4004)	(-1.4711)	(1.8875)
$\hat{\lambda}_{GMM}$	0.0097	0.0045	-0.000051
	(0.8875)	(1.7152)	(-1.3665)
$RMSE$	0.0060		

LN's critique does not apply to the scaled CCAPM because CCAPM is a beta representation corresponding to unconditional SDF representation (5) rather than conditional beta representation (6) conditioned down to the econometrician's information set. The former does not lead to additional economic constraints on regression slopes and, naturally, Fama-MacBeth estimates are consistent with GMM estimates as Tables A4.4.1 and A4.4.2 demonstrate.

Chapter Five: General Conclusion.

This thesis has studied asset pricing implications of incomplete information on fundamentals in LRR models and unconditional constraints on conditional affine factor asset pricing models, both theoretically and empirically. While contributions of each essay have been summarized in their respective conclusions, here we provide a brief summary.

Essay one extends the theoretical work on the LRR model. Unlike in previous studies, we assume that investors do not observe the LRR component directly, but can learn it by building a minimum mean square estimate from observations on state variables (consumption and dividend growth). Under this assumption, investor's optimal estimate of the LRR component is a linear combination of state variables. With the filtered LRR component process in hand, we derive the endogenous equilibrium pricing kernel under a new risk-neutral measure corresponding to information available to investors. Since the pricing kernel is affected by learning, the estimation risk of the unobserved LRR component affects all risk prices in the pricing kernel and, consequently, all contributions to the equity premium. Determined by the drift of the pricing kernel, the short rate in our model is also endogenous. It is an affine function of the estimated conditional mean of consumption growth and, thus, inherits its mean-reverting property. We demonstrate our analytical results by applying the model to a profitability-based equity valuation model proposed by Pastor and Veronesi (2003). In the combined model, the equity premium has two components: the component corresponding to risk in profitability and the contribution of discount rate risk. Calibration of the model demonstrates that the LRR model with learning has the potential to fit levels of price-dividend ratios of the S&P 500 Composite Index, equity premium, and the short term interest rate simultaneously. More importantly, information uncertainty reduces the total equity premium. Higher estimation risk (due to lower correlation between state and latent variables) implies a smaller component of the total risk premium which is due to the profitability risk. At the same time, it also leads to larger (in absolute value) discount risk component of the risk premium. Since the discount risk premium component is negative (due to the negative elasticity of the price-dividend ratio to the short rate), the total equity premium becomes smaller as learning becomes less effective.

The second essay contributes to the finance literature along the following dimensions. Theoretically, this is the first study on valuation of nominal term structure of interest rate in a LRR economy with incomplete information on fundamentals. First, on the empirical side, we propose a Bayesian MCMC method, which allows us to estimate not only latent variables and parameters in variable processes simultaneously, but also conditional correlations between latent and state variables as well. These correlations are crucial in the sense that they jointly determine the short rate, risk premium, and the learning efficiency in our model. With estimated values of these correlations, the model-implied risk premium turns out to be negative. Second, unlike previous literature on LRR term structure, we estimate the model parameters and latent variables using only information in the sample of state variables. In this way, parameters and latent variables are estimated outside the model, which imposes the tension between the time series and cross sectional properties of the model and makes empirical tests of pricing implications of the model more stringent. Third, we provide empirical evidence contrary to the base case LRR model (without economic uncertainty and estimation risk). In particular, for the base case model, we show that the model fails to produce an upward-sloping yield curve, which is a direct consequence of the negative risk premium. Finally, we study the impact of learning on the shape of the nominal yield curve. Given estimated parameter and latent variable values, we find that the estimation risk of the LRR component changes the conditional correlation between the LRR component and all other variables in the model, and therefore reduces both the model-implied short rate and the risk premium (in absolute value). Combining these two effects, learning shifts the nominal yield curve downward and makes it flatter. However, the overall effect is insufficient to produce an upward-sloping yield curve.

