Surplus Consumption, Habit Utility and Moody Investors

Jun Lou

The Graduate Center, City University of New York

How does access to this work benefit you? Let us know!

Follow this and additional works at: https://academicworks.cuny.edu/gc_etds

Part of the Finance Commons, Finance and Financial Management Commons, and the Macroeconomics Commons

Recommended Citation

https://academicworks.cuny.edu/gc_etds/2400

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
SURPLUS CONSUMPTION, HABIT UTILITY AND MOODY INVESTORS

by

JUN LOU

A dissertation submitted to the Graduate Faculty in Business in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2017
SURPLUS CONSUMPTION, HABIT UTILITY AND MOODY INVESTORS

by

JUN LOU

This manuscript has been read and accepted by the Graduate Faculty in Business in satisfaction of
the dissertation requirement for the degree of Doctor of Philosophy.

Professor Christos Giannikos

Date

Chair of Examining Committee

Professor Wim Vijverberg

Date

Executive Officer

Professor Christos Giannikos

Professor Wim Vijverberg

Professor Sebastiano Manzan

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

iii
Abstract

SURPLUS CONSUMPTION, HABIT UTILITY AND MOODY INVESTORS

by

JUN LOU

Adviser: Professor Christos Giannikos

The thesis examines a blend of Asset Pricing topics: joint stock-bond pricing, consumption-based asset pricing puzzles, time variation in risk preference, among others. In chapter one, I first review the literature on respective topics in search of a consolidated framework of resolution. I then propose one, a consumption-based affine model that jointly prices bond and stock in closed form. The tractable feature of the price solutions remains standard as in affine term-structure of interest rates, but presents novelty for the stock prices. In chapter two, I discuss the GMM based procedures for model estimation. In chapter three, I interpret the empirical results. I find the model broadly matching most first and second moments of stock, bond and macro variables, the time-series behavior and long-horizon predictability of returns. I contrast my model with prior frameworks to reveal some of their imprecise predictions and my model’s more plausible accountability in risk aversion. Specifically, a revisit to Campbell-Cochrane habit model using current data exposes the increasingly widening gap in post-1990s price-dividend ratio predictions. Meanwhile, an out-of-sample test indicates improved predictive power in my model for stock price dynamics particularly during more recent decades.
Acknowledgments

This thesis would not be possible without the constant mentorship of my advisor, Professor Christos Giannikos. I am forever grateful for his tireless expert guidance, understanding and encouragement throughout the entire process. I feel fortunate that I had the opportunity of being his student.

I would also like to sincerely thank Professor Wim Vijverberg for always being available to help. I greatly appreciate Professor Sebastiano Manzan's discussions and comments on this thesis. My sincere gratitude also goes to Professor Hany Guirguis for his patience, availability, valuable feedback and long-time support.

For my early studies, I am grateful to Professor Jessica Wachter, who offered me her Mablab files on habit models. I also thank Professor John Campbell and Professor John Cochrane for their email responses clarifying numerical iteration procedures used in their paper.

Finally, I want to thank my parents for their lifelong commitment and unconditional love, without which I can never be where I am today.
# Contents

<table>
<thead>
<tr>
<th>Contents</th>
<th>vi</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>vii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>ix</td>
</tr>
</tbody>
</table>

1 **Literature Review and Model of Joint Stock-Bond Pricing** 1

1.1 Literature Review and Introduction .................................................. 1

1.2 Model Specification .............................................................................. 8

  1.2.1 Proposition 1 .................................................................................. 14

  1.2.2 Proposition 2 .................................................................................. 16

2 **Estimation Methodology and Data Construction** 19

  2.1 Moment Conditions ............................................................................. 20
2.2 Two-Stage Procedure ................................................................. 22
2.3 Data Construction................................................................. 25

3 Empirical Evaluation ........................................ 28
  3.1 Parameters ................................................................. 28
  3.2 Stock and Bond Moments .............................................. 30
  3.3 Time Variation.............................................................. 34
  3.4 The Unobservable............................................................ 34
  3.5 Comparison with Related Models ............................... 36
  3.6 Time Series Movement.................................................. 38
  3.7 Out-of-Sample Test ....................................................... 41
  3.8 Long-Horizon Return Regression ............................... 42
  3.9 Variant Specification...................................................... 44
      3.9.1 Specification ...................................................... 45
      3.9.2 findings .......................................................... 46
  3.10 Concluding Remarks .................................................... 47

Appendix ............................................... 49
  I Proof of Joint Stock-Bond Pricing ......................... 49
  II Analytical Moments .................................................. 55

Bibliography ........................................... 71
List of Tables

1 Parameters Estimation ..................................................................................................... 57
2 Moment Comparison Summary ...................................................................................... 58
3 Cross Correlations ........................................................................................................... 59
4 Variance Decomposition ................................................................................................. 60
5 Risk Aversion Property by Bekaert et al. (2010) ............................................................ 60
6 Risk Aversion Property ................................................................................................... 61
7 Risk Aversion Correlations by Different Models ........................................................... 61
8 Long Horizon Return Regression .................................................................................... 62
9 Parameters Estimation of Variant (32) ........................................................................ 63
10 Std. Dev. & Correlation of State Variables for Variant (32) ........................................... 64
11 Cross Correlations of Variant (32) .................................................................................. 65
List of Figures

1. Price-Dividend Ratio as a Function of Surplus Consumption Ratio ..................66
2. Latent State Variable, Surplus Ratio $q$ Extracted by Kalman Filter ..................66
3. Distribution of Surplus Ratio $q$ ........................................................................67
4. Risk Aversion $RA = \gamma \exp(q)$ ........................................................................67
5. Predicted (Filtered) Risk-Free Rate (Dash Line) vs. Actual (Solid Line) ..........68
6. Predicted (Filtered) Yield Spread (Dash Line) vs. Actual (Solid Line) ..........68
7. Predicted Return (Dash Line) vs. Actual (Solid Line) ........................................69
8. 1970-2013 Surplus Ratio $q$ Predicted Out Of Sample ......................................69
9. 1970-2013 Campbell-Cochrane Predicted $P/D$ (Dot Line) vs. Actual (Solid Line)...70
10. 1970-2013 My Model Predicted $P/D$ (Dot Line) vs. Actual (Solid Line) ............70
Chapter 1

Literature Review and Model of Joint Stock-Bond Pricing

1.1 Literature Review and Introduction

The classical consumption-based asset pricing is a class of models characterized by a pricing kernel process that governs the joint processes of asset returns and per capita consumption. Despite the theoretical elegance, its inability to rationalize the reality of US financial markets over the past century was first observed by Mehra and Prescott (1985). Among all research efforts to enhance the model ability to replicate the empirical data, two primary mechanisms stand out without modifying the assumptions of complete market and representative agent. The first, pioneered by Bansal and Yaron (2004), attacks the i.i.d. consumption growth
assumption by introducing a small predictable component into it. This long run component increases the market price of consumption risk to generate large equity premia.

The second proposed mechanism, a particularly successful version of which is Campbell and Cochrane (1999) (hereafter CC1999), deviates from CRRA assumption by developing external habit. The habit formation generates time variation in risk appetite and hence the investors become moody. The notion of habit gains support in the sense that models of standard CRRA preference have also proposed habit-like channels that induce time-varying risk appetite. Examples include work by Piazzessi et al. (2007) who introduce housing into a consumption-based asset pricing model to give rise to an additional composition risk, i.e., risk of fluctuations in the share of housing expenditure relative to other nondurable consumption, other than the standard consumption growth risk. Composition share is a persistent and heteroskedastic state variable driving the pricing kernel, a role like the habit of CC1999. Flavin and Nakagawa (2008) devise their housing-CCAPM following Grossman and Laroque (1990). They show that, due to adjustment cost, infrequent home purchase decisions turn housing into a slow-moving state variable, a role again similar to the persistent habit in CC1999.

Thus, I take the paper of CC1999 as the origin of my thesis work. Despite the popularity and success with many observed features of the asset market and its influence in the field, the CC1999 paper inevitably has been questioned over certain aspects. (i) Their model is evaluated using calibration and simulation but not formal econometric estimation. (ii) Absent from the model is the analytical closed-form pricing solutions.\(^1\) As a result, they use numerical methods instead to determine numerically the functional relations among the variables of interest. (iii) The model relies on one single state variable along with the ideal assumption of perfect correlation

\(^1\) Closed-form prices are actually available for bonds but not for stocks in the working paper version of the model, Campbell and Cochrane (1995).
between consumption growth and surplus ratio. CC1999 does not estimate and test their model mainly because the absence of analytic representation for equities limits the availability of moment conditions that can be used for estimation. Specifically, second and higher moments on equities are not available but only the first moments such as the Euler condition can be used for testing. On the other hand, there is insufficient research on coherent bond-stock pricing in an analytical framework. More often the two classes of assets are priced separately. One difficulty is that almost any reasonably modeled dividend process lacks analytic tractability of the stock prices. In other words, even in a unifying pricing theory where the bond prices can often be straightforwardly derived in closed-form formulas, the stock prices however are usually left unsolved due to the presence of dividend. CC1999 is a typical example where the bond prices can be neatly formulated but not their stock prices or price-dividend ratios. My paper therefore contributes to both lines of research, addressing the challenges confronting CC1999 habit model while developing a joint pricing of bond and stock in closed-form.

With regard to aspect (i) in question, formal econometric techniques could help better assess the models beyond what calibrations and simulations can show. For example, Moller (2009) uses GMM estimation to uncover that CC1999 model can be rejected at value portfolio moments. Brandt and Wang (2003) who also estimate the model discover surprisingly that risk aversion in CC1999 exhibits little variation when only bond moments are fitted and that the model can produce substantial time-variation in risk aversion only when equity return moments are added. Brandt and Wang (2003) actually formulate a general model that can collapse into

---

2 Think about the Vasicek (1977) or CIR (1985) model that says nothing about stock prices. Clearly the bond-stock in this paper refers to term structure of interest rates and equity market respectively, not to be confused with the corporate bond and corporate stock of a same firm, whose values are not separately determined. See Merton (1974).

3 Technically in Campbell and Cochrane (1999, 1995), the bond prices in closed-form are only derivable for 1 and 2 period maturity bonds but not for any bonds that mature in longer periods due to their unique specification of the sensitivity function.
precisely CC1999 model under some restrictions. They estimate both models for the purpose of comparison. Others who also formally estimate and test the model include Fillat and Garduño (2005), Garcia et al. (2005), Tallarini and Zhang (2005). Møller’s (2009) estimation rejects value premium, but finds size premium well explained, which echoes Lettau and Wachter (2007) as well as Santos and Veronesi (2010), who conclude that the CC1999 model produces counterfactual predictions with respect to value premium documented in Fama and French (1993). Common in all five empirical articles which test and estimate, their GMM or EMM estimation utilizes only the first moment on equity returns.

As for challenge (ii) on tractable pricing formula, most follow-up articles have not made substantial differences. Watcher (2006) uses the exact CC1999 original model with the same calibration and grid-search technique to develop Campbell and Cochrane’s work into bond pricing. However, those questioned aspects remain unaddressed. The same is true with Brandt and Wang (2003), which appears worse in the sense that even their bond prices cannot be evaluated analytically.

In response to CC1999 concern (iii) of a single state variable and prefect correlation specification, some subsequent research does generalize by naturally inserting additional state variables and quite often relaxing the perfect correlation design. Those major contributions include Menzly et al. (2004), Bekaert et al. (2010) and Brandt and Wang (2003). Menzly et al. (2004) construct multiple dividend streams with inverse surplus ratio to explore equity return predictability, while the bond price is not covered. Bekaert et al. (2010) introduce three more state variables other than surplus ratio to form a four state variable VAR structure. And the perfect correlation situation stays easily incorporated into their generalized framework. However, their solutions in the equity prices, although managed to look linear in the state vector, actually
are computation heavy due to their complicated structures. Brandt and Wang (2003) append consumption and inflation dynamics to the risk aversion of the original Campbell-Cochrane model. They also let risk aversion freely correlate with shocks of both consumption growth and unexpected inflation in order to absorb the special case of perfect correlation into their general setup. Unfortunately, due to their chosen sensitivity function, neither bond nor equity prices in analytical expressions can be achieved. Furthermore, even the numerical method is computationally infeasible in their case. They end up relying on simulations for the prices.

On the second line of research, only a few articles try to merge the separation of bond and equity pricing. One such contribution is Bekaert et al. (2010), who price both bond and stock in closed-form and in a similar-looking affine pattern, following their earlier and original work developed in Bekaert and Grenadier (2002). Nevertheless, as mentioned, their equity part takes on very complicated expressions despite being affine in the final appearance. Specifically, they manage to represent the price-dividend ratio \( PD_t = \sum_{n=1}^{\infty} \exp(b_n^0 + b_n^1 Y_t) \), as an infinite sum of exponential affine functions of the state variables. One would have to sum up large number of terms (authors say 200 periods is what they use, and keep extending another 100 periods till desirable) with each term requiring \( n \) recursions to find the solution (say for example, the 200th term runs 200 times recursively to determine itself). As a result, when I replicate Bekaert et al. (2010) work, I notice that the computation of numerical GMM optimization turns out dramatically time consuming. Meanwhile, Mamaysky (2002) proposes another way to price bond and stock in a unified affine fashion. His theoretical model prices the stocks similarly to the way bonds are priced, which is, exponential affine in state variables. His paper is nevertheless silent on whether his tractable approach possesses desirable empirical properties.

---

4 Another article is Bakshi and Chen (1996), which formulates a monetary asset pricing model capable of jointly pricing bond and equity in closed form. But as a hybrid model of monetary economics and asset pricing, money-in-the-utility function is what they use, which my paper does not intend to borrow.
By reviewing the relevant literature, I come to establish the goals of my paper. First of all, the habit component is essential to form the time-varying conditional volatility of pricing kernel and therefore to generate countercyclical price of risk. That said, I inherit the surplus ratio as a key state variable to help reconcile the equity premium puzzle. Next, I generalize. Like Brandt and Wang (2003) and Bekaert et al. (2010), I allow for multiple state variables, specifically four, and allow for rather general inter-dependence among them. Unlike Brandt and Wang (2003), I require an analytical evaluation of both bond and equity prices. So the model is similar to Bekaert et al. (2010), but without the computational burden in their equity prices. I therefore re-model their stocks pricing portion by borrowing Mamaysky’s (2002) technique. This way I am able to unify the bond and equity prices into closed-form affine formulas, which are simply-looking, tractable and computation friendly. Moreover, I formally estimate and test my model using GMM. Because Mamaysky’s (2002) model is not empirically evaluated, it is tempting to see how it performs with the data. Different from earlier estimations by Moller (2009), Brandt and Wang (2003), Garcia et al. (2005) or Tallarini and Zhang (2005), now I have equity prices analytically ready. That enables my GMM to explore quantities like variances of stock return and its covariance with others variables by utilizing abundant second order moments on equity returns that have not been unused before. My paper does not address topics related to variations in cross-section returns but deal with aggregate equity returns only.

Model estimation in GMM uses an approach conceptually similar to Brandt and Wang’s (2003) two-stage procedure in which each stage handles structural and preference parameters separately. My results on parameter estimates are fairly accurate in that most parameter values are statistically significant with reasonable economic intuition. Estimates confirm a countercyclical and persistent surplus ratio consistent with major literature. For the 21 parameters in 23
moment conditions, the model fails to reject at 1 percent level but rejects at 5 percent level. Model can capture most means, variance, auto-covariance and cross-correlations practically well, a total of more than forty first and second moments, of which only 23 are meant to fit. The model implies that inflation together with risk aversion explains the majority of the time variation in nominal term yields and spreads. As for stock returns, most of the variation is attributed to the variance of dividend yield and risk aversion. I use a Kalman filter to extract the latent variable, surplus ratio, in order to find time series predicted by the model. The model-implied artificial series move rather similarly to the real data. The model is also able to replicate the long horizon return regression.

