THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITE DE TOULOUSE

Délivré par l'Université Toulouse Capitole

École doctorale : **Sciences Economiques-Toulouse School of Economics**

Présentée et soutenue par

TINANG NZESSEU Jules Valéry

le 25 juillet 2018

Essays in Financial Economics

Discipline : **Sciences Economiques**

Unité de recherche : **TSE-R (UMR CNRS 5314 – INRA 1415)**

Directeur de thèse : M. Christian GOLLIER, Professeur, Université Toulouse 1 Capitole

Co-directeur de thèse: M. Nour MEDDAHI, Professeur, Université Toulouse 1 Capitole

JURY

Rapporteurs M. George TAUCHEN, Professeur, Duke University
M. Ivan SHALIASTOVICH, Professeur, University of Wisconsin-Madison

Suffragants M. Roméo TEDONGAP, Professeur, ESSEC Business School
Mme. Marianne ANDRIES, Professeur, Université Toulouse 1 Capitole
M. Patrick FEVE, Professeur, Université Toulouse 1 Capitole

Essays in Financial Economics

Résumé

Cette thèse est composée de quatre chapitres autonomes visant à contribuer à une meilleure compréhension de la formation et de la dynamique des prix des actifs dans un modèle d'évaluation des actifs financiers fondé sur la consommation (C-MEDAF). Le chapitre 1 examine la structure par terme du rendement des actions dans les principaux modèles C-MEDAF et montre que permettre aux flux de consommation et de dividendes d'être affectés négativement par les chocs de volatilité comme observés empiriquement ("effet de levier") pourrait rendre les actifs de court terme plus risqués que les actifs de long terme comme on l'a récemment découvert dans certaines études empiriques. Cette modification donne plus de souplesse à ces modèles pour saisir différentes formes de la structure par termes des taux de rendement d'actifs risqués tout en respectant les niveaux observés de la prime de risque et du taux de rendement sans risque. Le chapitre 2 propose un modèle à changement de régimes pour s'adapter au comportement changeant de la pente de la structure par termes des rendements d'actifs risqués tel qu'observé dans les données. Nous montrons qu'un tel modèle permet de combiner les propriétés spécifiques à chaque régime des modèles à un régime telles que une pente en moyenne positive ou négative de la structure par termes des rendements, et donne plus de flexibilité dans la forme de la structure à terme des rendements d'actifs risqués. Le chapitre 3 étudie l'hypothèse d'anticipation sur le marché d'actions. Selon cette hypothèse, les rendements courant sur les actifs de long terme sont une moyenne pondérée de l'espérance des rendements futurs de court terme. Ce test a principalement été réalisé sur le marché des bons du trésor et dans beaucoup de cas rejeté. Cette hypothèse n'est pas rejetée sur le marché d'actions mais les rendements futurs sont aussi prévisibles. Le chapitre 4 examine l'estimation et l'inférence dans le modèle de risques de long terme en utilisant la méthode généralisée des moments.

Abstract

This thesis is made of four self-contained papers aiming at contributing to the better understanding of asset prices formation and dynamics in a Consumption-based Capital Asset Pricing Model (CCAPM). Chapter 1 looks at the term structure of equity return in leading CCAPM models and show that allowing the cash flows to be negatively affected by volatility shocks, as observed in the data ("leverage effect"), could make the short-term assets riskier than long-term assets as recently found in some empirical papers. This modification gives more flexibility to those models in capturing various shapes of the term structure of equity returns while still matching the observed level of the equity premium and the risk free rate. Chapter 2 proposes a regimes switching model to accommodate for the changing behavior of the term structure of equity returns as observed in the data. We show that such a model allows to combine the properties of the one regime models and it gives more flexibility in the shape of the average term structure of equity returns. Chapter 3 studies the Expectation Hypothesis on equity markets. This test has mainly been done for the bonds market. We find that the EH is not rejected but the future returns are also predictable. Chapter 4 examines the estimation and the inference in the LRR model using the Generalized Method of Moments.

Acknowledgments

This work would not have been possible without the help and the support of many people to whom I express my sincere gratitude.

First and foremost, I am heavily indebted to my advisors Prof. Christian Gollier and Prof. Nour Meddahi for their invaluable support, their advices and their guidance. I could not think about better advisors to introduce me into a research area that combines both finance theory and econometrics. I greatly thank them for their availability, their patience, and for their pedagogical approach in introducing me in research.

I greatly thank Prof. George Tauchen and Prof. Ivan Shaliastovich for accepting to to read and evaluate my work. Their comments and suggestions are invaluable for the improvement of this work.

I would also like to express my profound gratitude to Prof. Marianne Andries, Prof. Patrick Fève and Prof. Roméo Tédongap for being part of the jury. I have benefited from their ideas, comments and immense knowledge along these five years. I thank them for their availability to discuss about my work.

I have also benefited from the stimulant research environment at TSE and from the discussions with my classmates. I specially thank Simon Fuchs,Vitalijs Jascisen, Rohit Kumar, Ying Liang, Lucas Mahieux, Christian Nguenang, Serge Nyawa and Mamiko Yamashita. Over these years, I have also received insightful ideas, valuable comments and encouragements from Bruno Biais, Christian Bontemps, Michel-Benoît Bouissou, Catherine Casamatta, Jean Pierre Florens, René Garcia, Jihyun Kim, Pascal Lavergne, Thierry Magnac, Thomas Mariotti, Sophie Moinas, Francois Poignas, Sebastien Pouget. Special thanks to Caroline Tejedor and Aude Schloesing for their availability in helping me to solve various administrative problems.

The last but not the least, I would like to expressed my gratitude to all my great family members, especially to my father Nzesseu Benjamin, my mother Wandji Lisette, my siblings : Amélie, Lucile, Prisca, Stéphanie, Willy and Ephraïm. I thank them for their patience and for always pushing me to fulfill my dreams. This work is dedicated to my wife Liliane, and my children Joanna and Ethan. Every day with them is an endless source of happiness and motivation.

Contents

List of Figures

List of Tables

Introduction

The allocation of scarce economic resources to improve the human well being is at the heart of economic policy. In a competitive capital market, a risky project should be undertaken only if its present value is greater than the cost of the investment needed for its implementation. Evaluating the two parts of this decision rule can be a daunting task since on one hand the cost (but also the benefits) of a project involves the direct cash flow related to its realization but also the negative (or positive) externalities generated by the project. On the other hand, benefits and costs associated with the risky project are not only random but also usually spread over time such that the computation of their present value needs to take into account the risk associate with the project and the human preference for the present[1]. On top of that, a manager could have to choose between many risky and incompatible projects, and face the difficult question of which project to realize. Thus appropriately discounting the risky future net benefits expected from a project's realization, and thus determining its price or value, is of prime importance for resources allocation. The Capital Asset Pricing Model (CAPM) proposed by Sharpe [1964] and Lintner [1965] provides a theoretical and easily applicable answer to this problem. According to the CAPM, the expected return on a risky project in excess of the risk free rate is equal to the quantity of systematic risk embedded in the project or the beta of the project times the market price of risk. A risky asset that highly positively co-moves with the market portfolio will require a high return to be hold since adding that asset to the market portfolio will increase the overall risk of the new portfolio. However, the CAPM presents some limitations related to the fact that only financial market variables are involved in the computation of the expected return of a risky project. First the market portfolio could be unknown or nonexistent as it is the case for many developing countries, then it becomes difficult to assess the systematic risk embedded in a project. The second challenge posed by the CAPM is about the connection of the risk on financial markets with the overall economy : what drives the price of risk, the market return

[1] In general, 1 dollar today has more value than 1 dollar in 1 year for everyone.

itself or the risk free rate ? These questions are very important for pension funds, government or any wealth manager with long term goals.