In the third essay, we identify an unconditional constraint on unconditional betas for time-varying beta models. The constraint is a result of a model assumption about the functional relation between time-varying betas, instruments, and unconditional betas. Sometimes, this constraint may not be imposed in model estimation. For example, in the Fama-MacBeth procedure, this constraint cannot be incorporated. To incorporate this constraint in model estimation, we develop a three-stage estimation procedure. We also study the constraints on risk prices proposed by Lewellen and Nagel (2006). Following Ludvigson (2012), we show that Lewellen and Nagel's constraints are derived from the time-varying beta assumption, and therefore can only be applied to the time-varying beta models. To demonstrate the economic impact of the restrictions, we test our constraint on unconditional betas and Lewellen and Nagel's constraints on risk prices in the labor income ratio model of Santos and Veronesi (2006) using both the 25 Fama-French size and B/M portfolios (SBM25) and the 30 Fama-French value-weighted industry portfolios (30IND). Our results suggest that, for the labor income ratio model, the constraint on unconditional betas has a significant effect on the estimates of both unconditional betas and risk prices. Further, in the SBM25 sample, imposing the unconditional constraints on risk prices has a significant impact on estimated risk premiums.

Bibliography

Acharya, Viral V., and Pedersen, Lasse Heje. "Asset pricing with liquidity risk." *Journal of Financial Economics* 77, no. 2 (2005): 375-410.

Adrian, Tobias, and Franzoni, Francesco. "Learning about beta: Time-varying factor loadings, expected returns, and the conditional CAPM." *Journal of Empirical Finance* 16, no. 4 (2009): 537-556.

Ai, Hengjie. "Information Quality and Long-Run Risk: Asset Pricing Implications." *The Journal of Finance* 65, no. 4 (2010): 1333-1367.

Aitkin, Murray. "Posterior bayes factors." *Journal of the Royal Statistical Society. Series B (Methodological)* (1991): 111-142.

Aldrich, Eric M., and Gallant, A. Ronald. "Habit, long-run risks, prospect? a statistical inquiry." *Journal of Financial Econometrics* 9, no. 4 (2011): 589-618.

Ang, Andrew, and Chen, Joseph. "CAPM over the long run: 1926–2001."*Journal of Empirical Finance* 14, no. 1 (2007): 1-40.

Ang, Andrew, and Kristensen, Dennis. "Testing conditional factor models."*Journal of Financial Economics* 106, no. 1 (2012): 132-156.

Asness, Clifford S., Moskowitz, Tobias J., and Pedersen, Lasse Heje. "Value and momentum everywhere." *The Journal of Finance* 68, no. 3 (2013): 929-985.

Bai, Jushan, and Wang, Peng. "Identification and Bayesian estimation of dynamic factor models." *Journal of Business & Economic Statistics* 33, no. 2 (2015): 221-240.

Bakshi, Gurdip, and Chen, Zhiwu. "Cash Flow Risk, Discounting Risk, and the Equity Premium Puzzle." *Handbook of Investments: Equity Premium*(2006).

Bansal, Ravi, and Shaliastovich, Ivan. "A long-run risks explanation of predictability puzzles in bond and currency markets." *Review of Financial Studies* (2012): 1-33.

Bansal, Ravi, and Yaron, Amir. "Risks for the long run: A potential resolution of asset pricing puzzles." *The Journal of Finance* 59, no. 4 (2004): 1481-1509.

Bansal, Ravi, Dittmar, Robert, and Kiku, Dana. "Cointegration and consumption risks in asset returns." *Review of Financial Studies* 22, no. 3 (2009): 1343-1375.

Bansal, Ravi, Gallant, A. Ronald, and Tauchen, George. "Rational pessimism, rational exuberance, and asset pricing models." *The Review of Economic Studies* 74, no. 4 (2007): 1005-1033.

Bansal, Ravi, Khatchatrian, Varoujan and Yaron, Amir. "Interpretable asset markets?." *European Economic Review* 49, no. 3 (2005): 531-560.

Bansal, Ravi, Kiku, Dana, and Yaron, Amir. "Long run risks, the macroeconomy, and asset prices." *The American Economic Review* (2010): 542-546.

Bansal, Ravi, Kiku, Dana, and Yaron, Amir. "Risks for the long run: Estimation and inference." Working Paper, Wharton School, University of Pennsylvania, 2008.

Bansal, Ravi, Kiku, Dana, Shaliastovich, Ivan, and Yaron, Amir. "Volatility, the macroeconomy, and asset prices." *The Journal of Finance* 69, no. 6 (2014): 2471-2511.

Beeler, Jason, and Campbell, John Y.. "The long-run risks model and aggregate asset prices: An empirical assessment." *Critical Finance Review* 1, no. 1 (2012): 141-182.

Bollerslev, Tim, Engle, Robert F. and Wooldridge, Jeffrey M. "A capital asset pricing model with time-varying covariances." *The Journal of Political Economy* (1988): 116-131.