Risk aversion and business cycle are believed to correlate. But there is the question of how strongly. My finding suggests differently from the earlier habit formation papers. Earlier works such as CC1999 predicts an almost linear relation between risk aversion and dividend yield, which is a business cycle forecaster, or equivalently, a very high correlation coefficient, such as 0.93 predicted by Bekaert et al. (2010). I however point out that the strong relation might well be implausible. As for the question whether the risk-free rate is pro- or counter-cyclical, by Wachter’s (2006) calibration, the original Campbell-Cochrane model would imply a correlation of real risk-free with risk aversion that is close to 1. Wachter’s (2006) model, due to added state variables, suggests a slightly smaller coefficient for both real and nominal risk-free rate. Brandt and Wang (2003) also find it to be very large, at greater than 0.90 for nominal bond yield. My results remain in line with their claims that the real or nominal risk-free rate does fairly substantially correlate with risk aversion, but my calculated coefficient goes down to 0.69.

When comparing several related habit models, I briefly revisit CC1999, update their predictions to 2013, but find their price-dividend ratio hardly close to the actual level ever after
the early 1990s. On the other hand, I conduct an out-of-sample test with my model that confirms valid predictive ability in time of cyclical fluctuations. More noticeably, my predictions of equity price evolution from early 1990s to date and out of sample show considerably improved accuracy.

Other work related to Campbell-Cochrane model that contributes to the field of study includes Buraschi and Jiltsov (2007) and Verdelhan (2010). The former develops habit formation into a monetary nonaffine model for term structure of interest rates, and the latter extends into the area of exchange rate.

1.2 Model Specification

Following CC1999, the representative agents in a Lucas (1978) type economy maximize a utility function that contains an added element of habit:

\[ E \sum_{t=0}^{\infty} \beta^t \frac{(C_t - X_t)^{1-\gamma} - 1}{1-\gamma} \]

(1)

where \( X_t \) denotes the level of consumption habit. I define the habit to always be nonnegative but less than the actual consumption level \( C_t \) thus \( C_t - X_t \) represents surplus consumption that should be always positive. Parameter \( \gamma \) is the risk aversion coefficient and \( \beta \) is time preference. Notice that in contrast to the conventional utility of consumption, now it is the utility of surplus consumption that matters.

The notion of habit preference and surplus consumption is motivated, according to Cochrane (2005), by the idea that risk aversion must depends on the level of consumption relative to some ‘trend’ for a model to work around the conventional pitfalls. The habit design here is not innovative but follows the ‘keeping up with the Jones’ notion of Abel (1990) and the literature has long introduced habit for the study of consumption. See a survey by Deaton (1992).
Surplus ratio is therefore defined as \((C_t - X_t)/C_t\). This ratio ranges within \((0, 1)\) given my earlier confinement on the habit level \(X_t\). I differ from CC1999 but follow Bekaert et al. (2010) by using the inverse of the surplus ratio \(Q_t = C_t / (C_t - X_t)\), where apparently \(Q_t > 1\). The advantage of using the inverse is straightforward, because after taking natural logarithm it fits the need of square root process, which one will see immediately. Relative risk aversion now equals \(\gamma \cdot Q_t\). Let \(q_t = ln(Q_t)\) be my state variable, it follows that \(q_t > 0\) and I specify its evolution over time,

\[
q_{t+1} = \mu_q + \rho_q q_t + \sigma_q \sqrt{q_t} (\sqrt{1 - \lambda^2} \epsilon_{t+1}^q + \lambda \epsilon_{t+1}^c)
\]

(2)

where the parameters in this stationary process are \(\mu_q\), \(\rho_q\), \(\sigma_q\) and \(\lambda\). Apparently \(\rho_q\) is the feedback parameter and is supposed to be less than 1 for the process to remain stationary. \(\sigma_q\) is the volatility parameter. \(\epsilon_{t+1}^q\) and \(\epsilon_{t+1}^c\) are i.i.d. standard normal innovations inherent to \(q_{t+1}\) and \(\Delta c_{t+1}\) respectively. This process of inverse surplus ratio \(q_{t+1}\) adopts a square root process due to my objective of asset prices being neatly and tractably affine solutions in terms of the state vector. Another purpose for using square root format is to introduce time variation into the Sharpe ratio. As will be seen soon, my goals are indeed achieved in this way. I follow Bekaert et al. (2010) by including the parameter \(\lambda\) in the innovations of the process. The whole point of having \(\lambda\) is that it turns out to be the conditional correlation coefficient between the surplus ratio and consumption growth, as will be clear as soon as one sees the consumption growth specification in equation (3) the next paragraph.

Consumption growth is specified in a similar fashion, but depends on both its own past state and the past value of dividend yield. Equation (3) is clearly a more general formation than CC1999, who define consumption growth as a strict i.i.d. process written as \(\Delta c_{t+1} = g + \sigma_c \epsilon_{t+1}^c\). In particular, this simple formulation is simply one possibility out of many variations that
equation (3) can accommodate. Later on, the data show that the consumption growth in reality fits more like a mean reversion with moderate autocorrelation rather than a pure i.i.d. with zero autocorrelation as suggested by CC1999. By having dividend yield enter into consumption growth, I follow Bekaert et al. (2010) to allow for more practical dynamics among the state variables. But I differ from Bekaert et al. (2010) in the selection of state variable. I use \( dp_t \) the dividend yield rather than \( \Delta d_t \) the dividend growth as in Bekaert et al. (2010).\(^5\)

\[
\Delta c_{t+1} = \mu_c + \rho_{cc} \Delta c_t + \rho_{cd} dp_t + \sigma_{cc} \sqrt{q_t} \varepsilon_{t+1}^c
\]

(3)

Where \( \mu_c, \rho_{cc}, \rho_{cd} \) and \( \sigma_{cc} \) are parameters of consumption growth.

With the above two processes set forward, let me derive the pricing kernel as an immediate follow up in (4),

\[
M_{t+1} = \frac{u_c(c_{t+1},X_{t+1})}{u_c(c_t,X_t)} = \beta \left( \frac{Q_t}{Q_{t+1}} \frac{c_{t+1}}{c_t} \right)^{-\gamma}
\]

(4)

\[
= \beta \exp(-\gamma \Delta c_{t+1} + \gamma \Delta q_{t+1}).
\]

Apparently the model uses an endogenous pricing kernel due to consumption-based nature as opposed to a general kernel whose existence Harrison and Kreps (1979) theorize. It is easy to see how this pricing kernel works to fix the puzzles faced by the standard power utility model without habit formation. The added habit component \( q_t \) is taking the pressure off a very high \( \gamma \) that reconciles the 0.5 historical Sharpe ratio in standard model when consumption growth alone fights risk-free rate puzzle. Notice now because the discount factor concerns only two contemporary state processes \( \Delta c_t \) and \( q_t \), its conditional variance can be determined promptly,

\[
Var_t[m_{t+1}] = \gamma^2 (\sigma_{cc}^2 + \sigma_q^2 - 2 \sigma_{cc} \sigma_q \lambda) q_t
\]

(5)

\(^5\) Bekaert et al. (2010) aims to indirectly determine the dividend growth process through the consumption-dividend ratio process. So strictly speaking his state variable is the consumption-dividend ratio.
where \( m_{t+1} = \ln(M_{t+1}) \). What I pointed out earlier now becomes clear: the conditional standard deviation of the log pricing kernel, which equals the maximum Sharpe ratio, turns out to be an increasing function in \( q_t \). Therefore I have introduced time variation into the Sharpe ratio via the square root specification in the state variables \( \Delta c_t \) and \( q_t \). Accordingly, the Sharpe ratio varies positively with the fluctuation in surplus ratio \( q_t \).

Let me then decide how to model the dividend process. As pointed out by Mamaysky (2002), when dividend growth is used as a state variable, usually the closed-form formula of equity prices is hardly possible to come by, instead the price dividend ratio normally can be derived in closed form solutions. On the other hand, when dividend yield is used as the state variable, the equity prices may be solved in a closed-form. Since equity prices are my prime interest, I choose to use dividend yield as my state variable. That is one of my major departures from Bekaert et al. (2010) who derives the price dividend ratio as an analytic function of a state vector containing dividend growth,\(^6\) consistent with Mamaysky (2002) finding. In equation (6),

\[
d p_{t+1} = \mu_d + \rho_{dd} d p_t + \rho_{dc} \Delta c_t + \sigma_{dc} \sqrt{q_t} \varepsilon_{t+1}^{c} + \sigma_{dd} \varepsilon_{t+1}^{d}
\]

Where \( \mu_d, \rho_{dd}, \rho_{dc}, \sigma_{dc} \) and \( \sigma_{dd} \) are parameters of the process, dividend yield is modeled symmetrically to consumption growth. The two series depend on each other’s lag values. The square root structure stays in consistency to the two previous processes to insure tractable prices. Dividend yield has an innovation \( \varepsilon_{t+1}^{d} \) specific to itself and also an innovation to connect with consumption growth. This structure allows for an arbitrary condition of cross-correlations among the three state variables introduced so far with no loss of generality.

Lastly I specify a very simple inflation process,

\[
\pi_{t+1} = \mu_\pi + \rho_\pi \pi_t + \sigma_\pi \varepsilon_{t+1}^\pi
\]

---
\(^6\) Precisely, the state vector has consumption-dividend ratio as one of the variables.
where $\mu_\pi, \rho_\pi, \rho_{dc}$ and $\sigma_\pi$ are the parameters of the process. Brandt and Wang (2003) model unexpected inflation to correlate with risk aversion but I build no such a link into equation (7).

Now the full model is gathered,

$$q_{t+1} = \mu_q + \rho_q q_t + \sigma_q \sqrt{q_t} (\sqrt{1 - \lambda^2} \xi_{t+1}^q + \lambda \xi_{t+1}^e)$$
$$\Delta c_{t+1} = \mu_c + \rho_{cc} \Delta c_t + \rho_{cd} dp_t + \sigma_{cc} \sqrt{q_t} \xi_{t+1}^c$$
$$dp_{t+1} = \mu_d + \rho_{dd} dp_t + \rho_{dc} \Delta c_t + \sigma_{dc} \sqrt{q_t} \xi_{t+1}^e + \sigma_{dd} \xi_{t+1}^d$$
$$\pi_{t+1} = \mu_\pi + \rho_\pi \pi_t + \sigma_\pi \xi_{t+1}$$

and the pricing kernel is,

$$m_t = \ln(\beta) - \gamma \Delta c_t + \gamma \Delta q_t$$

In order to formulate a general pricing expression, I rewrite the whole structure into a compact matrix form,

$$Y_{t+1} = \mu + A Y_t + (\Sigma_F F_t + \Sigma_H) \epsilon_{t+1}$$

where I collect four state variables in vector $Y_{t+1} = [q_{t+1}, \Delta c_{t+1}, dp_{t+1}, \pi_{t+1}]'$. Thus specifically for my model the parameters $\mu$ and $\omega$ are $4 \times 1$ vectors and parameters $A$, $\Sigma_F$, $\Sigma_H$ and $\Omega$ are $4 \times 4$ matrices. Innovation $\epsilon_{t+1}$ is a $4 \times 1$ vector of zero mean unit variance i.i.d. noise $[\epsilon_{t+1}^q, \epsilon_{t+1}^c, \epsilon_{t+1}^d, \epsilon_{t+1}^\pi]'$. $\| \|$ denotes element-wise square root operator. $\odot$ denotes element-wise multiplication. $F_t$ is of size $4 \times 4$ and represents the square-root components in innovations of equation (8).

I provide the following details to link (8) to (10):

$$\Sigma_H = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{dd} & 0 \\ 0 & 0 & 0 & \sigma_\pi \end{bmatrix}, \quad \Sigma_F = \begin{bmatrix} \sigma_q \sqrt{1 - \lambda^2} & \sigma_q \lambda & 0 & 0 \\ 0 & \sigma_{cc} & 0 & 0 \\ 0 & 0 & \sigma_{dc} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
In the same spirit, the real pricing kernel can be written as,

\[ m_{t+1} = \mu_m + A_m' Y_t + (\Sigma_F F_t' + \Sigma_H H_t') \epsilon_{t+1} \]  \hspace{1cm} (11)

where parameter \( \mu_m \) is a scalar and parameters \( A_m, \Sigma_F, \) and \( \Sigma_H \) are \( 4 \times 1 \) vectors for my model but become \( k \times 1 \) when the general model has \( k \) state variables.

Again, I provide the following details to link (9) to (11):

\[ \mu_m = \ln(\beta) - \gamma \mu_c + \gamma \mu_q \quad A_m = \gamma \begin{bmatrix} \rho_q - 1 \\ -\rho_{cc} \\ -\rho_{cu} \\ 0 \end{bmatrix} \quad \Sigma_H = \gamma \begin{bmatrix} \rho_{cc} \sqrt{1 - \lambda^2} \\ -\rho_{cc} + \lambda \rho_{qq} \\ 0 \\ 0 \end{bmatrix} \quad \Sigma_F = \gamma \begin{bmatrix} \rho_{qq} & 0 & 0 & 0 \\ 0 & \rho_{cc} & 0 & 0 \\ 0 & 0 & \rho_{dd} & 0 \\ 0 & 0 & 0 & \rho_{uu} \end{bmatrix} \]

Some restrictions are imposed on the parameters from the general compact expression (10) (11),

\[ \Sigma_F F_t' \Sigma_H = 0 \]
\[ \Sigma_H F_t' \Sigma_F = 0 \]
\[ \Sigma_F F_t' \Sigma_F = 0 \]
\[ \Sigma_H F_t' \Sigma_H = 0 \]  \hspace{1cm} (12)

Basically these restrictions are to keep the square root terms out of the second order moments of the state variables. If the square root terms were to appear in the second moments, the recursive pricing procedures which are supposed to generate prices linear in the state vector at each recursion will cease to progress due to nonlinearity. In other words, these restrictions are needed to guarantee recursive pricing stay proceeding. See the proof in Appendix I for detailed algebra. My four state variable model (8), as a specific case of the general compact framework (10), certainly satisfies the above restrictions and can easily be verified.
Let me now price the nominal term structure of interest rates first. Let \( P^b_{n,t} \) be the time \( t \) price of a nominal bond which will mature in \( n \) periods. The nominal pricing kernel is \( \hat{m}_t = m_t - \pi_t \). The bond prices must always follow the recursive pricing rule,

\[
P^b_{n,t} = E_t [\exp(\hat{m}_{t+1}P^b_{n-1,t+1})]
\]  

(13)

where \( E \) is the expectation operator. Bond prices that satisfy this pricing rule are solved and summarized in the following proposition. I refer the proof to Appendix I.

1.2.1 Proposition 1. The price of a time \( t \) zero-coupon bond with maturity \( n \) is exponential affine in state vector \( Y_t \),

\[
P^b_{n,t} = \exp(a^c_n + a'_n Y_t)
\]  

(14)

where \( a^c_n \) is a scalar given by,

\[
a^c_n = a^c_{n-1} + (a'_n - e'_n)\mu + \mu_m + \frac{1}{2} \left((\Sigma'_F(a_{n-1} - e_n)) \odot (\Sigma'_F(a_{n-1} - e_n))\right) \omega + \frac{1}{2} \left(\Sigma_{Fm} \odot \Sigma_{Fm}\right) \omega + \frac{1}{2} \left(\Sigma_{hm} \odot \Sigma_{hm}\right) \omega + (a'_{n-1} - e'_n)(\Sigma'_F \odot \Sigma_F) \omega + (a'_{n-1} - e'_n)(\Sigma''_F \odot \Sigma'_F) \omega + (a'_{n-1} - e'_n)(\Sigma''_F \odot \Sigma''_F) \omega
\]  

(15.1)

and \( a_n \) is a \( 4 \times 1 \) vector of coefficients,

\[
a'_n = (a'_{n-1} - e'_n)A + A'_m + \frac{1}{2} \left((\Sigma'_F(a_{n-1} - e_n)) \odot (\Sigma'_F(a_{n-1} - e_n))\right) \Omega + \frac{1}{2} \left(\Sigma_{Fm} \odot \Sigma_{Fm}\right) \Omega + (a'_{n-1} - e'_n)(\Sigma'_F \odot \Sigma_F) \Omega + (a'_{n-1} - e'_n)(\Sigma''_F \odot \Sigma''_F) \Omega
\]  

(15.2)

and where \( a^c_0 = 0, a'_0 = [0 0 0 0], P^b_{0,t} = 1 \) and \( e'_\pi = [0 0 0 1] \).