The Consumption-based Capital Asset Pricing Model (CCAPM) proposed by Lucas [1978] provides a theoretical framework to connect asset prices with economic fundamentals such as consumer preferences, aggregate consumption shocks, macroeconomic uncertainty, etc. In its basic formulation, the CCAPM considered a representative agent with time separable utility function and a log-normally distributed consumption process[2]. This model states that the expected excess return on a risky asset equals the covariance between consumption growth and the asset returns, multiplied by the consumer risk aversion coefficient. This simple statement allows to understand why people do not put all their money on financial markets despite the relative high average returns observed in the past and also incoherently with the fact that in their daily life they are ready to take other kind of risks such as playing lotteries. The explanation provided by the CCAPM is that the returns on financial market reward a special kind of risk which is positively correlated with consumption; in other words the expected return for holding a stock is high because its value might fall at the same time you loose your job and your marginal utility is high (Cochrane [2017]). But this basic model failed in explaining many of the observed quantitative facts on financial markets. Along the 80's, many puzzles were raised against the canonical model such as the volatility puzzle (Shiller [1981]), the equity premium puzzle (Mehra and Prescott [1985]) and the risk free rate puzzle (Weil [1989b]). Indeed, the standard model could not rationalized the observed prices behavior given the economic fundamentals observed (e.g. the relative smooth consumption growth) or commonly accepted (e.g. a relative risk aversion coefficient below 10). Solving the puzzles raised by the standard CCAPM has been a major challenge in macro-finance and many models have been proposed. The more prominent models with a representative agent in a complete and competitive market are the Habit Formation Model (Cochrane and Campbell [1999]), the Long Run Risks model (Bansal and Yaron [2004]), the Rare Disaster Model (Barro [2006]). These models were successful in providing a credible explanation for why we could simultaneously observed an average high equity premium on the market, a high volatility of returns and, a low and smooth risk free rate. The Habit formation Model explains that the utility derived from consumption at a given time depends on the habit level which could be seen as the recent past average level of the consumption. When the consumption realization falls below the habit level for example during recessions, the consumer

[2]Thereafter, I will interchangeably use the words basic or standard or canonical to refer to this model.

becomes more risk averse as her marginal utility increases; thus she will require a high premium to hold risky assets. So, the relative smoothness of the consumption growth translates into more volatile risky asset prices because of the changes in the concavity of the utility function. The Long Run Risks model, henceforth LRR, attributes the observed high equity premium to the presence of a smooth and persistence predictable component in the consumption growth driving the long run risk. Because of that component and the recursivity of the utility function, negative consumption shocks will have long lasting effects and the consumer dislike those risks. The rare disaster model argues that the relative smoothness of the observed U.S consumption is a matter of luck given the disasters that were happening elsewhere in the world. The U.S consumers might be aware of that and in their decision making, they may have internalized the probability that such disasters could also occur in the U.S economy. This "subjective" disaster probability will increase the volatility of the consumer perceived consumption growth relative to what is observed and it rationalizes the observed high equity premium. The success of all those models have mainly been illustrated through some calibrations to match some key observed moments in the data. Recently, researchers have been moving beyond the calibrations; they have started to design methods in order estimate and formally test those models. The task is not easy given short length of the observed macroeconomic data and the complexity of the models involving non-linearity or the presence of latent unobserved variables. But, the estimation and the testing of those models give a bigger perspective in looking at how they work, what are the key ingredients and also their limitations.

Up to the beginning of the 21^{st} century, the models where designed to address the equity premium puzzle and researchers were agnostic about the maturity decomposition of the equity risk premium. However, a stock can be seen as a portfolio of zero coupon equities or dividend strips with maturities ranging from 1 year to infinity; each zero coupon equity paying the realized dividends at maturity and nothing else. An important question is then how do the elements in the portfolio (the dividend strips) contribute to the risk and the observed returns of the overall portfolio ? Recently this question has been investigated and some new puzzles have been raised against the leading frictionless consumption-based asset pricing models concerning the term structure of equity returns van Binsbergen et al. [2012], Maggiori et al. [2015]). Those models predict that the expected return on assets whose cash flows appear in the distant future are higher than or equal to the expected returns on assets which pay-off in the near future. Contrary to that prediction, some recent empirical studies (van Binsbergen et al. [2012], van

Binsbergen and Koijen [2016], van Binsbergen et al. [2013]) have found that short-term assets on average earn a higher expected return than long-term assets. In the Chapter 1, we design a consumption based asset pricing model that yields an average downward sloping term structure of equity returns as recently observed in the data. Our models builds on the leading asset pricing models framework by assuming a representative agent with the same kind of preferences (CRRA or recursive utility) as they do but the cash flows dynamics is modified in order to capture an important economic mechanism driving the result. More specifically, we show that allowing the cash flows to be negatively affected by volatility shocks, as observed in the data ("leverage effect"), could make the short-term assets on average riskier than long-term assets. Bloom et al. [2007], Bloom [2009] explained that an increase in uncertainty reduces investment response to demand shocks because it pushes investors to delay their investment decision; this will reduce the output and thus have a negative impact on the cash flows (consumption and dividends). Empirically, we observe on figure 1.1 that during recession periods in particular uncertainty increases and cash flows drop. Applying a reduced form VAR model with cash flows growth and a measure of uncertainty (see Table 1.5), we find that cash flows growth shocks are negatively correlated with uncertainty shocks. The model proposed in this paper gives more flexibility in capturing various shapes of the term structure of equity returns while still matching the observed level of the equity premium and the risk free rate.

However, the term structure slope is time varying as we can observe from figure 1.7. Furthermore, as we can see on figure 2.2, the dividend yields spread on the S&P 500 dividend strips happened to be on average positive during normal times and negative during recessions Bansal et al. [2017]. Thus, only matching the average slope of the equity term structure might not be full informative about it dynamics. The Chapter 2 proposes a state space extension of the previous model that allows to switch from a parametrization that gives an upward sloping term structure to a one that yields a downward sloping term structure. As emphasized in the previous chapter, the model presented there is able to capture both an upward or a downward sloping term structure of the risk premium depending on the calibration that is used. So, the regime switching model with two states will allow to move from states with a upward sloping term structure (normal times) to states with a downward sloping term structure (bad times). The regime switching model is able to capture the key asset pricing implications achieved by both state specific calibrations such as the high equity premium and high return volatility, the low risk free rate and low consumption volatility. Setting a high probability to remain within a given regime enables to stay within that regime for a while and to have a slope of the term structure that reflects the dominant regime. On average the term structure will be positive as observed in the data because the transition matrix enables to favor the regime with an upward sloping term structure by specifying a bigger probability to be in that regime compared to the alternative.

The expectation hypothesis is a consequence of the efficient market hypothesis (EMH). If markets are efficient then returns on stocks or bonds should not be predictable and thus all the variation in the long-term spot yield spread should come from the expected variation in the future short-term yield spread. So, when the EH holds, there is a one to one relationship between the current long term yield spread and the expected future short-term yield spread; meaning if the current yield curve is upward sloping it is simply because the market expects an increase in the short-term yields. The expectation hypothesis has mainly be tested (and rejected) on the bond market by regressing the future changes of the short-term yields on the current long term yield spread. In the Chapter 3, we show that this test can be extended to the equity market by applying the same type of regressions. We formulate an equivalent of the Campbell and Shiller. [1991] regressions that can be applied on equity dividend yields to test the EH. We find that the EH is not rejected when using those equivalent regressions of forward yield spreads. But using the Cochrane and Piazzesi [2005] type regressions, we find that excess return on dividend strips with various maturities are predictable in a one factor model, where the factor is obtained as a combination of forward rates up to 5 years of maturity. This unexpected result might sound contradictory but it is not. Indeed, the EH states that if returns are unpredictable then the current yield spreads should predict future yield spread with a slope coefficient of 1. So a failure to obtain a slope coefficient of 1 which means a rejection of the expectation hypothesis implies that the return are predictable. However, obtaining a slope coefficient of 1 (not rejecting the expectation hypothesis) does not say anything about the predictability of return. Thus, it is compatible to have a validation of the EH and a predictability of return as we do here.

In the Chapter 4, we analyse the Generalized Method of Moments (GMM) estimation and inference of the structural parameters of the LRR model that allows for the separation between the consumer optimal decision's frequency and the frequency by which the econometrician observes the data. Following Bansal et al. [2007b, 2012b, 2016], we use temporal aggregation to match the two decision frequencies and compute the unconditional moments of the variables in the model. Our contribution to the literature is twofold. First our inference procedure is robust to weak identification. Indeed, the Elasticity of Inter temporal Substitution is not well identified in this model and applying the standard method (the Delta method) for inference might lead to biased confidence intervals. Applying the method provided by Stock and Wright [2000], we construct confidence intervals for each parameter under the hypothesis that the parameter of interest might be weakly identified and while the other might be strongly identified. Second, we analyze the predictability implications of the model for the excess returns and the cash flows growth. This second point is made to address the critics of Beeler and Campbell [2012] stating that the LRR implies too much predictability of the cash flows growth and low predictability of excess returns compared to what is observed in the data. The key finding is that the Long Run Risk model adapts well to the data (the model is not rejected with basic moment conditions) but could not be so good at forecasting or telling the true story about what drives the evolution of asset prices (it becomes rejected when we add predictability moment conditions). Indeed, the model is able to reproduce the qualitative behavior of targeted moments in the long run when the corresponding estimates of the structural parameters are used for simulations, but it also faces a urge tension in keeping in track with all the observed moments considered.

1 Macro Uncertainty and the Term Structure of Risk Premium

Leading frictionless consumption-based asset pricing models (long-run risks and Habit formation) predict that the expected return on assets whose cash flows appear in the distant future are higher than or equal to the expected returns on assets which pay-off in the near future. Contrary to that prediction, some recent empirical studies have found that short-term assets earn a higher expected return than long-term assets. Here, I show that allowing the cash flows to be negatively affected by volatility shocks, as observed in the data ("leverage effect"), could make the short-term assets riskier than long-term assets. This modification gives more flexibility to those models in capturing various shapes of the term structure of equity returns while still matching the observed level of the equity premium and the risk free rate.