Brennan, Michael J., and Xia, Yihong. "Stock price volatility and equity premium." *Journal of monetary Economics* 47, no. 2 (2001): 249-283.

Campbell, John Y., and Cochrane, John H. *Explaining the poor performance of consumption-based asset pricing models*. No. w7237. National bureau of economic research, 1999.

Campbell, John Y., and Viceira, Luis M. *Strategic asset allocation: portfolio choice for long-term investors*. New York: Oxford University Press, 2002.

Carlin, Bradley P., Polson, Nicholas G., and Stoffer, David S. "A Monte Carlo approach to nonnormal and nonlinear state-space modeling." *Journal of the American Statistical Association* 87, no. 418 (1992): 493-500.

Carter, Chris K., and Kohn, Robert. "On Gibbs sampling for state space models." *Biometrika* 81, no. 3 (1994): 541-553.

Cochrane, John H. “A cross-sectional test of an investment-based asset pricing model.” *Journal of Political Economy* 104, (1996) 572-621.

Cochrane, John H. *Asset Pricing:(Revised Edition)*. Princeton: Princeton university press, 2005.

Constantinides, George M., and Ghosh, Anisha. "Asset pricing tests with long-run risks in consumption growth." *Review of Asset Pricing Studies* 1, no. 1 (2011): 96-136.

Doh, Taeyoung. "Long-Run Risks In The Term Structure Of Interest Rates: Estimation." *Journal of Applied Econometrics* 28, no. 3 (2013): 478-497.

Drechsler, Itamar, and Yaron, Amir. "What's vol got to do with it." *Review of Financial Studies* 24, no. 1 (2011): 1-45.

Duffie, Darrell, and Epstein, Larry G. "Stochastic differential utility."*Econometrica: Journal of the Econometric Society* (1992): 353-394.

Dybvig, Philip H., and Ross, Stephen A. "Differential information and performance measurement using a security market line." *The Journal of finance* 40, no. 2 (1985): 383-399.

Epstein, Larry G., and Zin, Stanley E. "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework." *Econometrica: Journal of the Econometric Society* (1989): 937-969.

Eraker, Bjørn. "Affine general equilibrium models." *Management Science* 54, no. 12 (2008): 2068-2080.

Fama, Eugene F., and French, Kenneth R. "Common risk factors in the returns on stocks and bonds." *Journal of financial economics* 33, no. 1 (1993): 3-56.

Fama, Eugene F., and French, Kenneth R. "Dissecting anomalies." *The Journal of Finance* 63, no. 4 (2008): 1653-1678.

Fama, Eugene F., and French, Kenneth R. "Multifactor explanations of asset pricing anomalies." *The journal of finance* 51, no. 1 (1996): 55-84.

Fama, Eugene F., and French, Kenneth R.. "Size, value, and momentum in international stock returns." *Journal of financial economics* 105, no. 3 (2012): 457-472.

Fama, Eugene F., and MacBeth, James D. "Risk, return, and equilibrium: Empirical tests." *The Journal of Political Economy* (1973): 607-636.

Fama, Eugene F., and. French, Kenneth R. "The cross-section of expected stock returns." *the Journal of Finance* 47, no. 2 (1992): 427-465

Ferson, Wayne E., Kandel, Shmuel, and Stambaugh, Robert F. "Tests of Asset Pricing with Time-Varying Expected Risk Premiums and Market Betas." *The Journal of Finance* 42, no. 2 (1987): 201-220.

Ferson, Wayne, Nallareddy, Suresh, and Xie, Biqin. "The “out-of-sample” performance of long run risk models." *Journal of Financial Economics* 107, no. 3 (2013): 537-556.

Fruhwirth-Schnatter, Sylvia. "Bayesian model discrimination and Bayes factors for linear Gaussian state space models." *Journal of the Royal Statistical Society. Series B (Methodological)* (1995): 237-246.

Gelfand, Alan E. "Model determination using sampling-based methods."*Markov chain Monte Carlo in practice*, edited by W.R. Gilks, S. Richardson, and David Spiegelhalter, 145-161. Boca Raton: Chapman & Hall/CRC, 1996.

Gelfand, Alan E., and Smith, Adrian F. M. "Sampling-based approaches to calculating marginal densities." *Journal of the American statistical association* 85, no. 410 (1990): 398-409.