Parameter \( e'_\pi \) serves to single out element \( \pi_t \) from \( Y_t \). Note that the two coefficients \( a^c_n \) and \( a_n \) obey a recursive pattern. Particularly, \( a_n \) develops through \( a_{n-1} \) but has nothing to do with \( a^c_n \). For bonds distant from maturity, \( a_n \) has to converge otherwise bond prices will diverge as \( n \)
approaches infinity. On the other hand \(a_n^c\) depends on both \(a_{n-1}^c\) and \(a_{n-1}\). Specifically, \(a_n^c\) is the sum of \(a_{n-1}^c\) and some specific function of \(a_{n-1}\), which indicates that, for distant-to-maturity bonds, \(a_n^c\) would keep differing from \(a_{n-1}^c\) by a fixed amount. This certainly makes sense, because the further away from maturity, the smaller a bond price becomes. However in practice, to numerically keep \(a_n\) in converged is not a straightforward task due to the large number of parameters involved that do not organize themselves linearly. That is one of the challenges in numerical application to minimize the GMM objective function. Many different combinations of parameter values can easily trigger \(a_n\) to become divergent. And there is no once-for-all safe zone for parameter value selection. I explain how I cope with this difficulty in Chapter 2 on methodology.

Applying Proposition 1, it is easy to show that the risk-free rate is affine in state vector \(Y_t\),

\[
    r_t^f = -\ln(p_{1,t}^b) = -p_{1,t}^b = -(a_1^c + a_1^c Y_t) = -a_1^c - a_1^c Y_t
\]

(16)

The term spread is the difference of yield to maturity between a long term bond and a short term bond. Again, applying Proposition 1, the term spread is affine in state vector \(Y_t\),

\[
    s_t^{pd} = \text{YTM of 10 yrs bond} - \text{riskless rate}
\]

\[
    = -\frac{1}{10} \ln(p_{10,t}^b) - r_t^f = -\frac{1}{10}p_{10,t}^b - r_t^f
\]

\[
    = -\frac{1}{10}(a_{10}^c + a_{10}^c Y_t) + (a_1^c + a_1^c Y_t)
\]

\[
    = -\frac{1}{10}a_{10}^c + a_1^c + \left(-\frac{1}{10}a_{10}^c + a_1^c\right) Y_t
\]

(17)

where I select 10 year bond as my long term bond.

Let me now price the equities. Similarly, equity prices must follow the recursive pricing rule,

\[
    P_t^e = E_t[exp(m_{t+1})(P_{t+1}^e + D_{t+1})]
\]

(18)
where I use the real pricing kernel $m_{t+1}$ because all prices and dividends are in real terms. The equity prices that satisfy this pricing equation can be solved in the following proposition.\(^7\) Again I refer the readers to Appendix I for the proof.

1.2.2 **Proposition 2.** At time $t$, the price of an equity paying infinity periods of dividends solves to an exponential affine form,

$$P_t^e = \exp(b_n^c + b'Y_t + Z_t)$$

(19)

where $b_n^c$ is a scalar given by

$$b_n^c = b_{n-1}^c + (e_{dp}' + b')\mu + \mu_m + \frac{1}{2}\left[\left(\Sigma_F'(e_{dp} + b)\right)\otimes \left(\Sigma_F'(e_{dp} + b)\right)\right]'\omega + \frac{1}{2}(e_{dp}' + b')\Sigma_H\Sigma_H'(e_{dp} + b) + \frac{1}{2}[\Sigma_F\otimes\Sigma_F]'-\omega + \frac{1}{2}\Sigma_{Hm}\Sigma_{Hm} + (e_{dp}' + b')\left(\Sigma_{Fm}\otimes\Sigma_F\right)\omega + (e_{dp}' + b')\Sigma_H\Sigma_{Hm}$$

(20.1)

and $b$ is a $4\times1$ vector of coefficients implicitly determined by

$$b' = B' + (e_{dp}' + b')A + A_m' + \frac{1}{2}\left[\left(\Sigma_F'(e_{dp} + b)\right)\otimes \left(\Sigma_F'(e_{dp} + b)\right)\right]'\Omega + \frac{1}{2}[\Sigma_F\otimes\Sigma_F]'\Omega + (e_{dp}' + b')\left(\Sigma_{Fm}\otimes\Sigma_F\right)\Omega$$

(20.2)

and $Z_t$ is a random walk component whose first difference is stationary and assumed to be linear in $Y_t$ with a $4\times1$ loading coefficient vector $B$, following the pattern (21),

$$\Delta Z_{t+1} = Z_{t+1} - Z_t = B'Y_t$$

(21)

and where $e_{dp}'$ is the selection vector $[0 \ 0 \ 1 \ 0]$.

\(^7\) The accuracy of the pricing solutions (14) and (19) from both propositions relies on non-negative $q_t$. As Bekaert et al. (2010) suggest, in the rare event that $q_t$ goes negative, I bound it at zero. However, in that case, my solutions become a close approximation but accurate. Therefore, it is crucial that $q_t$ stays non-negative most of the time. My simulation of 10,000 draws of $q_t$ has negative occurrence only less than 7% of the time.
Similar to the recursive structure in bond price of Proposition 1, parameter $b$ develops on its own. But more precisely now it is implicitly determined in equation (20.2). Parameter $b_n^c$ is the sum of its prior $b_{n-1}^c$ and a function of $b$. The increment in $b_n^c$, which is $b_n^c - b_{n-1}^c$, represents the equity prices growing over time.

Applying Proposition 2 I derive the price growth or capital gain, which is affine in state vector $Y_t$ and $Y_{t-1}$,

$$
\Delta p_t = p_t^e - p_{t-1}^e = b_n^c - b_{n-1}^c + b'Y_t + (B' - b')Y_{t-1}
$$

(22)

and the excess return,

$$
r_t^{ex} = p_t^e - p_{t-1}^e + dp_t + \pi_t - r_{t-1}^f
$$

$$
= b_n^c - b_{n-1}^c + a_n^c + (b' + e_{dp} + e_{\pi})Y_t + (B' - b' + a_1^c)Y_{t-1}
$$

(23)

With the two propositions in hand, prior to diving into the details of Appendix I for proof, one can still gain some insights into how these pricing expressions better capture the dynamics than the classical consumption-based model. I take a quick look at nominal risk-free rate for an example. By Proposition 1 at coefficients $a_1^c$ and $a_1$ I open up the risk-free rate term by term,

$$
r_t^f = -\ln\beta + \gamma(c - \mu_q) + [\gamma(1 - \rho_q) - \gamma^2(\sigma_{cc}^2 + \sigma_{qq}^2 - 2\sigma_{cc}^2\sigma_{qq}^2\lambda)/2]q_t +
$$

$$
\gamma\rho_{cc}\Delta c_t + \gamma\rho_{cd}dp_t + \mu_\pi + \rho_\pi\pi_t - \sigma_\pi^2/2
$$

(24)

where I see the four state variables all enter into the rate linearly. On the right hand side of equation (24), the third, fourth, fifth and seventh term each captures the sensitivity on the risk-free rate, in contrast to the conventional power utility model in which only consumption growth is the sole driving factor for the risk-free rate. Analytically I have many more channels than Mehra and Prescott (1985) to account for the yields, which is potentially how the risk-free rate puzzle may get resolved. Equation (24) would become real risk-free rate if the last three terms were removed. Compared to the standard power utility model, on the right hand side of (24),
\(-\gamma \mu_q\) of the second term and \(\gamma (1 - \rho_q) q_t\) of the third term together form the additional consumption smoothing effect while \(-q_t \gamma^2 (\sigma_{qq}^2 - 2\sigma_{cc}^2 \sigma_{qq}^2 \lambda)/2\) from the third term represents the additional precautionary saving effect. The richer contents in my expressions imply sources of flexibility. Moreover, because the effect of \(a_1\) breaks down into its four elements, I am able to identify the relative role each factor plays and to perform analysis by variance decomposition. Moller (2009) mentions that historically, the evidence is mixed as to whether risk-free rate is pro- or counter-cyclical. My model has potential power to investigate its cyclical nature.
Chapter 2

Estimation Methodology and Data

Construction

My economy has four state variables and four derived (endogenous) variables which are either linear in vector $Y_t$ only or linear in both $Y_t$ and $Y_{t-1}$ simultaneously as (22) (23) show. Thanks to these affine properties, I am able to construct all the moment conditions in neat analytic expressions of the parameters. To see this, I collect all variables of interest in $\Psi_t = [q_t, \Delta c_t, dp_t, \pi_t, r^f_t, s_t^{pd}, \Delta p_t, r^e_t]^\prime$, which by affine properties, can be written as,

$$\Psi_t = \mu_\Psi + \Phi_\Psi Y_t + \Theta_\Psi Y_{t-1}$$

(25)

therefore, the unconditional mean and variance of the vector $\Psi_t$ will all depend on that of the state vector $Y_t$,

$$E(\Psi_t) = \mu_\Psi + \Phi_\Psi \cdot E(Y_t) + \Theta_\Psi \cdot E(Y_{t-1})$$

(26)
\[ V(\Psi_t) = \Phi_{\Psi} \cdot V(Y_t) \cdot \Phi_{\Psi}' + \Theta_{\Psi} \cdot V(Y_{t-1}) \cdot \Theta_{\Psi}' + \Phi_{\Psi} \cdot C(Y_t, Y_{t-1}) \cdot \Theta_{\Psi}' + \Theta_{\Psi} \cdot C(Y_{t-1}, Y_t) \cdot \Phi_{\Psi}' \]

where \( V \) is notation of the variance-covariance matrix for a random vector and \( C \) denotes the covariance matrix for two random vectors, particularly,

\[
C(Y_t, Y_{t-1}) = \text{cov} \begin{bmatrix}
q_t q_{t-1} & q_t \Delta c_{t-1} & q_t d_{p_{t-1}} & q_t \pi_{t-1} \\
\Delta c_t q_{t-1} & \Delta c_t \Delta c_{t-1} & \Delta c_t d_{p_{t-1}} & \Delta c_t \pi_{t-1} \\
d_p q_{t-1} & d_p \Delta c_{t-1} & d_p d_{p_{t-1}} & d_p \pi_{t-1} \\
\pi_t q_{t-1} & \pi_t \Delta c_{t-1} & \pi_t d_{p_{t-1}} & \pi_t \pi_{t-1}
\end{bmatrix},
\]

\[
C(Y_{t-1}, Y_t) = [C(Y_t, Y_{t-1})]' \quad (27)
\]

which are not symmetric matrices. Because all unconditional moments of \( Y_t \) including \( E(Y_t) \), \( V(Y_t) \) and \( C(Y_t, Y_{t-1}) \) are all readily derivable as functions of parameters thanks to the VAR structure of \( Y_t \) in (8), the moments of vector \( \Psi_t \) which are my moment conditions to use in GMM estimation will consequently be available too as functions of parameters. Each moment in its specific functional expression is presented in Appendix II. Notice also that although my economy is conveniently linear in state vector \( Y_t \), the moments are not necessarily linear in parameters. In fact, they are highly nonlinear in the parameters, posing big challenges for the nonlinear optimization carried out in my GMM estimation.

### 2.1 Moment Conditions

The full structure of (8) and (9) contains a total of 21 parameters which may be combined in a parameter vector \( H \),

\[
H = \{ \mu_c, \mu_d, \mu_p, \rho_q, \rho_{cc}, \rho_{cd}, \rho_{dd}, \rho_{dc}, \rho_{\pi}, \sigma_q, \sigma_{cc}, \sigma_{cd}, \sigma_{dd}, \sigma_{\pi}, \lambda, \beta, \delta, \rho_{2q}, \rho_{2c}, \rho_{2d}, \rho_{2\pi} \}.
\]

Notice that for identification purpose, the long-run mean of the surplus ratio \( q_t \) is fixed at 1, therefore the parameter \( \mu_q \) in (8) is not included in \( H \). I use 23 moment conditions to estimate the
21 parameters. I leave the details of each moment’s functional form to the Appendix II, but I present here what these moment conditions are and explain the reasons they are selected. They can be grouped into the following five rows.

\[ E(\Delta c_t), E(dp_t), E(\pi_t), E(r_t^f), E(s_t^{pd}), E(\Delta p_t) \]
\[ V(\Delta c_t), V(dp_t), V(\pi_t), V(r_t^f), V(s_t^{pd}), V(\Delta p_t) \]
\[ C(\Delta c_t, \Delta c_{t-1}), C(dp_t, dp_{t-1}), C(\pi_t, \pi_{t-1}), C(r_t^f, r_{t-1}^f) \]
\[ C(\Delta c_t, dp_t), C(\Delta c_t, r_t^f), C(dp_t, \Delta p_t), C(\Delta c_t, \Delta p_t), C(\pi_t, \Delta p_t) \]  
\[ C(\Delta c_t, dp_{t-1}), C(dp_t, \Delta c_{t-1}) \]

The first row is simply the means of all observable variables, fundamental and derived. The second row pertains to the variances. I try to match my economy in means and volatilities to the real data. Notice I use the price increment \( \Delta p_t \) to proxy for the equity excess return. One reason is that the price increment constitutes the large majority of equity excess return. Other than that, the excess return in my model is merely the arithmetic sum of price growth, dividend yield and minus risk-free rate. By matching each to the data, the sum of them should fit accordingly.

The third row of moments is to match the autocorrelations of each variable. This intends to enhance the model fit to a higher level of precision. Moller (2009) has not fitted any autocovariance in his moment conditions for GMM. In CC1999, the auto-correlation of consumption growth is fixed at zero and parameters are chosen in order for the theoretical price-dividend ratio autocorrelation to match the actual data, which results in the serial correlations of risk-free rate and surplus ratio also being calibrated. That is because in the CC1999 model, a single parameter \( \phi \) determines the auto-correlations for all three series. I, on the other hand, let data decide the serial correlation of each process, which is no longer governed by a single parameter and...
therefore has the freedom to differ from each other although the estimates of serial correlation parameters of the three series are all fairly high.

The fourth row of moments captures the cross-correlations between the fundamentals and the derived variables as well as that within the fundamentals. And the fifth row accounts for lagged cross-correlations among the fundamentals.

I use these 23 moment conditions because they are valid and relevant moments. Many unused moments will add little information because they are either invalid or redundant moments (see, e.g., Breusch et al. 1999). For example, the moments of covariance between inflation and other state variables are not useful for identifying any parameters given the zero-correlation implied by the model (8). These are invalid moments. The Moments of covariance between a state variable and excess return is redundant, because such covariance is pre-determined by the linear combination of covariance with capital gain, dividend yield and risk-free rate respectively.

The GMM estimation is quite straightforward in two aspects. First, the objective function uses all the unconditional moments of the observed variables only.