1.1 Introduction

During the past 30 years, lots of efforts in asset pricing research have been devoted to solve the equity and risk free rate puzzles[1] posited by Mehra and Prescott [1985], Weil [1989a]. The leading successful frictionless models (the Habit formation model by Cochrane and Campbell [1999], the long-run risks, henceforth LRR model by Bansal and Yaron [2004] and the Rare disasters model by Barro [2006]) are able to match the observed historical high equity premium level and low risk free rate level with a relatively low value of the risk aversion coefficient. More recently, researchers have been investigating the term structure of equity returns[2], meaning how the holding-period return evolves as a function of the time to maturity of the cash flow. This question can also be seen as a term structure decomposition of the equity risk premium; Indeed we can view a stock index as a portfolio of zero-coupon equity with maturities ranging from 1 to infinity, each zero-coupon equity paying a unit of the realized stock dividends only at its maturity and nothing else. Some recent empirical studies have found that both the term structure of the one-period risk premium[3] and the Sharpe ratio[4] are downward sloping, meaning that everything else kept equal, assets which pay-off in the short-run earn a higher risk premium per period compared to assets which pay-off in the distant future. In other words, the value of the short-term asset is lower compared to the value of the long-term asset. This observation appeared to be at odd with the predictions of the leading asset pricing models where the term structure of equity risk premium is upward sloping (van Binsbergen et al. [2012], Maggiori et al. [2015]) and it challenges the common sense since we expect cash flows falling in the closed future to be more predictable and less risky compared to those occuring in the distant future. Even though the discussions about the true average slope of the equity returns term structure have not yet been settled (Cochrane [2017], Bansal et al. [2017]), reconciling asset pricing models with

[1]These puzzles refer to the inability of the standard consumption based asset pricing model featuring a frictionless economy with a representative agent having constant relative risk aversion utility function and maximizing its lifetime expected utility by smoothing a log-normally distributed consumption, to match both the US financial market high level of the historical average risk premium which is around 6.5 % and the low level of the average risk free rate around 1 % with a reliable value of the risk aversion coefficient.

[2]van Binsbergen et al. [2012] used option prices data to recover dividend strip prices for different maturities and found a decreasing term structure of the equity risk premium. Maggiori et al. [2015] looked at real estate prices to extract the term structure of equity risk premium, more specifically they used data on the leasehold contracts for very long maturities and the freehold contracts on the housing market in England and Singapore, and they found a decreasing term structure of leasehold returns. See van Binsbergen and Koijen [2016] for a review.

[3]All along this paper, by risk premium I mean the log expected excess return as in Belo et al. [2015], which under conditional normality reduces to expected excess return adjusted for the Jensen inequality term.

[4]The Sharpe ratio as defined by William F. [1975] is the ratio of expected excess return over its standard deviation. It is reward-to-variability ratio giving the expected excess return per unit of risk for an asset or a portfolio.

the downward sloping term structure of equity returns has recently become a very active area of research (Belo et al. [2015],Hasler and Marfé [2016],Marfe [2016], Andries et al. [2017]).

This paper argues that accounting for the negative correlation between the level of the cash flow growth and its volatility (the "leverage effect") enables to generate a downward sloping term structure of equity returns in the leading asset pricing models while still matching the historical levels of equity premium and risk free rate. Indeed, we show that the negative correlation between the cash flow growth shocks and the stochastic volatility shocks pushes the distribution of the cash flows to the negative realizations, and this negative skewness is more pronounced in the short-run. Similarly in the observed data as illustrated on figure 1.4, consumption and dividends growths are more negatively skewed over one year horizon than when aggregated over many years horizon. Furthermore, as shown on figure 1.8, the term structures of the skewness for both consumption and dividend growths are negative in the short-run, and increasing with the horizon. The asymetrie between the positive and the negative cash flows (consumption or dividends) growth realizations in the short-run can be explained by the fact that negative cash flows shocks are amplified by a higher volatility, and on the contrary, positive cash flow shocks are dampened by a lower volatility. But since the volatility mean-reverts to its long term level, the "weight effect" on the negative cash flows growth shocks disappears as we aggregate over many periods thus shifting the distribution of the cash flows growths to the right in the long-run. So, in the short-run cash flows are more likely to be lower than in the long-run; henceforth assets that pay-off in the short-run appear to be riskier than assets that pay-off in the long-run.

My extension of the standard LRR model focuses on the cross-correlation between cash flows and volatility processes. In the specification of the cash flows and volatility dynamics, I allow volatility shocks to have a direct negative effect on consumption and dividend growth[5]; thus introducing a conditional (negative) correlation between the stochastic volatility and the cash flows. Compared to the standard LRR model, this specification magnifies the effect of the stochastic volatility by emphasizing its price of risk and increasing the short-term exposure of cash flows to the volatility risk. This modification brings in two additional parameters and enables to match the three recent stylized facts about the equity returns which are : The downward sloping term structure of the risk premium, the downward sloping term structure of the cash flows volatility and the downward sloping term structure of the Sharpe ratio. Indeed,

[5]Similarly, I could allow consumption growth shocks to have a direct negative effect on the volatility and obtain the same implications for asset prices since the same risks will still be priced in the equilibrium through the stochastic discount factor. So, what matters is the correlation not the causality.

in this model setup, the spread between the risk premium on dividend strips maturing at two consecutive periods can be expressed as a weighted sum of the price of long-run consumption shocks and the price of volatility shocks. I show that, while under preference for early resolution of uncertainty the weight on the former price of risk (long-run consumption shocks) is always positive, the weight on the price of volatility risk can be positive or negative. When the negative weighted price of volatility risk dominates (is greater in absolute value than) the positive weighted price of the long-run risk, the term structure becomes downward sloping. Therefore, my extension allows to remove the upward sloping constraint on the term structure of the risk premium and to add more flexibility to the LRR model in other to capture the recent stylized facts.

To empirically motivate my extension of the LRR model, I checked the negative correlation between shocks on macro uncertainty and cash flows in the data. For that, I used the VIX index as the main measure of macro uncertainty[6]. The VIX index is a measure of the risk neutral expectation of the next 30 days volatility implied by at-the-money S&P 500 index option prices. When investors expect a high volatility on the market in the next month, the VIX will shoot up; this happens especially during recession periods simultaneously with a drop of the cash flow as we can see from figure 1.1. I confirmed this visual observation by running a reduced form VAR model with VIX and cash flows growths showing that there is indeed a negative correlation between innovations in the VIX and in the cash flows. I then assessed the effect of macro uncertainty on asset prices by running a factor model regression on 10 Fama-French portfolios sorted by book-to-market ratio. I find that the VIX index has a negative beta that is more negative for value stocks than for growth stocks. It shows that value stocks, usually considered as short-duration assets, are more exposed to macro uncertainty compared to growth stocks which are long-duration asset. Thus macro uncertainty has the potential to explain the difference in returns between short and long duration asset, hence it can explain the downward slope of the term structure of the risk premium.

Following Bansal and Yaron [2004], Lettau and Wachter [2007], I solved the extended model and derived the key asset pricing equations for aggregate market index, risk-less asset and zero coupon equities. I then moved to the estimation of the structural parameters of the model following the method developed by Bansal et al. [2016], Meddahi and Tinang [2016], which consist of using temporal aggregation and the log-linear approximation to compute the theoret-

[6] I also checked that the negative correlation between uncertainty and cash flows still hold for other measures of uncertainty such as economic policy uncertainty index or the realized volatility on the S&P 500.

ical unconditional moments of the variables involved in the model (see Table 1.5.1). Then the theoretical unconditional moments are matched to their annual empirical counterparts to form the Generalized Method of Moments (GMM) objective function following Hansen [1982]. The structural parameters of the model are estimated by minimizing the objective function under the constraint of a decreasing term structure of the risk premium. The constraint is added for two reasons: First to see whether it is possible to find a valid vector of parameters (such that the model is not statistically rejected at the standard confidence levels) implying a decreasing term structure of the equity risk premium, and second because the moment conditions, more precisely the empirical counterparts, are not very informative about the term structure of the risk premium since I am not using the dividends strip data in the estimation. Thus, the constraint could be removed and replaced by some moment conditions on the dividend strips returns, but this came at the price of losing the closed form solutions of the theoretical moments and induced the use of the Simulated Method of Moment (SMM) for the estimation. This is left for future research.

Related literature

This paper makes a slight extension of the LRR model to show that it can match the decreasing term structure of the risk premium under certain conditions. In this view, it complements the work of Croce et al. [2014] who showed that without stochastic volatility the only way to achieve a decreasing term structure of the risk premium in a full information model is to have a dividend process that is less exposed to long-run consumption risk compared to the consumption process. This paper is closely related to Backus et al. [2016] who showed that a wide range of levels and shapes of the term structures of claims can be achieved by modifying the dynamics of the pricing kernel, of the cash flow growth and their interaction.

This paper also contributes to the literature on assessing the effects of uncertainty shocks through a structural model. Drechsler and Yaron [2011] provided a broad extension of the LRR model in order to match the returns predictability by the variance risk premium. They allowed for stochastic volatility and jumps in the innovations but similarly to the standard model by Bansal and Yaron [2004], cash flows processes are independent of volatility process in their calibration.