Gelman, Andrew, and Rubin, Donald B. "Inference from iterative simulation using multiple sequences." *Statistical science* (1992): 457-472.

Geman, Stuart, and Geman, Donald. "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 6 (1984): 721-741.

Geweke, John, and Tanizaki, Hisashi. "Bayesian estimation of state-space models using the Metropolis–Hastings algorithm within Gibbs sampling."*Computational Statistics & Data Analysis* 37, no. 2 (2001): 151-170.

Hansen, Lars Peter, Richard, and Scott F. "The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models." *Econometrica: Journal of the Econometric Society* (1987): 587-613.

Hansen, Lars Peter. "Large sample properties of generalized method of moments estimators." *Econometrica: Journal of the Econometric Society*(1982): 1029-1054.

Harvey, Campbell R. "Time-varying conditional covariances in tests of asset pricing models." *Journal of Financial Economics* 24, no. 2 (1989): 289-317.

Hasseltoft, Henrik. "Stocks, bonds, and long-run consumption risks." *Journal of Financial and Quantitative Analysis* 47, no. 02 (2012): 309-332.

Hastings, W. Keith. "Monte Carlo sampling methods using Markov chains and their applications." *Biometrika* 57, no. 1 (1970): 97-109.

Jacoby, Gady, Paseka, Alex, and Wang, Yan. “A Generalized Earning-Based Stock Valuation Model with Learning.” Working Paper, Asper School of Business , University of Manitoba, 2014.

Jagannathan, Ravi, and Marakani, Srikant. "Price-Dividend Ratio Factor Proxies for Long-Run Risks." *Review of Asset Pricing Studies* 5, no.1 (2015): 1-47.

Jagannathan, Ravi, and Wang, Zhenyu. "The conditional CAPM and the cross-section of expected returns." *journal of finance* (1996): 3-53.

Jensen, Michael C. "The performance of mutual funds in the period 1945–1964." *The Journal of finance* 23, no. 2 (1968): 389-416.

Jong, Frank de. "Time series and cross-section information in affine term-structure models." *Journal of Business & Economic Statistics* 18, no. 3 (2000): 300-314.

Kiku, Dana. "Is the value premium a puzzle?." *Manuscript, Wharton Business School, University of Pennslyvania, Philadelphia* (2006).

Kim, Chang-Jin, and Nelson, Charles R. *State-space models with regime switching: classical and Gibbs-sampling approaches with applications. Vol. 2.* Cambridge: MIT press, 1999.

Lamont, Owen A. "Economic tracking portfolios." *Journal of Econometrics*105, no. 1 (2001): 161-184.

Lettau, Martin, and Ludvigson, Sydney. "Consumption, aggregate wealth, and expected stock returns." *Journal of Finance* (2001a): 815-849.

Lettau, Martin, and Ludvigson, Sydney. "Resurrecting the (C) CAPM: A cross-sectional test when risk premia are time-varying." *Journal of Political Economy* 109, no. 6 (2001b): 1238-1287.

Lettau, Martin, Ludvigson, Sydney C. and Wachter, Jessica A. "The declining equity premium: What role does macroeconomic risk play?."*Review of Financial Studies* 21, no. 4 (2008): 1653-1687.

Lewellen, Jonathan, and Nagel, Stefan. "The conditional CAPM does not explain asset-pricing anomalies." *Journal of financial economics* 82, no. 2 (2006): 289-314.

Lewellen, Jonathan, and Shanken, Jay. "Learning, asset-pricing tests, and market efficiency." *Journal of Finance* (2002): 1113-1145.

Lewellen, Jonathan, Nagel, Stefan, and Shanken, Jay. "A skeptical appraisal of asset pricing tests." *Journal of Financial Economics* 96, no. 2 (2010): 175-194.

Li, George. "Information quality, learning, and stock market returns." *Journal of Financial and Quantitative Analysis* 40, no. 03 (2005): 595-620.

Liptser, Robert, and Shiryaev, Albert N.. *Statistics of Random Processes: I. General Theory*. Vol. 5. New York: Springer Science & Business Media, 2013.

Ludvigson, Sydney C. *Advances in consumption-based asset pricing: Empirical tests*. No. w16810. National Bureau of Economic Research, 2011.

Lundblad, Christian. "The risk return tradeoff in the long run: 1836–2003."*Journal of Financial Economics* 85, no. 1 (2007): 123-150.