I collect the 23 moment conditions from (28) in a vector of sample moments $g_T(H)$, where $T$ is the number of observations. And I estimate the parameter vector $H$ by minimizing the quadratic form:

$$g_T(H)'Wg_T(H),$$

where the weighting matrix $W$ is the optimal weighting matrix, the inverse of the variance-covariance matrix of sample moment conditions:

$$W = \hat{S}^{-1}$$

and where $\hat{S}$ is the sample counterpart of $S$, the covariance of population moment conditions:

$$S = \text{avar} \left\{ \sqrt{T}g_T(H) \right\}$$
where \( avar \) indicates asymptotic variance.

Second, the optimal weighting matrix \( \hat{S}^{-1} \) is free from any parameter dependence. Since my moment conditions \( g_T(H) \) are neatly separable in data and parameters, this separability suggests a spectral density matrix conveniently and solely data based. The spectral density is calculated using the Newey and West (1987) method.

### 2.2 Two-Stage Procedure

For the 21 parameters in 23 moment conditions, with the reference of Appendix II, I examine more closely how the parameters enter into the moment equations. Basically I have two groups of parameters, structural and preference, and two groups of moments, fundamental and derived. Through equation (4), the state variable \( q_t \) and \( \Delta c_t \) enter the pricing kernel. Because the kernel goes into every pricing activities, any bond and stock related moment expressions involve parameters from state process \( q_t \) and \( \Delta c_t \). These moments include almost all first and second moment and cross moment related with endogenous variables \( r_t^f, s_t^{pd}, r_t^{ex} \). I call them derived moments because they are the moments involving derived variables. That said, in the first row of moments, \( E(\Delta c_t), E(dp_t), E(\pi_t) \) depend on parameters \( \mu_c, \mu_d, \mu_\pi, \rho_{cc}, \rho_{cd}, \rho_{dd}, \rho_{dc}, \rho_\pi \) only, but \( E(r_t^f), E(s_t^{pd}), E(r_t^{ex}) \), moments of derived variables or kernel related moments, depend on more than just these parameters. The additional parameters they use are \( \rho_q, \sigma_q, \lambda, \beta, \gamma, \rho_{2q}, \rho_{2c}, \rho_{2d}, \rho_{2\pi} \). The same is true among the second row of moments, where \( V(\Delta c_t), V(dp_t), V(\pi_t) \) depend only on parameters \( \rho_{cc}, \rho_{cd}, \rho_{dc}, \rho_\pi, \sigma_{cc}, \sigma_{dc}, \sigma_{dd}, \sigma_\pi \), but \( V(r_t^f), V(s_t^{pd}), V(r_t^{ex}) \) require more parameters such as \( \rho_q, \sigma_q, \lambda, \gamma, \rho_{2q}, \rho_{2c}, \rho_{2d}, \rho_{2\pi} \). This is not surprising since the moments of state variables depend on only the structural parameters of VAR in \( Y_t \), while
moments of derived variables depend on more than that because the pricing kernel now comes into play, and parameters of unobservable process $q_t$ have to enter now. Therefore moments of fundamental observable state variables use fewer parameters while derived variables are relatively more complicated because they are derived on top of the fundamentals and thus absorb more parameters. So, to summarize as well as to distinguish, I collect structural parameters for fundamental moments in $H_1$ and preference parameters for derived moments in $H_2$,

$$H_1 = \{\mu_c, \mu_d, \mu_\pi, \rho_{cc}, \rho_{cd}, \rho_{dd}, \rho_{dc}, \rho_{\pi}, \sigma_{cc}, \sigma_{dc}, \sigma_{dd}, \sigma_\pi\}$$

$$H_2 = \{\rho_{q}, \sigma_{q}, \lambda, \beta, \gamma, \rho_{2q}, \rho_{2c}, \rho_{2d}, \rho_{2\pi}\}$$

where $\rho_{2q}, \rho_{2c}, \rho_{2d}, \rho_{2\pi}$ are parameters of the loading vector $B = [\rho_{2q}, \rho_{2c}, \rho_{2d}, \rho_{2\pi}]$ from equation (21). So in my 23 moment conditions, fundamental moments depend on only the parameter $H_1$ and derived moments rely on both parameter $H_1$ and $H_2$. This way, I determine which parameters are used in which moment conditions, and leads to the allocation of the entire 23-moment system into smaller sub-systems that help eventually solve the whole system of equations sensibly. The idea is that I narrow down to firstly those basic moments in parameter $H_1$ only, in order to have fewer parameters and fewer moments to start with. Specifically I solve 12 by 12 inside of 21 by 23 first, partially due to a smaller and thus a less complex system of equations to minimize, also that $H_1$ pave the ground where derived moments rely on. $H_1$ enters almost all moment conditions but $H_2$ enters only the moments involving derived variables. Obviously to determine the values of $H_2$ parameters first is not feasible. The 12 by 12 system turns out just-identified. I verify that there exists a unique solution by minimizing the objective function numerically to literally nil, which seems a reasonable guess to start with. To search for a good candidate of $H_2$ I fixate the values of 12 parameters in $H_1$ at just-identified and minimize the remaining 11 moments. $H_2$ found this way along with the previously exactly identified $H_1$
form my starting candidate of parameter estimation. I go from there to further minimize the entire 23 moments in all the parameters $H_1$ and $H_2$.

Digging deeper into the expressions given in Appendix II, I can attain a better picture of why those functions of parameters forming lower dimension sub-systems are indeed self-contained to solve. In details, the moments $V(\Delta c_t)$ and $V(dp_t)$ from row 2, the moments $C(\Delta c_t, \Delta c_{t-1})$ and $C(dp_t, dp_{t-1})$ from row 3, the moment $C(\Delta c_t, dp_t)$ of row 4 and $C(\Delta c_t, dp_{t-1})$ and $C(dp_t, \Delta c_{t-1})$ of row 5 make a 7 by 7 just-identified system. These 7 equations depend on exclusively 7 parameters which are 4 feedback $\rho_{cc}, \rho_{cd}, \rho_{dd}, \rho_{dc}$ and 3 volatility $\sigma_{cc}, \sigma_{dc}, \sigma_{dd}$ parameters. Likewise, the three moments on inflation $E(\pi_t)\ V(\pi_t)\ C(\pi_t, \pi_{t-1})$ make another 3 by 3 self-sufficient system. The two more moments $E(\Delta c_t)$ and $E(dp_t)$ can determine another two parameters $\mu_c$ and $\mu_d$ at the solutions of the earlier 7 by 7 equations. This is how the 12 by 12 system gives rise to the structural parameters, which is just an initial candidate to start minimizing with. The other 9 parameters in addition to the 12 parameters freshly estimated enter into the remaining 11 moments. Unlike the 12 parameters in $H_1$, a good starting guess for the 9 parameters in $H_2$ could be quite difficult to obtain. My method is to try different zones to start with and move along slowly. The difficulty lies in keeping the expressions, such as (20), from diverging. The infinite recursive nature of equation (20) leads to either convergence or divergence. Convergence makes intuitive sense for prices of bond infinitely far away from maturity while divergence simply means absurd bond prices due to wrong parameter values landed on. However, to ensure convergence relies on no shortcut but large quantity of trials within the parameter space. I have tried bounding the values of $H_2$ in a safety zone, but such boundaries can be very hard to determine in a comprehensive fashion. Because the moments are highly nonlinear and discontinuous functions of the parameters, a
single parameter could have multiple areas to explode the expression, not to mention that numerous parameters in $H_2$ all have roles to play. Once a converging candidate is found, I carefully sweep the neighborhood in small steps to locate a local minimum. Many such converging candidates are evaluated, screened and discarded for a final winner.

### 2.3 Data Construction

The in-sample estimation uses US annual data for the period 1927 to 2000. Consumption is measured as expenditures on non-durables and services obtained from the National Income and Product Accounts (NIPA) table 2.3.5. Nominal consumption is converted to real term using the consumption deflator from NIPA table 2.3.4. Real per capita consumption is obtained using the population numbers in NIPA table 2.1.


I have made a few minor modifications in Shiller’s data to meet the needs of my empirical work. The first variable I adjust is the dividend yield. Notice in Shiller’s data, the dividend yield $d_{e_t}$ is not constructed as a simple ratio $D_t/P_t$, but is defined to be $d_{e_t} = ln(1 + D_t/P_{t-1})$. This is a desirable format which my model just happens to need. However in Shiller’s data, the dividend yield concerns actually $D_t$ and $P_{t-1}$, which are the dividend of the year and the stock price at the beginning of that year, namely the stock price prior to the dividend issuance. For equation (19) to
satisfy equity pricing rule in equation (18), the dividend yield must be \( ln(1 + D_t/P_t) \), rather than \( ln(1 + D_t/P_{t-1}) \). The latter would cause error in price derivation. Thus, I amend Shiller’s data to match equity prices after dividend issuance with the corresponding dividend issued.

Another change is made on the data of excess return which Shiller collects in simple arithmetic term \((P_t + D_t - P_{t-1})/P_{t-1}\), but I rearrange that into log return \( ln\left(\frac{P_t + D_t}{P_{t-1}}\right)\) to be consistent with my formula derivation. To see why this correction is needed, please refer to equation (23) where excess return is derived. Notice that my adjustment reduces the data mean of equity return to about 7.2% from 9.2% of Shiller’s figure due to the property that logarithm approximation only works precisely when the rate is small. But historical stock returns are very large and volatile, not uncommon to reach 20%-30% from time to time.
Chapter 3

Empirical Evaluation

Using the estimation strategy outlined earlier, my estimation results are presented in the following tables and figures along with my analysis and evaluations of the model’s empirical applicability.

3.1 Parameters

Table 1 summarizes the parameter estimates. I actually have located a few candidate estimates. After ruling out other estimates, this one is identified as a legitimate candidate. One primary reason the other candidates are discarded is, for example, parameters going out of bound. Some parameters are valid only in limited domains, say $\lambda \in (-1, 1)$ or $|\rho_q| < 1$ to maintain stationary. I discard those points ending outside of the reasonable area. I also remove those
complex roots which although might falsely achieve a better minimum on Matlab program. The estimates need to produce the real-valued solution to \( b \) in equation (21). Candidate estimates that result in complex root of \( b \) is then thrown away.

At this confirmed global minimum, the whole model is rejected at 5% significance level but fail to reject at 1% level. The first column contains mean related parameters, second column feedback parameters, third column volatility parameters and fourth column parameters are related to pricing kernel, agent preference and the loading coefficients of equity price random-walk element. In terms of individual parameter, the majority appears statistically significant. The three mean parameters are statistically different from zero, implying means are significant. The estimates of six feedback parameters appear consistent with the assumption that the four state variables are stationary time series. Feedback \( \rho_q \) at 0.91 appears a very precise estimate and consistent with many habit formation papers that the surplus consumption ratio is a rather persistent process. Recall that CC1999 have this parameter calibrated at 0.87 for the autocorrelation of price-dividend ratio to fit. I do not require this type of linkage because the dividend-price ratio in my model has its own free parameters to fit the serial correlation. Nevertheless the results turn out very close. Recall also that Bekaert et al. (2010) come up with 0.89 and Brandt and Wang (2003) find it in 0.90–0.98 range. The 0.46 \( \rho_\pi \) shows inflation stationary but more rapidly reverting to the long-run mean. \( \rho_{dc} \) is estimated at 0.034 and not statistically different from zero even at 20% level, implying dividend yield process is very much close to an AR(1) process. Basically the own past effect of dividend yield prevails while the effect of past consumption growth on dividend yield is very minor. Recall that long-run mean of surplus consumption ratio \( q_t \) is fixed at 1 for identification purpose, while the long-run mean of consumption growth \( \Delta c_t \) dividend yield \( dp_t \) and inflation \( \pi_t \) are, on the other hand, in the
0.02–0.04 range by actual data. This scale difference helps in some way explain the magnitude discrepancy of volatility in each process, therefore $\sigma_q$ at 0.428, considerably larger than other volatility parameters in column three. The time discount $\beta$ is estimated at 0.97, which is also in line with the conventional belief of a level close to but less than 1. The negative $\lambda$ shows the surplus ratio co-moves negatively with the consumption growth, as expected by the consensus of a countercyclical risk aversion. For the loading coefficient vector $B$ of the random-walk component in equity price, the estimates also appear in high significance and good precision. The four estimates show high $t$-statistic of 6.5, 4.0 and 4.2 except $\rho_2q$ a bit lower of 1.5.

### 3.2 Stock and Bond Moments

In Table 2, the means, the standard deviations and the first order auto-correlations of all observable variables are presented. For each moment, I show the model predicted value in the first row against the actual data value in the second row, along with the standard errors in the third row for benchmark purpose. The standard errors are calculated by standard Delta Method therefore are solely data determined. I see that in the majority of the 21 moments, model predicted value can match quite closely to its data. Notice that even the moments that are absent from my objective function (28), that in other words are not calibrated to necessarily fit, such as the autocorrelations of $spd_t$ and $rex_t$ are able to match up nicely. Specifically, yield spread displays moderate serial persistence which I predict at 0.48 against the actual of 0.56. Price growth and excess return have almost zero persistence and my model predicts that nearly precisely. The fact that my model can capture those extended moments puts me under the impression that my model is on the right track to replicate the economy based on the chosen state variables and the hypothesized pricing philosophy. The most distant prediction appears to be the
standard deviation of risk-free rate, which model implies 2.3% while actual data suggests 3.5%, much more than two standard errors apart from each other. I do not yet have a good explanation for that. Another slightly off-the-target incidence happens with the standard deviation of the inflation, which the model estimates to 3.2%, undershooting the real world figure of 4.3% for a gap of almost two standard errors. A possible explanation for that will shortly be suggested in evaluating the next Table 3.

The equity premium and risk free rate puzzle says for the standard power utility, in order to match the high equity return historically, the risk-free rate has to go unrealistically high. It obviously is not the case in my results. I generate an artificial excess equity return of 6.0%, matching closely to the sample moment of 5.6%, while keeping the nominal risk-free rate on target at 4.4% within one standard error from the actual of 4.8%. However, this is by no means some new achievement here. Because it is alone the result of habit induced time-varying risk appetite, which passes down from CC1999. Therefore I refrain from any further elaborations on these puzzles due to repetitiveness.

Table 3 is a complete array of cross correlations. Keep in mind that of all the 21 co-movements in this table only 5 are manipulated to the data. In other words, the vast majority of the correlations are not controlled intentionally. However again, I see quite many extended moments which do not belong to the 23 moment conditions of (28) nevertheless matching reasonably well to the real world, which could be another evidence that earns the model some credit. Most moment pairs can match to the sign as well as to the magnitude. Some pairs fail to precisely match the magnitude but the keep the sign correct, such as $\text{corr}(r_t^f, dp_t)$, which is un-calibrated. A few have opposite signs but both are close to zero at insignificant magnitude, such as $\text{corr}(\Delta p_t, r_t^f)$, which is again outside of the fitted moments (28).
I notice that a few less satisfactory moment pairs are related with the inflation. Among the six correlations with inflation, $corr(\pi_t, \Delta c_t)$ and $corr(\pi_t, dp_t)$ cannot match at all because they are not designed to. $corr(\pi_t, s^d_t)$ is 1.89 standard error apart in model’s prediction from the actual. This is not too surprising for an un-calibrated moment. $corr(\pi_t, s^d_t)$, $corr(\pi_t, \Delta p_t)$ and $corr(\pi_t, r^e_t)$ are able to match up very closely. The former is also an un-calibrated moment and the latter two are basically the results of their inclusion in the 23 moment conditions. To bridge the overall differences around these correlations with inflation, the model naturally would need a more delicate design of the inflation process. I do present one such enhanced structure in Section 3.9 for interested readers to follow up. I do not pursue it formally here because generally the inflation design stays indifferent to the affine structure in asset prices of my model.