I find that assuming the negative correlation between uncertainty shocks and aggregate output shocks helps to explain the observed decreasing term structure of equity risk premium. The negative correlation between macroeconomic uncertainty and the economic conditions (output) that I bring in the LRR model, has been emphasized in the literature. Indeed, there is a feed back loop between macro uncertainty and the output. On one hand, as shown by Bloom [2009], an increase in uncertainty can decrease the output through the reduction of investments and hiring. On the other hand, a drop in the output might also increase the macroeconomic uncertainty because of the uncertainty about the actions that will be taken by the government to remedy to the situation or the uncertainty about whether the incumbent government will remain on sit (Kelly et al. [2016]). The mechanism at play in our model can be related to the "real options effect" described by Bloom [2014], stating that in the face of uncertainty shocks, economic agents (consumers or investors) will prefer to postpone their decision to consume or to invest, thus reducing short-run hiring, investments and in our model increasing short-run risk premium; but at the same time uncertainty shocks stimulate research and innovation hence increase the upside from innovative new product and reduces the long-run risk premium. Indeed, according to the option approach of investment (Dixit and Pindyck [1994]), because of the irreversibility of an investment and the possibility to delay it, the decision of investing is taken when the difference between the expected benefits and costs related to the investment exceeds the value of the (call) option to delay the investment. So an increase in uncertainty will increase the value of the option and thus, will reduce investment.

The change in the term structure of the risk premium could come from a combination of changes in the price of risk and in the quantity of risk embedded in the asset. Andries et al. [2017] generalize the Epstein and Zin [1989] (henceforth EZ) preferences by allowing a change of the risk aversion coefficient used to compute the continuation value of the stream of present and future consumption. The stochastic volatility in the cash flow process enables the horizon dependent risk aversion to affect the equilibrium risk prices and to reverse the term structure of risk premium compared to the Bansal and Yaron [2004] model. Contrary to them, I maintain the standard EZ preferences which imply constant prices of risk in my model. Belo et al. [2015] modified the dividend dynamics in order to be consistent with capital structure policies that generate stationary leverage ratios. The implied dynamic features the negative effect of volatility shocks on the dividend growth which is the main change compared to the BY model that I highlight in this paper as being important to obtain a decreasing term structure for the risk premium and the dividend growth volatility. Marfe [2016] provided a labor income insurance explanation of the decreasing term structure of risk premium. Lettau and Wachter [2007, 2011] specified an exogenous mean reverting process driving the price of risk and also determining the pricing kernel used in their model. The innovations on that process come from changes in the preferences or changes in sentiment. The cash flow in their model features the one present in the Bansal and Yaron [2004] model with a persistent component corresponding to the conditional mean of the dividend growth but they allow for a negative correlation between shocks on the dividend growth and on its conditional mean. Our model share some similarities with the one developed by Lettau and Wachter [2007, 2011]. Indeed, as they explained in their paper, one of the key ingredient to obtain a decreasing term structure of equity and thus being able to explain the value premium is the fact that dividend loads directly on shocks affecting the stochastic discount factor which are priced. In our model, there are three of such shocks and differently from Bansal and Yaron [2004], Drechsler and Yaron [2011][7], I allow dividend growth to load on all those shocks. Contrary to Lettau and Wachter [2007, 2011] who specified an exogenous stochastic discount factor to price assets, the stochastic discount factor derived in our model is micro-founded and comes from the representative agent inter-temporal optimal allocation plan.

The empirical evidences about the downward sloping term structure of equity premium are recent and still fragile. For example Cochrane [2017] pointed out the lack of statistical significance of the main result in van Binsbergen et al. [2012] which is the difference of expected returns between short-term assets and the market index. Bansal et al. [2017] found that dividend strip returns are increasing with maturity (1 to 7 years) and their empirical evidence supports the implications of leading equilibrium asset pricing models. So, having a flexible model that allows to capture both the increasing or the decreasing term structure of the risk premium could be a good starting point for testing the sign of the slope in the data.

The remaining of the paper is organized as follow: Section 2 presents on one side the link between the macro uncertainty and the cash flows (consumption and dividends), and on the other side the link between macro uncertainty and asset prices. Section 3 presents the model and its solution. Section 4 presents the derivation of the risk premium term structure formulas for the dividend strips. Section 5 presents the estimation of the model. Section 6 studies the timing of risk implied by the dynamics in our model. Section 7 goes deeply in understanding the drivers of the term structure slope. Section 8 presents some simulations of the risk premium term structure and discusses the implications concerning the unobserved component of the wealth portfolio on

[7] They only allow dividend to load on short-run consumption growth shocks, while the most important loading in our case is on the volatility shocks.

top of the financial assets (usually regarded as the human capital). It also presents the economic policy implication for the pricing of long-term investment project. Finally, section 9 concludes.

1.2 Empirical support for the negative link between macro uncertainty and cash flows

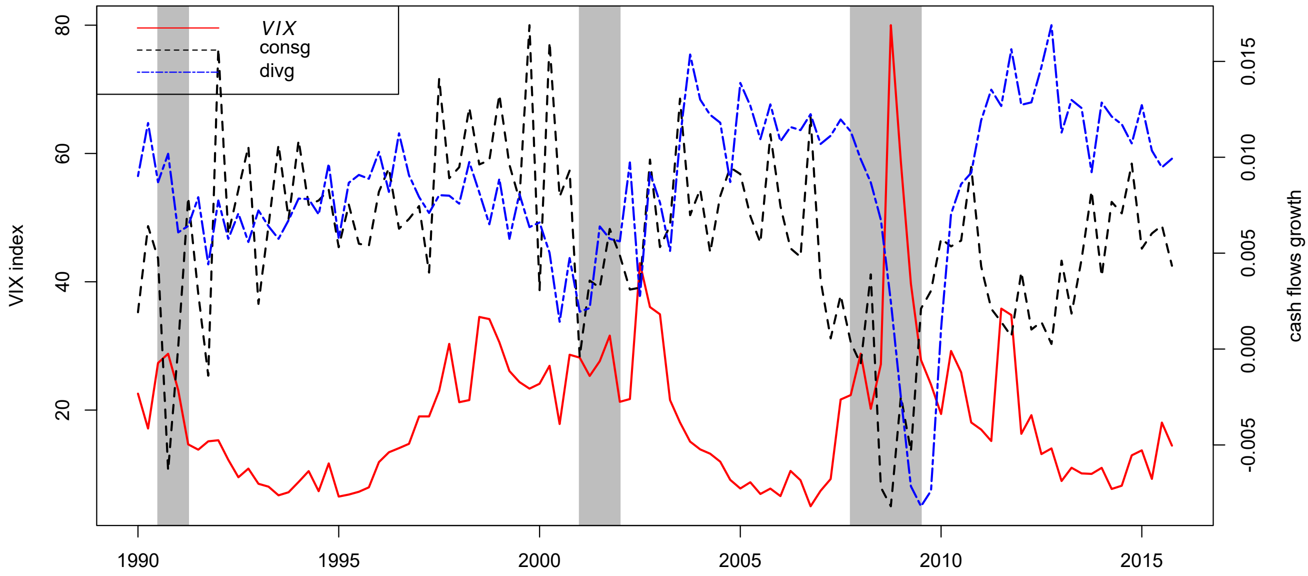
In this section, I provide the empirical support for the negative correlation between innovations on uncertainty, proxied by the VIX index, and cash flows shocks.

Macroeconomic effects of uncertainty have been the focus of many researchers during the recent decade. Bloom [2009] presented a set of uncertainty indicators (stock-market volatility, cross-sectional standard error of firm's pretax profit growth, standard deviation of industry TFP growth, dispersion across macro forecasters over the predictions for future gross domestic product) which he found to be positively correlated; all those indicators overshoot during recessions. Both at the micro and the macro levels, uncertainty rises during recessions and declines during booms (Bloom [2009], Bloom [2014]).

Our empirical analysis focuses on the VIX index as the indicator of macro uncertainty. The VIX index is an estimate of the next 30-day expected volatility on the S&P 500 index provided by the Chicago Board of Options and Exchange (CBOE). It is computed by averaging the weighted prices of the S&P 500 index put and call option prices over a wide range of strike prices (see CBOE [2015]). As the VIX index represents the risk neutral expectation of the future volatility of the market stock index, it shows the investors perception of the future risk. Figure 1.1 shows the evolutions of the VIX index, aggregate dividends growth rate, and the growth rate of consumption expenditures on non-durables and services. It shows that during recessions macro uncertainty increases while consumption expenditures decrease (the growth rate of consumption expenditures becomes negative). This visual observation is confirmed by a reduced form VAR model that I run in order to estimate the correlation between innovations in the VIX index and innovations in the growth rate of consumption expenditures. The results reported in Table 1.5 show that the two innovations are negatively correlated, meaning that a positive shocks on the VIX index (a raise in macro uncertainty) will decrease the growth rate of consumption expenditures and vice-verse.

Sources: Author using data from the Bureau of Economic Analysis, the Chicago Board of Options and Exchange and the National Bureau of Economic Research. Notes: This figure shows the evolutions of the dividend growth and the consumption growth on non-durable goods and services (right axis) and the VIX index of 30-day implied

Figure 1.1: Cash flows growth and Stock market Implied Volatility



volatility on the Standard & Poor's 500 stock market index (left axis). I use quarterly data from the first quarter of 1990 to the last quarter of 2015. Gray bars are NBER recessions.