Lustig, Hanno, Van Nieuwerburgh, Stijn, and Verdelhan, Adrien. "The wealth-consumption ratio." *Review of Asset Pricing Studies* 3, no. 1 (2013): 38-94.

Maio, Paulo, and Santa-Clara, Pedro. "Multifactor models and their consistency with the ICAPM." *Journal of Financial Economics* 106, no. 3 (2012): 586-613.

Malloy, Christopher J., Moskowitz, Tobias J., and Vissing-Jørgensen, Annette. "Long-run stockholder consumption risk and asset returns." *The Journal of Finance* 64, no. 6 (2009): 2427-2479.

Metropolis, Nicholas, Rosenbluth, Arianna W, Rosenbluth, Marshall N., Teller, Augusta H., and Teller, Edward. "Equation of state calculations by fast computing machines." *The journal of chemical physics* 21, no. 6 (1953): 1087-1092.

Nagel, Stefan, and Singleton, Kenneth J. "Estimation and evaluation of conditional asset pricing models." *The Journal of Finance* 66, no. 3 (2011): 873-909.

Ortu, Fulvio, Tamoni, Andrea, and Tebaldi, Claudio. "Long-Run Risk and the Persistence of Consumption Shocks." *Review of Financial Studies* (2013): hht038.

Pástor, Ľuboš, and Veronesi, Pietro. "Stock Valuation and Learning about Profitability." *The Journal of Finance* 58, no. 5 (2003): 1749-89.

Pástor, Ľuboš, and Veronesi, Pietro. "Was there a Nasdaq bubble in the late 1990s?." *Journal of Financial Economics* 81, no. 1 (2006): 61-100.

Pástor, Ľuboš, and Veronesi, Pietro. “Technological Revolutions and Stock Prices”, *The*

Annual Review of Financial Economics 99, (2009): 1451-83.

Petkova, Ralitsa, and Zhang, Lu. "Is value riskier than growth?." *Journal of Financial Economics* 78, no. 1 (2005): 187-202.

Piazzesi, Monika, and Schneider, Martin. "Equilibrium yield curves." In *NBER Macroeconomics Annual 2006, Volume 21*, pp. 389-472. MIT Press, 2007.

Piazzesi, Monika, and Schneider, Martin. "Equilibrium yield curves." In *NBER Macroeconomics Annual 2006, Volume 21*, edited by Daron Acemoglu, Kenneth Rogoff and Michael Woodford, 389-472. Cambridge: MIT Press, 2007.

Robert, Christian, and Casella, George. *Monte Carlo statistical methods*. New York: Springer Science & Business Media, 2013.

Rubin, Donald B. "Bayesianly justifiable and relevant frequency calculations for the applied statistician." *The Annals of Statistics* 12, no. 4 (1984): 1151-1172.

Santos, Tano, and Veronesi, Pietro. "Labor income and predictable stock returns." *Review of Financial Studies* 19, no. 1 (2006): 1-44.

Schorfheide, Frank, Song, Dongho, and Yaron, Amir. *Identifying long-run risks: A bayesian mixed-frequency approach*. No. w20303. National Bureau of Economic Research, 2014.

Shanken, Jay. "Intertemporal asset pricing: An empirical investigation."*Journal of Econometrics* 45, no. 1 (1990): 99-120.

Singleton, Kenneth J. *Empirical dynamic asset pricing.* Princeton: Princeton University Press, 2006.

Timmermann, Allan G. "How learning in financial markets generates excess volatility and predictability in stock prices." *The Quarterly Journal of Economics* (1993): 1135-1145.

Vassalou, Maria. "News related to future GDP growth as a risk factor in equity returns." *Journal of financial economics* 68, no. 1 (2003): 47-73.

Veronesi, Pietro. "How does information quality affect stock returns?." *The Journal of Finance* 55, no. 2 (2000): 807-837.

Wang, Kevin Q. "Asset pricing with conditioning information: A new test."*The journal of finance* 58, no. 1 (2003): 161-196.

West, Mike, and Harrison, Jeff. *Bayesian Forecasting and Dynamic Models*. New York: Springer Science & Business Media, 1997.

Zhou, Ji, and Alex Paseka. "Equity Pricing and Risk Premium under Long-Run Risks and Incomplete Information." *Journal of Mathematical Finance* 4, no. 04 (2014): 279.

Zhu, Yingzi. "Dynamic volatility strategy with recursive utility." *Available at SSRN 890040* (2006).