First let us look at what Table 3 shows about the relations among the fundamentals. Within the three visible state variables, dividend yield and consumption growth show a negative connection. The intuition is the following. High dividend yield positively forecasts future high return and economic boom. Because dividend yield is a business cycle forecaster that high dividend yield implies a boom tomorrow and usually a recession today. Thus today’s high dividend yield tends to associate with the low consumption growth of the current period. Therefore a negative correlation is normally expected as in the data, but not a very strong one. It is because the dividend yield and the future return do not go together one to one, only roughly 50% variation of future return is explained by dividend yield. Likewise economic boom and bust do positively co-move with high and low consumption growth but not in perfect linkage. My model correctly captures this relation, although it slightly overshoots at -0.34 vs. -0.26. Also within the state variables, there are some positive correlations in the real data at 0.42 between inflation and consumption growth as well as between inflation and dividend yield at 0.10, but my model
generate none of it at all. This is no surprising due to my minimal design of the inflation process. My specification of inflation as an AR(1) with no noises influence from other processes seems a simplistic setup. One might argue that a more sophisticated inflation process could potentially correct for these wide differences and therefore render the overall model a better fit. However the focus of this article is primarily on the affine prices, particularly the equity prices, which inflation has no direct impact on. I consider its effect on the entire model limited.

As for the covariance between state and derived variables, consumption growth or dividend yield does not have particularly strong records in history to co-move with the risk-free rate or yield spread except a quite strong negative connection between dividend-price ratio and equity return. Understandably, at a low dividend yield, end of period stock price is high relatively to the dividend which results in a relatively large price increment for the year, therefore usually low dividend yields are accompanied by large stock returns. The model generates this quantitatively strong correlation, although it overestimates a little bit at -0.61 against the actual of -0.49. The model conforms to other weak connections quite closely as well. Specifically, I catch both signs correct for consumption growth and dividend yield’s correlations with risk-free rate, the latter of which is an unintended moment and undershoots a little. Also correct is the sign of the weak positive correlation between consumption growth and price growth. Naturally in good days, consumption expansion and stock appreciation have reasons to move in the same direction. For inflation, as pointed out earlier as an underperforming factor causing mixed results, its correlations succeed with risk-free at 0.29 vs. 0.33 and excess return at -0.08 vs. -0.03 but fail with yield spread at -0.43 vs. -0.17, of which the first and last moment are unpicked in (28).
For the correlations among derived variables, the risk-free rate and the yield spread appear quite strongly connected due to the same term structure pricing scheme. My model affirms this high correlation at -0.78 however overshooting the real value -0.60. Risk-free rate and excess return appear fairly loosely linked and my model mimics that correctly. Term spread correlates weakly with capital gain at 0.24 in the dataset while my model agrees to a low connection but fails to generate a quantitatively close enough prediction, leaving a two standard error difference.

### 3.3 Time Variation

Table 4 explains the variances of the endogenous variables contributed by each of the four state variables. Variation in risk-free rate is mostly driven by the inflation and the risk aversion, accounting for about 40% and 60% each. It is consistent with the model’s findings in Table 3 that risk-free rate bears no significant correlation with consumption growth or dividend yield, but having quite a larger correlation of 0.29 with the inflation. Same is true with the price growth. The majority of variations in capital gains is due to the variance of dividend price ratio and risk aversion, accounting for 47% and 38% respectively, while consumption growth and inflation rate contribute only 10.8% and 5.8%. This allocation conforms to the capital gain’s large correlation of -0.61 with dividend price ratio as well as minor correlations with consumption growth and inflation at 0.11 and 0.21 respectively. Past state variables contribute very little to the variation of price growth. Notice I derive the price growth $\Delta p_t$ in its real rather than the nominal term using real discount factor $m_t$ that excludes the inflation element. Nevertheless price growth still has a small correlation with the inflation due to that the random walk component in the price $p_t^\theta$ has a nonzero connection with the inflation.
3.4 The Unobservable

Because the latent variable, risk aversion \( q_t \) is not directly observable, I provide a separate Table 5 to analyze its property and connections with other variables. Surplus ratio is negatively correlated with consumption growth at a coefficient of -0.39, consistent with the negative \( \lambda \) estimated in Table 1. So both the conditional and unconditional correlation between \( q_t \) and \( \Delta c_t \) are negative, reinforcing the consensus that risk aversion is counter-cyclical. Campbell-Cochrane model specifies this correlation at -1. Surplus ratio has quite a large positive correlation coefficient 0.688 with risk-free rate, consistent with the finding that 60% of variation in risk-free rate is due to variation of surplus ratio as reported in Table 4.

The dynamics of the risk aversion with respect to the term structure displayed in Table 5 are more consistent with the finding of Brandt and Wang (2003), whose model reveals that risk aversion is highly correlated with yields of almost all maturities. Likewise, my correlations of risk aversion with 1 year and 10-year bond yields are noticeably strong. But this finding appears not so consistent with that of Bekaert et al. (2010), which discover low correlations between \( q_t \) and bond yields, specifically 0.21 for one year bond. Because I borrow the setup of Bekaert et al. (2010), particularly the identical specification of the surplus ratio \( q_t \), it is tempting to compare the outputs of \( q_t \) on the two models. Most noticeably, \( q_t \) is seen in Bekaert et al. (2010) having a large correlation 0.93 with dividend-price ratio and small correlation of 0.21 with risk-free rate (see Table 5), while my results differ quite vastly at 0.014 and 0.688 respectively. Due to the unobservable property of risk aversion and psychological nature of habit formation, it could remain debated as for which output is closer to reality. Nevertheless, Brandt and Wang (2003) suggests that risk aversion and all maturities of bond yields are highly correlated (see Brandt and Wang 2003 Table 5). Particularly, his model predicted risk aversion time series has a 0.84
correlation with actual one-year bond yield and 0.98 with model simulated yields. Brandt and Wang (2003) seems more supportive than Bekaert et al. (2010) in terms of my 0.69 finding on correlation with the short rate.

Other than the risk-free rate, my correlation with consumption growth at -0.39 is more evident than Bekaert et al. (2010) at -0.15 although both confirm the counter-cyclical risk aversion. Correlations with excess returns are weak in both ours and Bekaert et al. (2010) model, which is a consensus capturing the lack of persistence in annualized equity return. Also weak are the correlations with yield spreads in both models, which however expose a discrepancy. The weak correlation in my results conveys a quite large correlation between risk aversion and the yield of long-term bond, while for Bekaert et al. (2010) it implies quite a weak one.

The distribution properties of Relative Risk Aversion are outlined in the first row of Table 6. Recall that for identification purpose I have pinned the long-run mean of surplus ratio $q_t$ at unity, therefore giving the risk aversion a long-run mean of 0.81 at my point estimates. The distribution of risk aversion is right skewed, getting very high but in small chances, staying low in majority of the time. Risk aversion has interquartile range of (0.39, 1.3) with median at 0.63. In 90% of the time risk aversion is no more than 3.8 and only shoot up to over 40 in less than 1% chance.

3.5 **Comparison with Related Models**

To find out more about the correlation of risk aversion with the risk-free rate, I outline the findings of more peer models in Table 7. By CC1999, this correlation is destined a perfect
correlation\textsuperscript{8} if $b \neq 0$ because recall that the real risk-free rate is linear alone in surplus ratio, which is the only state variable. Moller (2009) estimates $b$ at a negative value which suggests a -1 correlation. But his estimation of $b$ is imprecise and the term structure output thus turns counterfactual, which renders his results hardly usable. Wachter (2006), whose overall outputs fit data well, calibrates $b$ at positive 0.011 to imply a perfect positive correlation at 1 for real risk-free rate. For nominal risk-free rate, due to additional state variable introduced, including inflation as one, the correlation naturally drops lower from 1. Wachter (2006), with consumption and inflation added as two more state variables, finds the nominal short rate still strongly correlated with risk aversion which I infer at a level around 0.8-0.9. Wachter (2006) provides no specific figures for that coefficient but offers graphic plots documenting the relationship that I infer from. Brandt and Wang (2003) also appoints consumption and inflation as the two added state variables but construct the two series differently and his model when estimated with bond moments only concludes a strong correlation of nominal short rate with risk aversion in 0.92-0.98 range. His results are robust to using bond and equity moments together. Bekaert et al. (2010) who adds a fourth state variable, dividend-consumption ratio, however finds the nominal short rate only 0.21 correlated with risk aversion. My result of 0.688 supports more to the works other than Bekaert et al. (2010).

CC1999 by numerical method establishes that $P/D$ is nearly entirely linear in surplus ratio $S$, an almost perfectly positive correlation (Figure 1). Wachter (2006) and Moller (2009) simply reiterate the result because they essentially use the same model. This result appears well supported by Bekaert et al. (2010) that shows a high 0.93 correlation between $D/P$ and $q$ (Table A), where $q=-s$. On the other hand, my result of 0.014 correlation contrasts hugely with all of

\textsuperscript{8} Because RRA is driven by surplus ratio $q$ or $s$ alone, I usually loosely refer the ratios as the risk aversion. So strictly, the prefect correlation here means that between real risk-free rate and $s$. 

37
them. Are my results far away from the truth? Actually the finding of an almost linear relation demonstrated by Figure 1 may not be that ironclad. By feeding $s$ with real consumption noises, the Campbell-Cochrane model implied surplus ratios in time series do not move one on one with the actual price-dividend ratio, conflicting with the prediction of Figure 1. Moller (2009) points out this contradiction in his Fig 7. CC1999 themselves are candid about it in their Fig 9. Same technique can apply to show that CC1999 prediction of real risk-free linear in $s$ also may requires scrutiny. If they were correct about the correlation of risk aversion with the real risk-free and dividend yield, one should infer that real risk-free and dividend price ratio must be close to perfectly correlated, but data suggests about only 0.2. Either that CC1999 is inaccurate about one of the two relations or both correlations deviate from the truth.

3.6 Time Series Movement

I have so far seen the model mimic quite closely to the sample moments. In other words the model can match the mean level of last century’s data. One might be eager to know what about the time series fluctuations. Let us next see how it performs in terms of depicting the economy’s evolution path over the years. To form the time series of model predicted fundamentals and endogenous variables as well as to reveal the hidden values of surplus ratio $q_t$ year by year, I use the Kalman filter for the task. The idea is to filter what’s unobservable from the observables. I extract underlying state variables from the observable or measureable dataset, and then use the extracted or filtered state variables to construct the endogenous variables and the entire economy. I collect the measurable variables in $Z_t = [\Delta c_t, d_p t, \pi_t, r_t^f, s_t^{pd}, \Delta p_t]'$, by equation (25), $Z_t$ can be written as

$$Z_t = \mu_z + \Phi_z Y_t + \Theta_z Y_{t-1} + \text{measurement error} \quad (29)$$
which, jointly with equation (10), forms the standard two-equation ground for applying Kalman filter. Measurement error is simply assumed to be 5% of the sample variance. From the observable \( Z_t \) over a certain period, accounting for the measurement error, Kalman filter can function to extract out the state vector \( Y_t \) for that period. In order to use the standard linear Kalman filter of Harvey (1989), I slightly rewrite (29) to convert lag state vector into current ones,

\[
Z_t = \mu_z + (\Phi_z + \Theta_z A^{-1})Y_t - \Theta_z A^{-1}(\Sigma_F F_t + \Sigma_H)\varepsilon_t + \text{measurement error} \tag{30}
\]

where the parameter notations newly appearing here are merely just a matter of compact representation in terms of the original 21 parameters collected in \( H \).

With the state vector extracted by Kalman filter and therefore endogenous variables plotted, the time series property of the model is uncovered in a few figures below. The latent variable surplus ratio \( q_t \) is graphed for the data sample period in Figure 2. Because surplus ratio wholly dominates risk preference, graph shows certain years of high and low risk aversion. The peaks in early 1930s and 1980s appear consistent with the timing of recessions as conventional wisdom believes that the investors are more risk averse when economy is bad. However the height of sharp rises in surplus ratio is not fully justified in proportion to the severity of economic downturns. The 1930s depression, usually considered worse, does not spur risk aversion as strong as the early 1980s recession.

I also present the distribution of surplus ratio \( q_t \) from simulation in Figure 3. \( q_t \) is almost always positive as it should be but right skewed. It is most densely distributed between 0 and 1, could go as high as 8 but only very occasionally, exhibiting long tail. This is consistent with the distribution property of \( s_t \) in CC1999 and \( q_t \) in Bekaert et al. (2010). Because risk aversion is increasing in \( q_t \), it inherits the distribution pattern which I document in percentile in first row of
Table 5. One concern is that $q_t$ can get negative in small chances, which implies minus habit level, a situation making no practical sense. On one hand I want to strictly rule out minus $q_t$ which will require mean reversion process to be designed with a stop-loss abrupt point, usually a piece-wise defined function for the square root part of the formulation (8). The sensitivity function $\lambda(s_t)$ designed in CC1999 is an example that works to guarantee positivity. But the change of design could complicate the analytics of moment conditions for the state vector $Y_t$ resulting in closed-form analytics unavailable or not tractable. The chance of $q_t$ going negative, per my simulation, is less than 5%, which I consider immaterial. I therefore keep the simple square root format unchanged.

Chronically risk aversion is graphed in Figure 4 for the sample period. Risk aversion time series as expected resembles that of surplus ratio $q_t$, shooting up considerably in the 1930s and the 1980s to a level of 15 to 30 while remains relatively low below 5 in majority of the time. Risk aversion values predicted over the last century are in line with the distribution summarized in table 5. Interestingly, Brandt and Wang (2003), Bekaert et al. (2010) and my model all produce the time-series of RRA or surplus ratio $q$ quite similarly for the period after 1960s. Risk aversion rises from1960s to peak in early 1980s and then goes downhill. However for the period before 1960s, my model and Bekaert et al. (2010) do not quite coincide with each other.

Predicted risk-free rate, yield spread, and equity return are plotted in Figure 5 to 7 and are paired with the actual to contrast. Because my model asserts that these derived variables are all linear in state vectors and I already have the four state variables extracted out, in this way I straightforwardly obtain the derived variables in time series. Predicted risk-free rate appears quite close to the actual evolution in the last century, there are slight over and understatement in some years, but capture the overall time series trend accurately enough. So is the predicted yield
spread. Because my spread is the difference between 10-year bond yield and risk free, I can say with certainty that model’s prediction of 10-year bond yield is a success as well, although its graph is not provided to avoid redundancy. Graphically, I can tell that the risk aversion in Figure 2 and the one-year bond yield in Figure 5 look fairly alike, conforming to the high correlation of 0.688 reported in Table 5 and 6. Brandt and Wang (2003) have quite the similar finding, but not so in Bekaert et al. (2010).

3.7 Out-of-Sample Test

As shown earlier in Table 2, 3 and 4, my model can fit the base moments as well as many moments not explicitly required to fit in the estimation stage, therefore it is already a successful joint stock-bond pricing model. Nevertheless, I want to further examine the predictive ability of my framework out of sample and relative to benchmark models. By far the model estimation and variable filtering use the sample period 1927-2000. I then perform a straightforward out-of-sample test using estimated model to predict for period after 2000. In Figure 8, the predicted (filtered) surplus ratio appears to capture the latest rising in risk aversion which corresponds to the most recent financial crisis. However, the timing is a little embarrassing. Actual crisis did not happen until 2008, but my model predicts a peak in risk aversion in 2006. A possible excuse links to the decreasing correlation between $q$ and consumption growth $∆c_t$ as a result of out-of-sample participation. This correlation is estimated 0.391 in-sample and decreases to 0.35 when mapping year by year for 1927-2000. Now it falls to 0.295 when extra years of one on one mapping further reduces the degrees of freedom.