There is also a negative correlation between asset returns and uncertainty. Indeed, Ang et al. [2006] found that the aggregate volatility of the market is negatively priced in the cross-section of expected stock returns and thus, stocks with large, positive sensitivities to volatility risk should have low average returns. We check that using the monthly value weighted returns on deciles book-to-market sorted portfolios. Value stocks have a higher (in absolute value) negative exposure to the VIX compared to growth stocks. As table 1.9 shows, stocks with a high Book to Market ratio (value stock) appeared to be less exposed to the VIX index; they have a negative and significant VIX-beta compared to low Book to Market ratio's stocks (growth stocks). Thus everything else being kept equal, an increase in the market uncertainty reduces more the expected excess return on value stocks compared to growth stocks.

This finding goes in the same direction of Lettau and Wachter [2007] who related the value premium to the cash flow duration. Contrary to them, I relate the value premium to the cash flow exposure to volatility shocks; the more the exposure the lower the expected return. Thus the value premium rewards stocks dividend's exposure to volatility risk, as positive shocks on volatility (increase in uncertainty) lead to a more pronounce drop in value stocks dividend. As we know from Lettau and Wachter [2007], a model (or a variable) that explains the value premium might also be able to reproduce the downward slope of the term structure of equity returns since Value stocks with a high Book-to-Market ratio are associated to short-duration assets while Growth stocks with a low Book-to-Market ratio correspond to long-duration assets.

1.3 The model

The empirical investigations in the previous section show that the VIX (which is a proxy for macro uncertainty) is negatively correlated with the aggregate consumption and dividends. In this section, I specify a consumption-based asset pricing with macro uncertainty introduced through the stochastic volatility driving the consumption and dividends growth processes. My specification of the cash flows dynamics follows the one of Bansal and Yaron [2004] but allows for a negative correlation between the innovations on the level of the cash flows and the innovation on the stochastic volatility process. As I will make it clearer in the next sections, this negative correlation makes the cash flows to be left skewed and more risky. More important, the skewness term structure is increasing such that in the short term cash flows are more risky compared to the long term; hence delivering a decreasing term structure of the risk premium.

1.3.1 Preferences and cash flow dynamics

I consider a rational representative agent embedded with Epstein and Zin [1989] recursive utility function[8] given by equation 1.3.1 who maximizes its continuation value subject to its inter-temporal budget constraint.

$$V_t = \left[(1-\delta)C_t^{1-\frac{1}{\psi}} + \delta\left(E_t\left(V_{t+1}^{1-\gamma}\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right)\right]^{\frac{1}{1-\frac{1}{\psi}}} \tag{1.3.1}$$

Where δ is the pure discount factor, γ is the relative risk aversion coefficient and ψ is the Elasticity of Inter-temporal Substitution (EIS). This preference specification allows to disentangle the EIS from the risk aversion coefficient and to break the tight link imposed between them by the time additive preference where $\gamma = \frac{1}{\psi}$.

[8]The full definition of the utility function is given by equation 1.10.1.

The cash flows dynamics are given by:

$$\Delta c_{t+1} = \mu_c + x_t + \sigma_t \varepsilon_{c,t+1} + \varphi_\sigma \sigma_w \varepsilon_{w,t+1} \tag{1.3.2}$$

$$x_{t+1} = \rho x_t + \varphi_e \sigma_t \varepsilon_{x,t+1} \tag{1.3.3}$$

$$\sigma^2_{t+1} = \nu \sigma^2_t + (1-\nu)\bar{\sigma}^2 + \sigma_w \varepsilon_{w,t+1} \tag{1.3.4}$$

$$\Delta d_{t+1} = \mu_d + \phi x_t + \pi_c \sigma_t \varepsilon_{c,t+1} + \pi_\sigma \sigma_w \varepsilon_{w,t+1} + \varphi_d \sigma_t \varepsilon_{d,t+1} \tag{1.3.5}$$

$$(\varepsilon_{c,t+1}, \varepsilon_{x,t+1}, \varepsilon_{d,t+1}, \varepsilon_{w,t+1}) \sim N.i.id(0, I)$$

Where $\varepsilon_{c,t+1}$,$\varepsilon_{d,t+1}$,$\varepsilon_{x,t+1}$ and $\varepsilon_{w,t+1}$ represent respectively the short-run consumption growth shock, the short-run dividend growth shock, the expected (or long-run) consumption growth shock and the consumption growth volatility shock. The consumption growth equation 1.3.2 assumes that at time $t+1$, it depends on the past period expected consumption growth $\mu+x_t$ and it is affected by both the current short-run shock and the stochastic volatility shock. Equation 1.3.3 describes the persistent component of the consumption growth process as an AR(1) process. Equations 1.3.4 presents the stochastic volatility process. Finally equation 1.3.5 describes the dividend growth process as a levered consumption. All the shocks are assumed to be normally and independently distributed with 0 as mean and 1 as standard deviation.

The dynamics described by equations 1.3.2-1.3.5 embed[9] the one present in the LRR model of Bansal and Yaron [2004] as a special case. The key difference lies in the facts that I introduce the possibility of explicit correlations between the stochastic volatility process and the cash flows dynamics. These correlations, highlighted in red in equations 1.3.2 and 1.3.5, when negative imply a negatively skewed cash flows growth which are more consistent with observed consumption and dividend growth processes[10].

1.3.2 Model's solution

In order to derive asset prices formulas in closed form while avoiding the use of the log-linear approximation, I first restrict myself to the case where the representative agent has an EZ utility function with EIS=1. The formulas using the log-linear approximation for the case with EIS≠1 are derived in the appendix. The solution to the model is standard in the literature. Assuming that the log-value consumption ratio vc_t is an affine function of the state variables :

[9]The dynamics in Bansal and Yaron [2004] can be obtained by setting the restrictions: $\varphi_\sigma = \pi_\sigma = 0$
[10]See the summary statistics in Table 1.3

$$vc_t = A_0 + A_1 x_t + A_2 \sigma_t^2 \tag{1.3.6}$$

and substituting 1.3.6 into the log-price consumption ratio allows to obtain the coefficients of the affine function as follows:

$$\begin{aligned}
A_1 &= \frac{\delta}{1-\delta\rho} \\
A_2 &= \frac{\delta(1-\gamma)}{2(1-\delta\nu)} \left[1 + \left(\frac{\delta\varphi_e}{1-\delta\rho} \right)^2 \right] \\
A_0 &= \frac{\delta}{1-\delta} \left[\mu_c + A_2(1-\nu)\bar{\sigma}^2 + \frac{1}{2}(A_2+\varphi_\sigma)^2\sigma_w^2 \right]
\end{aligned} \tag{1.3.7}$$

So under the standard calibrations ($\gamma > 1$ and $0 < \delta < 1$), the loading on the expected consumption growth (A_1) is positive and the loading on the volatility (A_2) is negative. This means that positive shocks on the expected consumption growth (respectively on the volatility of the consumption growth) increases the value to consumption ratio (respectively decreases the value to consumption ratio). Thus the representative consumer is better off when a positive shock on the expected consumption happens (which means a better future prospect of consumption) and she is worse off when a positive shock on the volatility of consumption growth happens (which means more uncertainty surrounding future consumption growth).

The log of the stochastic discount factor (sdf) is expressed as a function of the states variables and the different shocks priced in the model multiplied by their prices of risk.

$$m_{t+1} = a_{0m} + a_{1m} x_t + a_{2m} \sigma_t^2 + \lambda_c \sigma_t \varepsilon_{c,t+1} + \lambda_x \sigma_t \varepsilon_{x,t+1} + \lambda_w \sigma_w w_{t+1} \tag{1.3.8}$$

Where

$$\begin{aligned}
a_{0m} &= \log(\delta) - \mu_c - \frac{1}{2}(\gamma-1)^2(A_2+\varphi_\sigma)^2\sigma_w^2 \\
a_{1m} &= -1 \\
a_{2m} &= -\frac{1}{2}(1-\gamma)^2 \left[1 + \left(\frac{\delta\varphi_e}{1-\delta\rho} \right)^2 \right]
\end{aligned}$$

The price of the short-run consumption risk, the price of the long-run consumption risk which prices shocks happening to the expected consumption growth and the price of the volatility risk are respectively given by 1.3.9 , 1.3.10 and 1.3.11.

$$-\lambda_c = \gamma \tag{1.3.9}$$

$$-\lambda_x = (\gamma - 1)\frac{\delta\varphi_e}{1-\delta\rho} \tag{1.3.10}$$

$$-\lambda_w = \gamma\varphi_\sigma + (\gamma - 1)A_2 \tag{1.3.11}$$

On one hand, as in standard calibrations of the LRR model, I maintain the rescaling parameter φ_e that governs the variance of the expected consumption growth to be positive and the prices of the short and the long-run risks on consumption growth are both positive. So positive consumption growth shocks will decrease the log stochastic discount factor, thus reducing the marginal utility of the consumer who will be better off in that case. On the other hand, I expect that a positive shock on the volatility of consumption growth (an increase of macro uncertainty) will have a negative effect on the consumption growth ($\varphi_\sigma < 0$). This assumption is supported by both the data (see Table 1.5) and the literature on the effects of macroeconomic uncertainty (Bloom [2009]). It means that during bad times (recessions), cash flows fall because of the negative shocks on the aggregate output but also because of the increase in the aggregate uncertainty. So compared to the model in Bansal and Yaron [2004] where the preference for early resolution of uncertainty implies a negative price of volatility risk, here that price will be further negative because of the negative exposure of the consumption growth to volatility shocks. A positive volatility shock increases the marginal utility of the consumer who is then worse off in that case. This increase in absolute value of the price of volatility risk due to direct consumption growth exposure to volatility shocks is consistent with the intuition in Boguth and Kuehn [2013] stating that stocks with volatile cash flows in uncertain aggregate times require higher expected returns.

The return on the risk free asset at time t is determined through the pricing kernel by:

$$r_{f,t} = -a_{0m} - \frac{1}{2}\lambda_w^2\sigma_w^2 - a_{1m}x_t - \left(a_{2m} + \frac{1}{2}(\lambda_c^2 + \lambda_x^2)\right)\sigma_t^2 \tag{1.3.12}$$

Compared to the standard BY model, only the constant term has changed; it has become smaller because of the negative exposure of the consumption growth to the volatility shocks. Equation 1.3.12 shows that the precautionary motive to save is higher compared to the one in the standard BY model because of the additional exposure of the consumption to the volatility risk and the risk free rate is lower since the consumer prefers to shift its wealth portfolio toward the risk-less asset.

The return on the aggregate market portfolio can be obtained[11] in closed form using the Campbell and Shiller [1988]'s log-linear approximation. Let us denote by P_{t+1} the price at $t+1$ of the market portfolio and by D_{t+1} its dividend, then the log-return on the market portfolio is given by:

$$\tilde{r}_{m,t+1} = \log\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right) \approx k_{0,m} + k_{1,m} z_{m,t+1} - z_{m,t} + \Delta d_{t+1} \tag{1.3.13}$$

The market risk premium at time t and the excess return volatility are respectively given by 1.3.14 and 1.3.15[12].

$$rp_{m,t} = -\lambda_w \beta_{m,w} \sigma_w^2 - (\lambda_c \beta_{m,c} + \lambda_x \beta_{m,x})\, \sigma_t^2 \tag{1.3.14}$$

$$\sigma_{m,t} = \sqrt{\beta_{m,w}^2 \sigma_w^2 + \left(\beta_{m,c}^2 + \beta_{m,x}^2 + \beta_{m,d}^2\right) \sigma_t^2} \tag{1.3.15}$$

Where $\beta_{m,c}$,$\beta_{m,x}$,$\beta_{m,w}$,$\beta_{m,d}$ are given by 1.10.51 in the appendix .

1.4 The term-structures of equity and bond returns

1.4.1 The term structure of zero-coupon equity

The zero coupon equity or dividend strip is a derivative contract that pays a unit of the underlying asset realized dividends once at maturity. Let us denote $P_t^{(n)}$ the price at time t of a zero coupon equity maturing n-periods later at time $t+n$. From the no-arbitrage condition, $P_t^{(n)}$ should satisfy the Euler condition stating that the price at time t of the n-periods dividend strip is the discounted expected value of the one period ahead future price of the $n-1$-periods dividend strip:

$$P_t^{(n)} = E_t\left[M_{t+1} P_{t+1}^{(n-1)}\right] \tag{1.4.1}$$

with the boundary condition that :

$$P_t^{(0)} = D_{t+1} \tag{1.4.2}$$

[11] Another method to compute this return uses the fact that the market portfolio is made by dividend strip at all maturities, thus its price is the infinite sum of dividend strips prices for all maturities. See 1.10.3 for the derivation.

[12] See Appendix 1.10.2 for details

Let $R_{d,t+1}^{(n)}$ denotes the one holding-period return at time $t+1$ of a zero-coupon equity that matures n-periods and $r_{d,t+1}^{(n)} = \log R_{d,t+1}^{(n)}$. Then, we have :

$$R_{d,t+1}^{(n)} = \frac{P_{t+1}^{(n-1)}}{P_t^n} \tag{1.4.3}$$

We then deduce the risk premium on the n-period dividend strip as follows:

$$rp_t^{(n)} = -\lambda_w \left(\pi_\sigma + A_2(n-1)\right)\sigma_w^2 + \left[-\lambda_c\pi_c - \lambda_x\left(A_1(n-1)\varphi_e\right)\right]\sigma_t^2 \tag{1.4.4}$$

So the risk premium is a weighted sum of the risk prices. Under preference for early resolution of uncertainty ($\gamma > 1/\psi$), the price of the volatility risk is negative while the prices of long-run and short-run consumption risks are both positive.

As $A_1(0) = 0$ and $A_2(0) = 0$, the risk premium on the 1-period dividend strip return is :

$$rp_t^{(1)} = -\lambda_w\pi_\sigma\sigma_w^2 - \lambda_c\pi_c\sigma_t^2 \tag{1.4.5}$$

Compared to the standard BY model, there is one new term in the one month risk premium because of the cross correlation between cash flow shocks and the volatility shocks. When $\pi_\sigma < 0$, meaning that the dividend growth reacts negatively to an increase of uncertainty, the short-run risk premium is higher compared to the case where the cross correlation is not taken into account, since in the face of more risk, investors require a higher risk premium to bear the risks. So, the correlation of the dividend growth process and the volatility is very important to increase the short-term risk premium which helps to match the historical level of the equity premium, especially when the term structure of risk premium is decreasing.

Proposition 1 (Short term spread) : The one period term spread defines as the difference between the n-periods dividend strip risk premium and the $n-1$-periods dividend strip risk premium is given by (for $n \geq 2$) :

$$\begin{aligned} S_{n,t}^{(1)} &= rp_t^{(n)} - rp_t^{(n-1)} \\ &= -\lambda_w\left(A_2(n-1) - A_2(n-2)\right)\sigma_w^2 - \lambda_x\left(\left(\phi - \frac{1}{\psi}\right)\varphi_e\rho^{n-2}\right)\sigma_t^2 \end{aligned} \tag{1.4.6}$$

where

$$A_2(1) = \frac{1}{2}\left(\pi_c - 1\right)\left(\pi_c + 1 - 2\gamma\right) + \frac{1}{2}\varphi_d^2$$

This term spread can be expressed as the weighted sum of the price of the volatility shock and the price of the expected consumption growth shock. As we know, the price of volatility risk is negative while the price of expected consumption growth risk is positive. When the weight on each price is positive, the sign of the short term spread will depend on which weighted price dominates the other : It is negative if the weighted price of the volatility risk is greater (in absolute value) than the weighted price of the expected consumption growth risk. The key equation 1.4.6 allows to pin down the sign of the slope of the term structure of equity risk premium. The weight on the (negative) price of volatility risk is driven by the sequence $A_2(n)$ which is positive and increasing.

Notice that the short term spread here is time varying because of the conditional volatility. It is determined as a weighted sum of the price of the volatility risk (which is negative) and the price of the long-run consumption risk (which is positive). The weight on the former depends on the volatility of the volatility (σ_w , which is constant) and the weight on the later depends on the conditional volatility (σ_t, which is time varying). So assuming that the two weights are positive, when the expected volatility increases at time t , the short term spread increases because of the increase in the weighted price of the long-run consumption risk. Thus the equity risk premium term-structure becomes more upward sloping or can move from downward to upward. On the contrary, when the expected volatility at time t decreases, the negative weighted price of the volatility risk becomes more important and, the short term spread decreases and can become negative. So, the slope of the term structure of the equity risk premium is counter-cyclical meaning that it is upward sloping in bad time (when the expected volatility is high) and downward sloping in good time (when the expected volatility is low).

Let us now look closer at the weights on the prices of risks in order to see under which conditions they are positive. Firstly, for $n = 2$, $S^{(1)}_{2,t} = -\lambda_w A_2(1)\sigma^2_w - \lambda_x \left(\phi - \frac{1}{\psi}\right) \varphi_e \sigma^2_t$ as the prices of the volatility risk and the expected consumption growth risk are respectively negative and positive, a necessary condition for $S^{(1)}_{2,t}$ to be negative is that $A_2(1)$ should be positive or ϕ being lower than the inverse of the EIS. On one hand for $A_2(1)$ to be positive, we need that the dividend growth loading on short-run consumption growth shock should be either "very low" ($\pi_c < 1$) or "very big" ($\pi_c > 2\gamma - 1$). It means that the short-run consumption growth shock should either have a lower effect (scaled down) or a very big effect (scaled "very" up) on the dividend growth. In the estimation, it happened that the loading of the dividend on short-run consumption growth is low. Indeed a very high loading of dividend on consumption growth will result into too much correlation between consumption and dividends, but with its low exposure to short-run consumption growth shock and its additional exposure to volatility shocks, the dividends can still command a high risk premium in equilibrium without being too much correlated with consumption. On the other hand, ϕ being lower than $\frac{1}{\psi}$ would mean that dividend growth load less on the expected consumption growth and thus the market portfolio will be less risky compared to the wealth portfolio as far as long-run risks on consumption growth are considered.