Next, I show a contrast of out-of-sample predictive power between CC1999, the benchmark model, and my model. In Figure 9, I reproduce price-dividend ratio prediction by
CC1999 up to year 2013 using exactly their approach of feeding the actual consumption data and precisely their calibrated parameters in-sample up till 1995. I discover increasingly wide prediction gap ever after early 1990s. Meantime, in Figure 10 I generate my corresponding predictions using in-sample estimates of Table 1. Notice that my model can only derive prices while CC1999 can only derive price-dividend ratio. To be comparable, I divide my predictions of prices by actual historical dividend to form my price-dividend ratio. Visual interpretation appears to favor my prediction of stock price dynamics, especially after 1990s, where my predictions catch the fluctuations a lot more closely than CC1999. The causes of dramatic rises in price-dividend ratio after 1990s remain still arguable, and may include, for example, dividend policies transform briefly implied by CC1999.

### 3.8 Long-Horizon Return Regression

Predictions of time-series in equity returns appear the least successful as the Figure 7 demonstrates. There are spikes or abysses in the actual curve that prediction fails to follow and even when the zigzag is captured in shape and timing the level is not accurately close. Perhaps the format that the excess return or price growth involves two consecutive state vectors to co-play, unlike risk-free rate which is linear in only the current state vector, makes it subtler to catch up closely to the actual returns. After all, the linear Kalman filter is designed to be most applicable for cases where the observed variables are linear only in contemporary state vector. Nevertheless, my predicted returns can still work reasonably well for the long-horizon regressions.

Table 8 presents the long-horizon regression of equity excess returns on dividend yields in model predicted and actual data. I do see the well-known pattern documented by Campbell and
Shiller (1988) and Fama and French (1988) showing up. The regression coefficients are all positive and increasing along horizons. Higher dividend yields imply higher future returns. The forecasting power rises with horizon to substantial level as indicated by R square. The model’s prediction however exhibits less significantly in R square although the pattern remains evidently discernible.

The real intuitive reasons why the dividend-price ratios and some other macroeconomic variables\(^9\) can forecast future long-horizon returns is that they can forecast or are correlated with business cycles. These variables are usually quite persistent, refusing to frequently revert to the long-run mean, and stay in the vicinity of a prior year level for duration of periods, therefore they coincide with the business cycles which also only change after some extended years. My model is able to capture this cyclicality. Judging by auto-correlations, the surplus ratio \(q\) (0.91) and risk-free (0.74) are fairly persistent and consumption growth \(\Delta c_t\) (0.43) is moderately persistent. As expected, they turn out quite heavily correlated with one another, say \(\rho(q_t, rf_t)\) at 0.69 and \(\rho(q_t, \Delta c_t)\) at -0.39. On the other hand, annualized equity excess return (-0.01) or price growth (0.02) is hardly persistent and thus does not correlate with risk aversion or any other business cycle forecasters. Therefore I see examples like \(\rho(q_t, \Delta p_t)\) at 0.149, \(\rho(q_t, rex_t)\) at 0.07, \(\rho(\Delta c_t, \Delta p_t)\) at 0.12 and \(\rho(\Delta c_t, rex_t)\) at 0.11. Persistence of equity return is barely evident till accumulated over longer span. I have established that risk aversion is counter-cyclical and dividend yields forecast business cycles. It seems natural to infer that risk aversion and dividend yield should strongly co-move. But my model shows that correlation to be 0.014. The catch here is the following, it is true that the risk aversion tends to counter move with economic climate, but not perfectly, the same is true that dividend yield does forecast business cycle, but only around

\(^9\) See, for example, Lettau and Ludvigson (2001).
50% or less in R square. These two mild correlations are not enough to imply a strong connection between the risk aversion and the dividend yield. Therefore it is not surprising to see a low correlation generated from my model.

Another variable not to be ignored that also shares the cyclical properties is the Sharpe ratio. Like other cyclical quantities, Sharpe ratio is persistent. It is usually believed increasing in risk aversion. Recall that it is constructed by CC1999 to decease in \( s \) and hence increase in \( q \) in my model. So the risk aversion rises when economy is down and the Sharpe ratio drifts higher. Expected return, the numerator in Sharpe ratio, could also be increasing in risk aversion. According to CC1999, expected return and expected volatility both are decreasing in \( s \), hence increasing in risk aversion. Therefore expected returns vary over business cycle and people expect higher risk premium to hold stocks during the recessions. However the annual realized return hardly has any cyclicity despite the fact that the expected return of a single year does. It is because the realized return is the net effect of expected and unexpected return. The combined unconditional return tends to retain little cycicality and minimal correlation with risk aversion.

### 3.9 Variant Specification

Under the general framework (10) and (11) of Joint Stock-Bond Pricing, a few variants to the one used in this study, (8) and (9), are possible. These alternative specifications potentially aim at a modified inflation design to enhance the overall data fitting performance. Particularly motivated by Brandt and Wang (2003) that risk aversion may react to news about inflation, and by Bekaert et al. (2010) that intricate modeling of the inflation process is a possible candidate for advancement, I consider some variants (31) and (32) thereafter.
The impact of inflation risk on asset prices, particularly equity prices, is one of the long-standing questions in finance. A traditional view is that unlike nominal bonds, claims to real assets, such as stocks, serve as a hedge against inflation. However, empirically, shocks to expected inflation depress both stock and bond prices (see, e.g., Fama and Schwert 1977; Bekaert and Wang 2010). A study on inflation and asset prices is not the focus and is beyond the scope of this thesis. But my model can at least shed light on the dynamics between unexpected inflation and risk appetite, the preface to the study of inflation and asset prices. Very few research has provided direct evidence for the response of risk appetite to the inflation news. This is either because risk appetite is not directly measurable or many models are built on constant relative risk aversion (CRRA) utility. Eraker et al. (2016) investigate the pattern of inflation non-neutrality: the asset pricing implications of inflation risk disparity on durable and non-durable consumption, but lack examination in risk preference. Brandt and Wang (2003) find evidence for the hypothesis of strong and positive response of risk aversion to unexpected inflation news.

3.9.1 Specifications

\[ q_{t+1} = \mu_q + \rho_q q_t + \sigma_q \sqrt{q_t} (\sqrt{1 - \lambda^2} \varepsilon^q_{t+1} + \lambda \varepsilon^c_{t+1}) \]

\[ \Delta c_{t+1} = \mu_c + \rho_{cc} \Delta c_t + \rho_{cd} d p_t + \sigma_{cc} \sqrt{q_t} \varepsilon^c_{t+1} \]

\[ d p_{t+1} = \mu_d + \rho_{dd} d p_t + \rho_{dc} \Delta c_t + \sigma_{dc} \sqrt{q_t} \varepsilon^c_{t+1} + \sigma_{dd} \varepsilon^d_{t+1} \]  \hspace{1cm} (31)

\[ \pi_{t+1} = \mu_\pi + \rho_\pi \pi_t + \sigma_\pi \varepsilon^\pi_{t+1} + \sigma_{\pi e} \sqrt{q_t} \varepsilon^e_{t+1} \]

The specification (31) makes change in the inflation process by adding a consumption noise \( \varepsilon^c_{t+1} \). Consequently, the inflation could correlate with three other state variables both conditionally and unconditionally. The one additional parameter introduced is \( \sigma_{\pi e} \). Extra moment conditions are needed for identification of more parameters. Fortunately, there are
abundant moment conditions available for use. For instance, the three visible state variables give the 3 by 3 variance-covariance matrix as well as the first-order auto-covariance matrix, making a total of 15 moment conditions. That should be more than enough to identify the added parameters.

By Brandt and Wang (2003), rising unexpected inflation increases agent’s risk aversion. This relation is governed by parameters $\lambda$ and $\sigma_{pc}$ in specification (31).

\[
q_{t+1} = \mu_q + \rho_q q_t + \sigma_q \sqrt{q_t} (\sqrt{1 - \lambda^2} \varepsilon_{t+1}^\pi + \lambda \varepsilon_{t+1}^c)
\]

\[
\Delta c_{t+1} = \mu_c + \rho_{cc} \Delta c_t + \rho_{cd} dp_t + \sigma_{cc} \sqrt{q_t} \varepsilon_{t+1}^c
\]

\[
dp_{t+1} = \mu_d + \rho_{dd} dp_t + \rho_{dc} \Delta c_t + \sigma_{dc} \sqrt{q_t} \varepsilon_{t+1}^c + \sigma_{dd} \varepsilon_{t+1}^d
\]

\[
\pi_{t+1} = \mu_\pi + \rho_{\pi\pi} \pi_t + \sigma_{\pi\pi} \sqrt{q_t} \varepsilon_{t+1}^\pi + \sigma_{\pi c} \sqrt{q_t} \varepsilon_{t+1}^c
\]

System (32) presents another variant, which resembles more closely that of Brandt and Wang (2003) in the spirit that risk aversion is only driven by consumption growth risk and inflation risk, but no other factors. Therefore specification (32) discards $\varepsilon_{t+1}^q$, the innovation inherent to surplus ratio. For the hypothesis of Brandt and Wang (2003) to hold, the signs and levels of three parameters $\lambda$, $\sigma_{\pi\pi}$ and $\sigma_{pc}$ are crucial.

3.9.2 Findings

I use the same econometric technique for estimating the system (32). Now the GMM estimation uses 24 moment conditions to identify 22 parameters. For identifying the additional volatility parameter $\sigma_{pc}$, I add another moment condition, the covariance between consumption growth and inflation $C(\Delta c_t, \pi_t)$. Major empirical results are presented in Table 9 to Table 11.
Estimate of $\lambda$ in Table 9 confirms the negative connection between risk aversion and consumption growth, as in the main model. More importantly, a positive conditional correlation of 0.32 between inflation and risk aversion can be calculated from estimated volatility parameters. This confirms, on average, a positive response of risk aversion to unexpected inflation news, lending support to the finding of Brandt and Wang (2003).

However, inconsistent with the findings of Brandt and Wang (2003) is the magnitude of such response. They conclude that the sensitivity of relative risk aversion to inflation shock is more statistically significant and economically important relative to consumption growth shock as a source for time-varying risk aversion. Their empirical evidence reveals that the effect of a one standard deviation inflation shock on risk aversion is approximately 17 times that of a one standard deviation consumption growth shock. In my results, risk aversion varies conditionally in response to inflation and consumption growth at the correlation coefficients of 0.32 and -0.76 respectively, and unconditionally at 0.21 and -0.41 (see Table 10) respectively. My results point at consumption growth shock as a more economically influential driver for aggregate risk aversion, opposite to the claim of Brandt and Wang (2003).

3.10 Concluding Remarks

The thesis develops a consumption-based affine asset pricing model that simultaneously prices both bond and equity in a coherent formality. The model inherits the time-varying risk aversion by form of surplus ratio of CC1999 and resembles the setup of Bekaert et al. (2010). But rather than finding the expressions for price-dividend ratio, I derive solutions for the equity prices themselves. Few papers have ever presented stock prices in an affine format. The theoretical model alone has merits in several aspects: (i) tractability, and (ii) generality. All the
asset prices are now exponential affine in the state vector. This ease of closed-form appearance allows almost all variables in the economy, fundamental or derived, and their moments, first-order or second-order, conditional or unconditional, to have tractable analytical expressions. This simple pricing structure avoids the endless recursion normally required in equity prices or the numerical approximation approach needed when the functional form of one variable in terms of another is not obtainable, hence lessens computational burdens particularly for numeric optimization in parameter estimation. As concluded by Mamaysky (2002), a less tractable model would underperform in this regard supposing that both models are equally eligible in empirical fit. Second, my model is designed more flexibly in the fundamental processes and therefore can theoretically be closer to a full-scaled realization of the actual economy.

The empirical results generally support an economy characterized by the four state variables I propose and the pricing methodology the economy is built upon. Specifically, when fed with the actual data, the model can reasonably match a large quantity of data moments, which include more than forty means, standard deviations, autocorrelations and cross-correlations. The variance decomposition analysis uncovers that variation in bond yields is largely driven by the variance of inflation and risk aversion while volatility in equity returns is mainly attributable to the variance of dividend yield and risk aversion. When all state variables are filtered out, the model-implied time series evolve rather similarly to the real process. The model is also able to replicate the long horizon returns regression. Comparisons with several peer models show that my model in some aspects outperforms, say more plausible connections of risk aversion with dividend yield and risk-free rate. Additionally, I let the model predict out of sample, and I find some realistic fit in cyclical behaviors and notice that my model outperforms CC1999.
Appendix

I Proof of Joint Stock-Bond Pricing

The proof of joint stock-bond pricing provided in this appendix I is for the general case, of which my four-state-variable model represents one particular example. The general case is,

\[ Y_{t+1} = \mu + AY_t + (\Sigma_F F_t + \Sigma_H) \epsilon_{t+1} \]  \hspace{1cm} (1)
\[ F_t = (\| \omega + \Omega Y_t \|) \odot I \] \hspace{1cm} (2)

Specifically, for my particular example developed in the main text, I collect the four state variables in vector \( Y_{t+1} = [q_{t+1}, \Delta c_{t+1}, dp_{t+1}, \pi_{t+1}]' \). Thus the parameters \( \mu \) and \( \omega \) are 4×1 vectors, parameters \( A, \Sigma_F, \Sigma_H \) and \( \Omega \) are 4×4 matrices. \( \epsilon_{t+1} \) is 4×1 vector of zero mean unit variance i.i.d. noise: \( \epsilon_{t+1} \sim iid(0, I_4) \). \( \| \| \) denotes element-wise square root operator. \( \odot \) denotes element-wise multiplication. Apparently \( F_t \) is of size 4×4 and represents the square-root components in innovations of equation (1). In the same spirit, log real pricing kernel can be written as,

\[ m_{t+1} = \mu_m + A_m' Y_t + (\Sigma_{Fm} F_t + \Sigma_{Hm}) \epsilon_{t+1} \] \hspace{1cm} (3)

where parameter \( \mu_m \) is 1×1, and parameters \( A_m, \Sigma_{Fm}, \Sigma_{Hm} \) are 4×1 vectors in my model. Remember 4×1 parameter vector could become \( k \times 1 \) if the general model has \( k \) state variables.

Some restrictions are imposed on the parameters,

\[ \Sigma_F F_t \Sigma_H' = 0 \] \hspace{1cm} (4)
These restrictions rule out the possibilities of square-root state variable terms in the conditional second moments, therefore guaranteeing the affine property in the recursive derivations going forward. As a side note, my model set out in the main text is merely a special example of the general case (1) and (2), hence it clearly satisfies the general restrictions (4), (5), (6) and (7).

I provide the following details to link the general structure (1) (2) (3) to the special case examined throughout the main text:

\[
\Sigma_H = \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & \sigma_{dd} & 0 \\
0 & 0 & 0 & \sigma_\pi
\end{bmatrix}, \quad \Sigma_F = \begin{bmatrix}
\sigma_q\sqrt{1 - \lambda^2} & \sigma_{q\lambda} & 0 & 0 \\
0 & \sigma_{cc} & 0 & 0 \\
0 & 0 & \sigma_{dc} & 0 \\
0 & 0 & 0 & \sigma_{cc}
\end{bmatrix}
\]

\[
\mu = \begin{bmatrix}
\mu_q \\
\mu_c \\
\mu_d \\
\mu_\pi
\end{bmatrix}, \quad A = \begin{bmatrix}
\rho_q & 0 & 0 & 0 \\
0 & \rho_{cc} & \rho_{cd} & 0 \\
0 & \rho_{dc} & \rho_{dd} & 0 \\
0 & 0 & 0 & \rho_\pi
\end{bmatrix}, \quad \Omega = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}, \quad \omega = \begin{bmatrix}
0 \\
0 \\
0 \\
0
\end{bmatrix}
\]

\[
\mu_m = \ln(\beta) - \gamma \mu_c + \gamma \mu_q, \quad A_m = \gamma \begin{bmatrix}
\rho_q - 1 \\
0 \\
0 \\
0
\end{bmatrix}, \quad \Sigma_{Hm} = \begin{bmatrix}
0 \\
0 \\
0 \\
0
\end{bmatrix}, \quad \Sigma_{Fm} = \gamma \begin{bmatrix}
\sigma_q\sqrt{1 - \lambda^2} & 0 & 0 & 0 \\
-\sigma_{cc} + \lambda \sigma_{qq} & 0 & 0 & 0
\end{bmatrix}
\]

The following Lemma applies to the general framework.