Secondly, notice that the weight on the price of the long-run consumption growth risks depend on the persistence of the expected consumption growth process (ρ). Thus with a lower persistence parameter[13], the (positive) contribution of the price of long-run consumption growth to the one period spread will quickly fade away such that in the long-run the term structure of the spread will only be governed by the weighted price of volatility risk. In the extreme case where there is no expected consumption growth persistence ($\rho = 0$), the one period spread given by 1.4.7, after the second period only depends on the price of volatility risk and thus will be negative when the volatility risk price's weight is positive. We can easily see that when the dividend loadings on the short-run consumption growth shock and on the expected consumption growth are sufficiently high or low enough, the weight on the price of volatility risk will be positive.

$$S_{n,t}^{(1)} = -\frac{1}{2}\lambda_w \nu^{n-3} \left(\left[(\pi_c - 1)(\pi_c + 1 - 2\gamma) + \varphi_d^2 \right] \nu + \varphi_e^2 \left(\phi - \frac{1}{\psi} + 2(1-\gamma)\delta \right) \left(\phi - \frac{1}{\psi} \right) \right) \sigma_w^2 \quad (1.4.7)$$

To further understand the key ingredients that drive the slope of the term structure of equity risk premium, let us restrict ourselves to the case of an asset whose dividend growth shares the same long-run component as consumption growth, meaning that $\phi = 1$ and thus $\phi - \frac{1}{\psi} = 0$ as we assumed $\psi = 1$. So, the short term spread (the difference on equity risk premium for dividend strips with two consecutive maturities) in (1.4.6) becomes[14] :

$$\begin{aligned} S_{n,t}^{(1)}|_{\phi=1} &= -\lambda_w \left(A_2(1)\nu^{n-2} \right) \sigma_w^2 \\ &= -\lambda_w \left(\frac{1}{2}(\pi_c - 1)(\pi_c + 1 - 2\gamma) + \frac{1}{2}\varphi_d^2 \right) \nu^{n-2} \sigma_w^2 \end{aligned} \quad (1.4.8)$$

Equation 1.4.8 shows that the short term spread here only depends on the volatility of volatility

[13] For example, taking half of the value of the persistence parameter usually used in calibration (0.987/2), will move the half-life of the weighted price of the expected consumption growth risk from 4.5 years to 2 months.

[14] In the case this asset mimics the wealth portfolio and delivers consumption bundles as dividends, the term structure will be flat given that we will have $\pi_c = 1$ and $\varphi_d = 0$.

and its sign is driven by a combination of preference parameters (pure discount factor and risk aversion) and cash flows parameters. As I maintain $-\lambda_w < 0$, having the loading of the dividend growth on the consumption growth shock (π_c) "small" or "large"[15] will make the short term spread constant and negative, which implies a decreasing term structure for the risk premium. Furthermore, the curvature of the equity yields curve is driven by the persistence of the volatility (ν).

Proposition (Long term spread or term-structure slope) : The long term spread between n-periods dividend strip return and 1-period dividend strip return and the infinite maturity term spread are respectively given by 1.4.9 and 1.4.10.

$$S_t^{(n)} = rp_t^{(n)} - rp_t^{(1)} = -\lambda_w \left(A_2(n-1)\right)\sigma_w^2 - \lambda_x A_1(n-1)\varphi_e\sigma_t^2 \tag{1.4.9}$$

$$S_t = \lim_{n\to\infty} S_t^{(n)} = -\lambda_w \left(A_2(\infty)\right)\sigma_w^2 - \lambda_x A_1(\infty)\varphi_e\sigma_t^2 \tag{1.4.10}$$

where

$$A_1(\infty) = \left(\frac{\phi - \frac{1}{\psi}}{1-\rho}\right)$$
$$A_2(\infty) = \frac{1}{2(1-\nu)}\left[(\pi_c - 1)(\pi_c + 1 - 2\gamma) + \varphi_d^2 + \left(\frac{\phi - \frac{1}{\psi}}{1-\rho}\right)\left(\frac{\phi - \frac{1}{\psi}}{1-\rho} + 2\frac{(1-\gamma)\delta}{1-\delta\rho}\right)\varphi_e^2\right]$$

Given that $A_1(\infty) \geq 0$, the term-structure of the equity risk premium is downward sloping if and only if $A_2(\infty) > 0$ and $\left|-\lambda_w\left(A_2(\infty)\right)\sigma_w^2\right| > \left|-\lambda_x A_1(\infty)\varphi_e\sigma_t^2\right|$.

Notice that when the volatility is held constant ($\sigma_w = 0$), the long term spread between n-periods dividend strip return and 1-period dividend strip return is given by equation 1.4.11 and we can see that the term spread will be positive once $\phi > \frac{1}{\psi}$ and it increases with the horizon. Proposition 1.4.1 complements the result highlighted by Croce et al. [2014] stating that in a full information "long-run risks" model with constant volatility, as soon as the loading of the risky asset under consideration on the expected growth rate of consumption is greater than 1 ($\phi > \frac{1}{\psi}$), the term structure of equity return is always upward sloping[16]. With the stochastic volatility

[15] The loading of dividend growth on consumption growth is considered "small"; meaning that when $\pi_c < 1$. In that case, the dividend is less exposed than the consumption to the short run consumption shock. The loading of dividend growth on consumption growth is considered "large" when $\pi_c > 2\gamma - 1$.

[16] Indeed the term-spread becomes:

$$S_t^{(n)} = \left[-\lambda_x\varphi_e\right]\left(\phi - \frac{1}{\psi}\right)\left(\frac{1-\rho^n}{1-\rho}\right)\bar{\sigma}^2 \tag{1.4.11}$$

in the consumption growth, the term structure of equity return can be downward sloping even when the dividend growth loads more on the long run risk than the consumption growth.

1.4.2 The real yield curve

Given the conditional normality of the state variables and the sdf, the real price at t of a zero coupon bond maturing n-period later ($P_t^{B,(n)}$) is an affine function of the state variable.

$$P_t^{B,(n)} = \exp\left(B_0(n) + B_1(n)x_t + B_2(n)\sigma_t^2\right) \tag{1.4.12}$$

Using the no-arbitrage condition, the price of the zero coupon bond can be expressed recursively as follows:

$$P_t^{B,(n)} = E_t\left(M_{t+1}P_{t+1}^{B,(n-1)}\right) \tag{1.4.13}$$

Equation 1.4.13 can be used with the boundary condition that $P_t^{B,(0)} = 1$ to solve for the unknown coefficients in the closed form expression in equation 1.4.12 for $P_t^{B,(n)}$ given that the zero coupon bond pays a single unit at maturity.

The coefficient in front of the expected consumption growth is given by:

$$B_1(n) = -\left(\frac{1-\rho^n}{1-\rho}\right) \tag{1.4.14}$$

The coefficient in front of the volatility is given by the recursive formula:

$$\begin{aligned} B_2(n) = B_2(n-1)\nu - \frac{1}{2}(1-\gamma)^2\left[1+\left(\frac{\delta\varphi_e}{1-\delta\rho}\right)^2\right] - \frac{1}{2}\gamma^2 \\ + \left(\frac{\varphi_e^2}{\nu-\rho}\right)\left[\frac{(\gamma-1)\delta}{1-\delta\rho} + \frac{1}{1-\rho}\right]\left(\frac{1-\nu^n}{1-\nu} - \frac{1-\rho^n}{1-\rho}\right) \end{aligned} \tag{1.4.15}$$

With the starting conditions that $B_2(0) = 0$ and $B_2(1) = (-\frac{1}{2} + \gamma)$. The constant term is given by equation 1.10.7 .

As equation 1.4.12 shows, bond price reacts negatively in response to an increase in the expected growth and the decline in the bond prices is more pronounced for longer maturities. This is the duration effect described by Lettau and Wachter [2011], meaning that given the persistence of the expected consumption growth, a higher expected consumption growth today predicts that the future expected growth will also be high. So, longer maturity bond's prices will be more affected and thus will decrease more in response to an increase in the expected consumption growth compared to shorter maturity bond's prices. As equation 1.4.15 shows, the effect of volatility on the bond price depends on the parameters governing the dynamics of the consumption growth process and on the prices of current and expected consumption growth risks. For example, in the case of i.i.d consumption growth with constant volatility ($\rho = \nu = 0$), the bond price is the same for all the maturities and it increases with the volatility[17]. This is not surprising given that a higher macroeconomic uncertainty makes the bonds become more attractive for the precautionary motive. When I allow for the persistence in the expected consumption growth process or the persistence in the stochastic volatility to enter into play, the bond price loading on the volatility is positive and increases with the maturity.