**Lemma 1** The conditional expectation of an exponential affine function of the state vector plus pricing kernel is exponential affine in the current state vector,

\[
E_t[\exp(a + c'Y_{t+1} + d'Y_t + m_{t+1})] = \exp(\mu + g'Y_t) \tag{8}
\]

**Proof of Lemma 1**
By lognormality,

\[ E_t[\exp(a + d'Y_t + c'Y_{t+1} + m_{t+1})] = \exp\left(a + d'Y_t + c'E_t[Y_{t+1}] + E_t[m_{t+1}] + \frac{1}{2}(V_t[c'Y_{t+1}] + V_t[m_{t+1}] + 2C_t[c'Y_{t+1}, m_{t+1}]\right) \tag{9} \]

Using the equations (1) to (3), the terms from RHS of the (9) are,

\[ c'E_t[Y_{t+1}] = c'(\mu + AY_t) \tag{10} \]
\[ E_t[m_{t+1}] = \mu_m + A'_mY_t \tag{11} \]
\[ V_t[c'Y_{t+1}] = c'(\Sigma_F F_t + \Sigma_H)(\Sigma_F F_t + \Sigma_H)'c = c'\Sigma_F F_t F'_t \Sigma_F'c + c'\Sigma_H \Sigma_H'c \]
\[ = [(\Sigma_F'c)\otimes(\Sigma_F'c)]'(\omega + \Omega Y_t) + c'\Sigma_H \Sigma_H'c \tag{12} \]

where the derivation uses restriction (4);
\[ V_t[m_{t+1}] = [\Sigma_{Fm} \otimes \Sigma_{Fm}](\omega + \Omega Y_t) + \Sigma_{Hm} \Sigma_{Hm} \tag{13} \]

which is derived similarly to (12) and uses restriction (5);
\[ C_t[c'Y_{t+1}, m_{t+1}] = c'[(\Sigma_{Fm} \otimes \Sigma_F)(\omega + \Omega Y_t) + \Sigma_H \Sigma_{Hm}] \tag{14} \]

which uses restrictions (6) and (7).

Substitute back each term:

\[ E_t[\exp(a + d'Y_t + c'Y_{t+1} + m_{t+1})] = \exp\left(a + d'Y_t + c'(\mu + AY_t) + \mu_m + \right. \]
\[ A'_m Y_t + \frac{1}{2}[(\Sigma_F'c)\otimes(\Sigma_F'c)]'(\omega + \Omega Y_t) + \frac{1}{2}c'\Sigma_H \Sigma_H'c + \frac{1}{2}[\Sigma_{Fm} \otimes \Sigma_{Fm}]'(\omega + \Omega Y_t) + \]
\[ \frac{1}{2}\Sigma_{Hm} \Sigma_{Hm} + c'(\Sigma_{Fm} \otimes \Sigma_F)(\omega + \Omega Y_t) + c'\Sigma_H \Sigma_{Hm} \right) \tag{15} \]

Collect the terms:

\[ g^c = a + c'\mu + \mu_m + \frac{1}{2}[(\Sigma_F'c)\otimes(\Sigma_F'c)]'\omega + \frac{1}{2}c'\Sigma_H \Sigma_H'c + \frac{1}{2}[\Sigma_{Fm} \otimes \Sigma_{Fm}]'(\omega + \Omega Y_t) + \]
\[ \frac{1}{2}\Sigma_{Hm} \Sigma_{Hm} + c'(\Sigma_{Fm} \otimes \Sigma_F)(\omega + \Omega Y_t) + c'\Sigma_H \Sigma_{Hm} \]
\[ g' = d' + c'A + A'_m + \frac{1}{2}[(\Sigma_F'c)\otimes(\Sigma_F'c)]'\Omega + \frac{1}{2}[\Sigma_{Fm} \otimes \Sigma_{Fm}]'\Omega + c'(\Sigma_{Fm} \otimes \Sigma_F)\Omega \]
The Lemma 1 is used in the following proof of proposition 1 and proposition 2.

Proof of Proposition 1

Nominal bond prices must satisfy,

$$P_{n,t}^b = E_t[\exp(\hat{m}_{t+1})P_{n-1,t+1}^b]$$  \hspace{1cm} (16)$$

where $\hat{m}_{t+1} = m_{t+1} - \pi_{t+1}$ is the log nominal pricing kernel. The price of a nominal bond maturing in one period of time is,

$$P_{1,t}^b = E_t[\exp(m_{t+1} - e^\pi_Y t_{t+1})] \quad \text{and} \quad r_t^f = -\ln(P_{1,t}^b)$$  \hspace{1cm} (17)$$

where $e^\pi_Y = [0 \ 0 \ 0 \ 1]$ which extracts the state variable from the state vector.

By Lemma 1, it follows that

$$p_{1,t}^b = a_1^c + a_1^r Y_t$$  \hspace{1cm} (18)$$

where $a_1^c$ and $a_1^r$ follow Lemma 1 with $a = 0$, $c = -e_\pi$ and $d = 0$. The risk-free rate is

$$r_t^f = -p_{1,t}^b = -a_1^c - a_1^r Y_t$$

Now for induction purpose, I let

$$p_{n-1,t}^b = a_{n-1}^c + a_{n-1}^r Y_t$$  \hspace{1cm} (19)$$

and plug it into equation (16). This yields

$$P_{n,t}^b = E_t[\exp(m_{t+1} - e^\pi_Y t_{t+1} + a_{n-1} + a_{n-1}^r Y_{t+1})] = \exp(a_n^c + a_n^r Y_t)$$

where the second equal sign uses the Lemma 1 again, with the following recursive coefficients,

$$a_n^c = a_{n-1}^c + (a_{n-1}^r - e_\pi^r)\mu + \mu_m + \frac{1}{2}[(\Sigma^F(a_{n-1} - e_\pi))\otimes(\Sigma^F(a_{n-1} - e_\pi))] \hspace{1cm} (20)$$

$$e_\pi)']\omega + \frac{1}{2}(a_{n-1}^r - e_\pi^r)\Sigma_H\Sigma^F_H(a_{n-1} - e_\pi) + \frac{1}{2}[\Sigma^F_{im}\Sigma^F_{im}]'\omega + \frac{1}{2}\Sigma^H_{im}\Sigma^H_{im} + (a_{n-1}^r - e_\pi^r)(\Sigma^F_{im}\Sigma^F)\omega + (a_{n-1}^r - e_\pi^r)\Sigma_H\Sigma^H_{im}$$
\[ a'_n = (a'_{n-1} - e'_\pi)A + A'_m + \frac{1}{2}[(\Sigma'_F(a_{n-1} - e_\pi)) \otimes (\Sigma'_F(a_{n-1} - e_\pi))]' \Omega + \]
\[ \frac{1}{2} [\Sigma_{Fm} \otimes \Sigma_{Fm}]' \Omega + (a'_{n-1} - e'_\pi)(\Sigma'_{Fm} \otimes \Sigma_F) \Omega \quad (20) \]

**Proof of Proposition 2**

Real equity prices must satisfy,
\[ P^e_t = E_t[exp(m_{t+1})(P^e_{t+1} + D_{t+1})] \quad (21) \]

Suppose prices take the form,
\[ p^e_t = b^n_c + b'_n Y_t + Z_t \]
\[ p^e_{t+1} = b^n_{n-1} + b'_{n-1} Y_{t+1} + Z_{t+1} \quad (22) \]

where \( Z_t \) is a random walk component assumed to follow,
\[ \Delta Z_{t+1} = Z_{t+1} - Z_t = \mathbf{B}' Y_t \]

where \( \mathbf{B}' \) is a \( 1 \times k \) vector loading onto the state vector, with \( k \) being 4 in my model. Now plug (22) into the pricing equation (21). This yields
\[ P^e_t = E_t[exp(m_{t+1})P^e_{t+1} exp(dp_{t+1})] \]
\[ = E_t[exp(m_{t+1} + b'^n_{n-1} + Z_t + (e'_{dp} + b'_{n-1}) Y_{t+1} + \mathbf{B}' Y_t)] \]
\[ = exp(b^n_c + b'_n Y_t + Z_t) \quad (23) \]

where \( e'_{dp} = [0 \ 0 \ 1 \ 0] \). Notice that the dividend yield \( dp_t \) is not a simple ratio \( D_t/P_t \), but is defined to be \( dp_t = ln\left(1 + \frac{D_t}{P_t}\right) \) and the third equal sign uses the Lemma 1 again. Therefore this determines the recursive coefficients,
Unlike bonds that discount payoff in finite periods, a stock has an infinite horizon of dividends inflow. Therefore \( b_n \) and \( b_{n-1} \) converge in the long run\(^1\) for \( b_n = b_{n-1} \), so I drop the subscript to use \( b \). On the other hand \( b_n \) might have a non-zero increment over time, which is \( b_n - b_{n-1} \), implying a stock price that is trending upward. Equations (24) and (25) are rewritten as,

\[
\begin{align*}
  b_n &= b_{n-1} + (e'_d + b')\mu + \mu_m + \frac{1}{2} \left( \left( \Sigma'_F(e_d + b) \right) \otimes \left( \Sigma'_F(e_d + b) \right) \right)' \omega + \\
        &\quad + \frac{1}{2} \left( e'_d + b' \right) \Sigma_H \Sigma'_H (e_d + b) + \frac{1}{2} \left[ \Sigma_{Fm} \otimes \Sigma_{Fm} \right]' \omega + \frac{1}{2} \Sigma'_{Hm} \Sigma_{Hm} + \\
        &\quad + (e'_d + b') (\Sigma'_{Fm} \otimes \Sigma_F) \omega + (e'_d + b') \Sigma_H \Sigma_{Hm} \quad (26)
\end{align*}
\]

\[
\begin{align*}
  b' &= B' + (e'_d + b')A + A'_m + \frac{1}{2} \left( \left( \Sigma'_F(e_d + b) \right) \otimes \left( \Sigma'_F(e_d + b) \right) \right)' \Omega + \\
     &\quad + \frac{1}{2} \left[ \Sigma_{Fm} \otimes \Sigma_{Fm} \right]' \Omega + (e'_d + b') (\Sigma'_{Fm} \otimes \Sigma_F) \Omega \quad (27)
\end{align*}
\]

Also unlike in bond pricing, where the starting point is \( a_0^c = 0, a'_0 = [0 \ 0 \ 0 \ 0] \) and \( P_{b,0} = 1 \), equity pricing equations do not have an initial level for \( b_n \) or \( Z_t \). Nor is that required, because it is not the absolute level of stock price that matters, but rather the rate at which prices change over the horizon. Same is true with bond prices which are used mainly to compute all types of rates and yield to maturity.

\(^1\) The same technique is used in Li (2002) and d’Addona and Kind (2006). I differ however to their method by inclusion of a random walk component in stock prices.
II Analytical Moments

Unconditional moments of $Y_t$ are available analytically as functions of the parameters to be estimated.

\[ E[Y_t] = (I - A)^{-1} \mu \] (28.1)

\[ V[Y_t] = A \cdot V[Y_t] \cdot A' + (\Sigma_F F_t + \Sigma_H)(\Sigma_F F_t + \Sigma_H)' \] (28.2)

where $V[Y_t]$, the variance of $Y_t$, is implicitly solved.

\[ C(Y_t, Y_{t-1}) = A \cdot V[Y_t] \] (28.3)

As pointed out in Chapter 2, these three moments of the state vector $Y_t$ are the foundations of all analytical representations that follow:

\[ E[\Delta c_t] = e'_c E[Y_t], \ E[dp_t] = e'_e E[Y_t], \ E[\pi_t] = e'_\pi E[Y_t] \] (29.1)

\[ E[r_t^f] = -a'_c E[Y_t] \] (29.2)

\[ E[s_t^{pd}] = -a_{10}'/10 + a'_c + (a'_1 - a_{10}'/10)E[Y_t] \] (29.3)

\[ E[\Delta p_t] = b'_c - b'_{n-1} + B'E[Y_t] \] (29.4)

\[ E[r_t^{ex}] = b'_c - b'_{n-1} + a'_c + (B' + e'_d + e'_\pi + a'_1)E[Y_t] \] (29.5)

$V(\Delta c_t), V(dp_t), V(\pi_t)$ are the diagonal elements of matrix $V[Y_t]$ (29.6)

\[ V[r_t^f] = (-a'_1)V[Y_t](-a_1) \] (29.7)

\[ V[s_t^{pd}] = (a'_1 - a_{10}'/10)V[Y_t](a_1 - a_{10}/10) \] (29.8)

\[ V[\Delta p_t] = b'V[Y_t]b + (B' - b')V[Y_t](B - b) + 2b'C(Y_t, Y_{t-1})(B - b) \] (29.9)

\[ V[r_t^{ex}] = (b' + e'_d + e'_\pi)V[Y_t](b + e_d + e_\pi) + (B' - b' + a'_1)V[Y_t](B - b + a_1) + 2(b' + e'_d + e'_\pi)C(Y_t, Y_{t-1})(B - b + a_1) \] (29.10)
\( C(\Delta c_t, \Delta c_{t-1}), C(dp_t, dp_{t-1}), C(\pi_t, \pi_{t-1}) \) are the diagonal elements of matrix \( C(Y_t, Y_{t-1}) \)

\[
C(r_t^f, r_{t-1}^f) = (-a_1^i)C(Y_t, Y_{t-1})(-a_1) \tag{29.12}
\]

\[
C[s_t^{pd}, s_{t-1}^{pd}] = (a_1^i - a_{10}^i/10)C(Y_t, Y_{t-1})(a_1 - a_{10}/10) \tag{29.13}
\]

\[
C[r_t^{ex}, dp_t] = e_{dp}'V[Y_t](b + e_{dp} + e_{\pi}) + e_{dp}' \cdot C(Y_t, Y_{t-1})(B - b + a_1) \tag{29.14}
\]

\[
C[r_t^f, \Delta c_t] = e_c' \cdot A \cdot V[Y_t](-a_1) \tag{29.15}
\]

In this way, all the unconditional moments of the state and derived variables can therefore be written in terms of my 21 parameters. These expressions are then equated to the real data moments counterparts to form the 23 moment conditions used in my GMM objective function.
<table>
<thead>
<tr>
<th>parameters</th>
<th>estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_c$</td>
<td>0.0379</td>
</tr>
<tr>
<td>(0.0175)</td>
<td></td>
</tr>
<tr>
<td>$\mu_d$</td>
<td>0.0191</td>
</tr>
<tr>
<td>(0.0058)</td>
<td></td>
</tr>
<tr>
<td>$\mu_\pi$</td>
<td>0.0136</td>
</tr>
<tr>
<td>(0.0009)</td>
<td></td>
</tr>
<tr>
<td>$\rho_q$</td>
<td>0.9129</td>
</tr>
<tr>
<td>(0.0203)</td>
<td></td>
</tr>
<tr>
<td>$\rho_{cc}$</td>
<td>0.3539</td>
</tr>
<tr>
<td>(0.0598)</td>
<td></td>
</tr>
<tr>
<td>$\rho_{cd}$</td>
<td>-0.6587</td>
</tr>
<tr>
<td>(0.4342)</td>
<td></td>
</tr>
<tr>
<td>$\rho_{dd}$</td>
<td>0.5000</td>
</tr>
<tr>
<td>(0.1455)</td>
<td></td>
</tr>
<tr>
<td>$\rho_{dc}$</td>
<td>0.0341</td>
</tr>
<tr>
<td>(0.0274)</td>
<td></td>
</tr>
<tr>
<td>$\rho_\pi$</td>
<td>0.4501</td>
</tr>
<tr>
<td>(0.0346)</td>
<td></td>
</tr>
<tr>
<td>$\sigma_q$</td>
<td>0.4257</td>
</tr>
<tr>
<td>(0.1982)</td>
<td></td>
</tr>
<tr>
<td>$\sigma_{cc}$</td>
<td>0.0204</td>
</tr>
<tr>
<td>(0.0013)</td>
<td></td>
</tr>
<tr>
<td>$\sigma_{dc}$</td>
<td>-0.0012</td>
</tr>
<tr>
<td>(0.0008)</td>
<td></td>
</tr>
<tr>
<td>$\sigma_{dd}$</td>
<td>0.0088</td>
</tr>
<tr>
<td>(0.0017)</td>
<td></td>
</tr>
<tr>
<td>$\sigma_\pi$</td>
<td>0.0288</td>
</tr>
<tr>
<td>(0.0021)</td>
<td></td>
</tr>
<tr>
<td>$\lambda$</td>
<td>-0.7406</td>
</tr>
<tr>
<td>(0.1288)</td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.9778</td>
</tr>
<tr>
<td>(0.0024)</td>
<td></td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.2994</td>
</tr>
<tr>
<td>(0.1509)</td>
<td></td>
</tr>
<tr>
<td>$\rho_{zq}$</td>
<td>0.0392</td>
</tr>
<tr>
<td>(0.0255)</td>
<td></td>
</tr>
<tr>
<td>$\rho_{zc}$</td>
<td>5.2536</td>
</tr>
<tr>
<td>(0.8114)</td>
<td></td>
</tr>
<tr>
<td>$\rho_{zd}$</td>
<td>-3.0693</td>
</tr>
<tr>
<td>(0.7647)</td>
<td></td>
</tr>
<tr>
<td>$\rho_{z\pi}$</td>
<td>-0.8662</td>
</tr>
<tr>
<td>(0.1655)</td>
<td></td>
</tr>
</tbody>
</table>

$J$-stat(2) = 8.0592

$P$-val = 0.0178

GMM Standard error in parentheses. Data are annual from 1927-2000. Test of over-identifying restrictions fails to reject at 1% level.
Table 2
Moment Summary

<table>
<thead>
<tr>
<th></th>
<th>$\Delta c_t$</th>
<th>$dp_t$</th>
<th>$\pi_t$</th>
<th>$r^f_t$</th>
<th>$s^p_t$</th>
<th>$\Delta p_t$</th>
<th>$r^e_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0184</td>
<td>0.0394</td>
<td>0.0247</td>
<td>0.0435</td>
<td>0.0023</td>
<td>0.0333</td>
<td>0.0603</td>
</tr>
<tr>
<td>Std dev</td>
<td>0.0237</td>
<td>0.0102</td>
<td>0.0323</td>
<td>0.0227</td>
<td>0.0124</td>
<td>0.1883</td>
<td>0.1783</td>
</tr>
<tr>
<td>Auto Corr.</td>
<td>0.4258</td>
<td>0.4797</td>
<td>0.4501</td>
<td>0.7350</td>
<td>0.4770</td>
<td>0.0236</td>
<td>-0.0074</td>
</tr>
</tbody>
</table>

## Table 3
Cross Correlations

<table>
<thead>
<tr>
<th></th>
<th>$\Delta c_t$</th>
<th>$dp_t$</th>
<th>$\pi_t$</th>
<th>$r_t^f$</th>
<th>$s_t^{pd}$</th>
<th>$\Delta p_t$</th>
<th>$r_t^{ex}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$dp_t$</td>
<td>-0.2550</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.3376</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1238)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\pi_t$</td>
<td>0.0000</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4215</td>
<td>0.1020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1719)</td>
<td>(0.1296)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_t^f$</td>
<td>-0.0334</td>
<td>-0.0668</td>
<td>0.2881</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0777</td>
<td>-0.2393</td>
<td>0.3355</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1113)</td>
<td>(0.1176)</td>
<td>(0.1376)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$s_t^{pd}$</td>
<td>-0.0822</td>
<td>0.0695</td>
<td>-0.4311</td>
<td>-0.7832</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1330</td>
<td>0.0331</td>
<td>-0.1676</td>
<td>-0.5969</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0837)</td>
<td>(0.0960)</td>
<td>(0.1394)</td>
<td>(0.1089)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta p_t$</td>
<td>0.1184</td>
<td>-0.6107</td>
<td>-0.2153</td>
<td>0.0003</td>
<td>-0.0064</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1092</td>
<td>-0.4870</td>
<td>-0.1600</td>
<td>-0.0825</td>
<td>0.2459</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1521)</td>
<td>(0.1206)</td>
<td>(0.2288)</td>
<td>(0.1167)</td>
<td>(0.1272)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_t^{ex}$</td>
<td>0.1148</td>
<td>-0.5796</td>
<td>-0.0830</td>
<td>-0.0786</td>
<td>0.0189</td>
<td>0.9826</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.1874</td>
<td>-0.3741</td>
<td>-0.0254</td>
<td>-0.2199</td>
<td>0.3163</td>
<td>0.9610</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1592)</td>
<td>(0.1682)</td>
<td>(0.2165)</td>
<td>(0.1043)</td>
<td>(0.1130)</td>
<td>(0.0097)</td>
<td></td>
</tr>
</tbody>
</table>

Table 4
Variance Decomposition

<table>
<thead>
<tr>
<th></th>
<th>$\Delta c_t$</th>
<th>$d_p_t$</th>
<th>$\pi_t$</th>
<th>$q_t$</th>
<th>$\Delta c_{t-1}$</th>
<th>$d_p_{t-1}$</th>
<th>$\pi_{t-1}$</th>
<th>$q_{t-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_t^f$</td>
<td>-0.0197</td>
<td>0.0093</td>
<td>0.4099</td>
<td>0.6006</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$s_t^{pd}$</td>
<td>0.0190</td>
<td>0.0179</td>
<td>0.9176</td>
<td>0.0455</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\Delta p_t$</td>
<td>0.1082</td>
<td>0.4711</td>
<td>0.0581</td>
<td>0.3761</td>
<td>-0.0050</td>
<td>-0.0242</td>
<td>0.0000</td>
<td>0.0157</td>
</tr>
<tr>
<td>$r_t^{ex}$</td>
<td>0.1108</td>
<td>0.4394</td>
<td>0.0086</td>
<td>0.1882</td>
<td>-0.0115</td>
<td>-0.0010</td>
<td>0.0000</td>
<td>0.2654</td>
</tr>
</tbody>
</table>

By model specification, each derived variable on the far left column is a linear combination of current and lag state variables. The table exhibits the contribution percentage of every state factor in the total variation of the derived variables.

Table 5
Risk Aversion Property by Bekaert et al. (2010)

<table>
<thead>
<tr>
<th>$q_t$ correlation</th>
<th>$\Delta c_t$</th>
<th>$d_p_t$</th>
<th>$\pi_t$</th>
<th>$r_f$</th>
<th>$sp_{d_t}$</th>
<th>$re_{x_t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.15</td>
<td>0.93</td>
<td>0.000</td>
<td>0.21</td>
<td>0.20</td>
<td>-0.21</td>
<td></td>
</tr>
</tbody>
</table>

Source: Bekaert et al. (2010) Table 5
Table 6
Risk Aversion Property

<table>
<thead>
<tr>
<th>$RA_t = rQ_t$ percentile</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.26</td>
<td>0.30</td>
<td>0.32</td>
<td>0.39</td>
<td>0.63</td>
<td>1.3</td>
<td>3.8</td>
<td>9</td>
<td>39</td>
</tr>
</tbody>
</table>

$q_t$ correlation $\Delta c_t$ $dp_t$ $\pi_t$ $rf_t$ $spd_t$ $\Delta p_t$ $rex_t$

-0.391 0.0140 0.688 -0.165 0.149 0.0705

At the point estimates from Table 1, first row of Table 5 shows the percentile distribution of the simulated risk aversion time series and second row shows model predicted correlations of the unobserved surplus ratio with all other variables.

Table 7
Risk Aversion Correlations by Different Models

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho(RRA, rf)$ almost 1 or -1 if $b \neq 0$</td>
<td>close to 1</td>
<td>close to -1</td>
<td>0.21</td>
<td>close to 1</td>
<td>0.688</td>
<td></td>
</tr>
<tr>
<td>$\rho(RRA, dp)$ close to 1</td>
<td>close to 1</td>
<td>close to 1</td>
<td>0.93</td>
<td>N/a</td>
<td>0.014</td>
<td></td>
</tr>
</tbody>
</table>

Six models including ours are displayed for comparison in their predictions on the connections between risk aversion and other few key time-series processes.
Table 8
Long Horizon Return Regression

<table>
<thead>
<tr>
<th>Horizon k in year</th>
<th>Data ( \text{rex}<em>{t \rightarrow t+k} = a + b \ast d</em>{pt} )</th>
<th>Model ( \Delta p_{t \rightarrow t+k} = a + b \ast d_{pt} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( b )</td>
<td>( \sigma(b) )</td>
</tr>
<tr>
<td>1</td>
<td>6.3</td>
<td>1.9</td>
</tr>
<tr>
<td>2</td>
<td>11.2</td>
<td>2.4</td>
</tr>
<tr>
<td>3</td>
<td>13.9</td>
<td>2.6</td>
</tr>
<tr>
<td>4</td>
<td>17.5</td>
<td>2.9</td>
</tr>
<tr>
<td>5</td>
<td>22.4</td>
<td>3.4</td>
</tr>
</tbody>
</table>

OLS regression of future \( k \) year excess return or price growth rate on time \( t \) dividend yield. Left column uses sample period 1945-1996 of actual data. Right column uses model predicted data 1957-1996. Model predicted data are generated with point estimates from Table 1.
Table 9
Parameters Estimation of Variant (32)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Feedback</th>
<th>Volatility</th>
<th>Preference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_c$</td>
<td>0.0365</td>
<td>0.9086</td>
<td>0.4175</td>
<td>$\beta$ 0.9807</td>
</tr>
<tr>
<td></td>
<td>(0.0171)</td>
<td>(0.0210)</td>
<td>(0.2041)</td>
<td>(0.0023)</td>
</tr>
<tr>
<td>$\mu_d$</td>
<td>0.0193</td>
<td>0.3540</td>
<td>0.0201</td>
<td>$\gamma$ 0.2874</td>
</tr>
<tr>
<td></td>
<td>(0.0061)</td>
<td>(0.0716)</td>
<td>(0.0021)</td>
<td>(0.1429)</td>
</tr>
<tr>
<td>$\mu_\pi$</td>
<td>0.0129</td>
<td>-0.6558</td>
<td>-0.0014</td>
<td>$\rho_{zq}$ 0.0380</td>
</tr>
<tr>
<td></td>
<td>(0.0010)</td>
<td>(0.4403)</td>
<td>(0.0009)</td>
<td>(0.0242)</td>
</tr>
<tr>
<td>$\rho_{dd}$</td>
<td>0.5005</td>
<td>0.0083</td>
<td></td>
<td>$\rho_{zc}$ 5.6112</td>
</tr>
<tr>
<td></td>
<td>(0.1395)</td>
<td>(0.0020)</td>
<td></td>
<td>(0.8905)</td>
</tr>
<tr>
<td>$\rho_{dc}$</td>
<td>0.0352</td>
<td>$\sigma_{\pi\pi}$ 0.0271</td>
<td></td>
<td>$\rho_{zd}$ -3.1214</td>
</tr>
<tr>
<td></td>
<td>(0.0297)</td>
<td>(0.0093)</td>
<td></td>
<td>(0.7881)</td>
</tr>
<tr>
<td>$\rho_{\pi}$</td>
<td>0.4509</td>
<td>$\sigma_{\pi} $ 0.0106</td>
<td></td>
<td>$\rho_{z\pi}$ -0.8929</td>
</tr>
<tr>
<td></td>
<td>(0.0284)</td>
<td>(0.0037)</td>
<td></td>
<td>(0.1598)</td>
</tr>
<tr>
<td>$\lambda$</td>
<td></td>
<td></td>
<td>-0.7590</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.1323)</td>
<td></td>
</tr>
</tbody>
</table>

$J$-stat(2) 8.8038
$PVal$ 0.0123

Parameter estimation of variant model (32). GMM Standard error in parentheses. Data are annual from 1927-2000. Test of over-identifying restrictions fails to reject at 1% level.

\[
q_{t+1} = \mu_q + \rho_q q_t + \sigma_q \sqrt{q_t} (\sqrt{1 - \lambda^2} \epsilon_{t+1}^q + \lambda \epsilon_t^q) \\
\Delta c_{t+1} = \mu_c + \rho_{cc} \Delta c_t + \rho_{cd} \Delta d_p_t + \sigma_{cc} \sqrt{q_t} \epsilon_{t+1}^c \\
d_p_{t+1} = \mu_d + \rho_{dd} d_p_t + \rho_{dc} \Delta c_t + \sigma_{dc} \sqrt{q_t} \epsilon_{t+1}^c + \sigma_{dd} \epsilon_{t+1}^d \\
\pi_{t+1} = \mu_\pi + \rho_{\pi} \pi_t + \sigma_{\pi\pi} \sqrt{q_t} \epsilon_{t+1}^\pi + \sigma_{\pi} \sqrt{q_t} \epsilon_{t+1}^c \\
(32)
\]
The table summarizes standard deviations and correlations of state variables based on estimated variant model (32). Standard deviations of state variables are on the diagonal. Off-diagonal entries are correlation coefficients. All moments are unconditional.

<table>
<thead>
<tr>
<th></th>
<th>$q_t$</th>
<th>$\Delta c_t$</th>
<th>$dp_t$</th>
<th>$\pi_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q_t$</td>
<td>1.00</td>
<td>-0.414</td>
<td>0.026</td>
<td>0.208</td>
</tr>
<tr>
<td>$\Delta c_t$</td>
<td>-0.414</td>
<td>2.32%</td>
<td>-0.274</td>
<td>0.341</td>
</tr>
<tr>
<td>$dp_t$</td>
<td>0.026</td>
<td>-0.274</td>
<td>0.96%</td>
<td>-0.044</td>
</tr>
<tr>
<td>$\pi_t$</td>
<td>0.208</td>
<td>0.341</td>
<td>-0.044</td>
<td>3.26%</td>
</tr>
</tbody>
</table>
This table presents cross correlations among variables of the economy that features variant model (32). The first row is model predicted moments at estimates given by Table 9. The second row is sample moments of actual dataset. Standard error in parentheses.
Figure 1: Price-Dividend Ratio as a Function of Surplus Consumption Ratio. Source: CC1999

Figure 2: Latent State Variable, Surplus Ratio $q$ Extracted by Kalman Filter
Figure 3: Distribution of Surplus Ratio $q$

Figure 4: Risk Aversion $RA = \gamma \exp(q)$
Figure 5: Predicted (Filtered) Risk-Free Rate (Dash Line) vs. Actual (Solid Line)

Figure 6: Predicted (Filtered) Yield Spread (Dash Line) vs. Actual (Solid Line)
Figure 7: Predicted Return/Price Growth (Dash Line) vs. Actual (Solid Line)

Figure 8: 1970-2013 Surplus Ratio $q$ Predicted Out Of Sample
Figure 9: 1970-2013 Campbell-Cochrane Model Predicted $P/D$ (dot line) vs. Actual (solid line)

Figure 10: 1970-2013 My Model Predicted $P/D$ (Dot Line) vs. Actual (Solid Line)
Bibliography


Li, L., 2002, Macroeconomic Factors and the Correlation of Stock and Bond Returns. Yale working paper.