The yield to maturity on a real zero coupon bond is linear in the state variables and is given by:

$$y_t^{(n)} = -\frac{1}{n}\log P_t^{B,(n)} = -\frac{1}{n}\left[B_0(n) + B_1(n)x_t + B_2(n)\sigma_t^2\right] \tag{1.4.16}$$

Notice that the zero-coupon bond can be seen as a "zero-coupon equity" with a fixed dividend. So, the formulas for the zero coupon bond are special cases of the zero coupon equity formulas with restrictions on the parameters making the dividend to be a random variable ($\pi_c = \varphi_d = \phi = \mu_d = 0$). So from 1.4.1 and 1.4.1, it follows that the bond term premium is negative since the volatility risk and the long run consumption risk would induce investors to prefer long term bonds compared to short-term bonds. The price of the long term bonds will increase relative to the price of short-term bonds; thus the yield curve will be downward sloping.

1.5 Estimation and results

In this section, I provide two approaches for the estimation of the structural parameters of the extended model. The first approach uses the Generalized Method of Moments (GMM) method with and without the constrained of the negative term structure of the risk premium. The second method uses the Simulated Method of Moments and includes among the moment conditions the difference in expected returns on dividend strips with different maturities.

1.5.1 Estimation

We used annual data on the U.S consumption of non-durable goods and services, the aggregate dividend growth of the S&P 500 index, the aggregated market return, the 3-month T-bill rate

[17] Under the standards calibration, $B_2(1)$ is positive. It could be negative under the assumptions of a very low risk aversion coefficient ($\gamma < \frac{1}{2}$) and/or a very negative loading of the consumption growth on the expected consumption growth shocks.

and the price dividend ratio. The data span the period from 1926 to 2016.

The parameters of the model are estimated using the Generalized Method of Moments (GMM) under constraints. More specifically, the moment conditions have been formed to match first and second order unconditional moments observed in the data. I follow the same procedure used by Bansal et al. [2016], Meddahi and Tinang [2016] which allows to temporally aggregate theoretical moments to match the quarterly or annually observed ones. The constraint of the negativity of the long term spread is then added to guarantee that the estimates of the model's parameters which minimizes the GMM objective function should also imply a decreasing term structure of equity risk premium. Thus, our estimation procedure consist of looking in the set of parameter's vectors that deliver a decreasing term structure of the equity risk premium, the (best) one which minimizes the distance between the theoretical moments derived from the model and the empirical moments computed using the observed data. Hence, we solve the following problem:

$$\hat{\zeta} = \underset{\zeta \in \Theta}{\arg\min}\, T \left[\frac{1}{T} \sum_{t=1}^{T} h(y_t, \zeta) \right]' W_T(\bar{\zeta}_T(\zeta))^{-1} \left[\frac{1}{T} \sum_{s=1}^{T} h(y_s, \zeta) \right] \tag{1.5.1}$$

$$s.t$$

$$\left(\mathrm{S}^{(n)}, \mathrm{S}^{(\infty)} \right) < 0$$

ζ is the vector of parameters, $h(y_t, \zeta)$ is the vector of moment conditions and $W_T(\bar{\zeta}_T(\zeta))$ is a symmetric and positive semi-definite weighting matrix. $W_T(\bar{\zeta}_T(\zeta))$ is obtained by computing the variance-covariance matrix of the moment conditions evaluated at $\bar{\zeta}_T(\zeta)$; when $\bar{\zeta}_T(\zeta) = \zeta$, the Continuously Updated Estimator (CUE) is obtained. $\mathrm{S}^{(n)}$ and $\mathrm{S}^{(\infty)}$ are respectively the n-period maturity and the infinite maturity spread at the steady state ($\sigma_t^2 = \bar{\sigma}^2$ and $x_t = 0$)[18].

I estimate 15 parameters : 3 preference parameters and 12 cash flows parameters using 21 moment conditions. The parameters and the moment conditions are summarized in the table 1.5.1:

Notes: This table shows the parameters and the moment conditions involved in our GMM estimation. E(.) is for the mean, Var(.) is for the variance, ACV1(.) is for the first order auto-covariance, $E(.)^3$is for the third order central moment, CV(.,.) is for the covariance.

The moment conditions are form by the mean, the variance, the first order auto-covariance,

[18] In the application, I only put $\mathrm{S}^{(\infty)}$ as the constraint to account for the possibility that the average term-structure can be hump shaped

Table 1.1: Structural parameters and moment conditions for the GMM

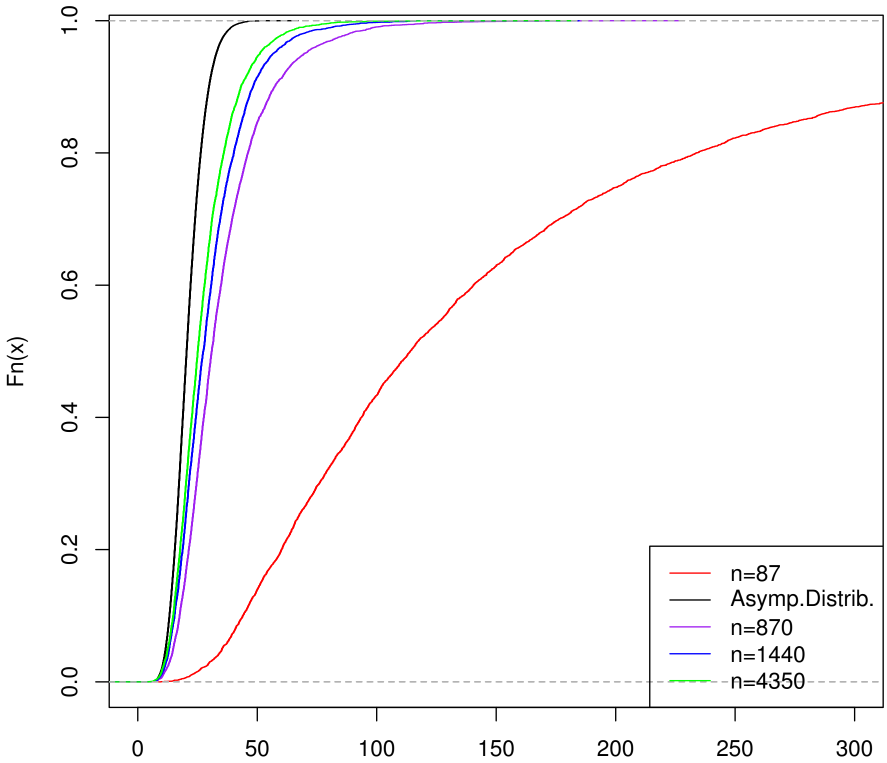
Model's structural parameters
$\underbrace{\delta, \gamma, \psi}_{\text{preference parameters}}, \underbrace{\mu_c, \mu_d, \phi, \varphi_d, \rho, \varphi_e, \bar{\sigma}, \nu, \sigma_w, \pi_c, \varphi_\sigma, \pi_\sigma}_{\text{cash flows parameters}}$
Moment conditions
$\underbrace{\mathrm{E}(g_t^a), \mathrm{Var}(g_t^a), \mathrm{ACV1}(g_t^a), \mathrm{E}(g_t^a - \mathrm{E}(g_t^a))^3}_{\text{consumption growth}}$
$\underbrace{\mathrm{E}(g_{d,t}^a), \mathrm{Var}(g_{d,t}^a), \mathrm{ACV1}(g_{d,t}^a), \mathrm{E}(g_{d,t}^a - \mathrm{E}(g_{d,t}^a),)^3, \mathrm{CV}(g_t^a, g_{d,t}^a)}_{\text{dividend growth}}$
$\underbrace{\mathrm{E}(pd_t^a), \mathrm{Var}(pd_t^a), \mathrm{ACV1}(pd_t^a)}_{\text{log-price dividend ratio}}$
$\underbrace{\mathrm{E}(r_{f,t}^a), \mathrm{Var}(r_{f,t}^a), \mathrm{ACV1}(r_{f,t}^a)}_{\text{risk free rate}}$
$\underbrace{\mathrm{E}(r_{m,t}^a), \mathrm{Var}(r_{m,t}^a), \mathrm{ACV1}(r_{m,t}^a)}_{\text{market return}}$
$\underbrace{\mathrm{CV}(r_{t+1}^{ex}, pd_t^a), \mathrm{CV}(r_{t+1,t+3}^{ex}, pd_t^a), \mathrm{CV}(r_{t+1,t+5}^{ex}, pd_t^a)}_{\text{Excess return predictability}}$

the third order central moment of consumption and dividend growth, the covariance between consumption growth and dividend growth; the mean, the variance and the first order auto-covariance of the market return, the risk free rate and the price dividend ratio; the covariance between the price dividend ratio and the year, three years and five years ahead excess return to account for the well recognized predictability of excess returns by the price dividend ratio. The identification of the parameters come from the fact that they appear in the theoretical moments for which we have the analytical formulas and thus despite the non-linearity of the moment conditions, we know exactly which moment conditions help to pin down each parameter.

Before running the estimation with the observed data, I first verify the validity of my approach by checking the convergence in distribution of the GMM objective function to the chi-squared distribution with the number of moment conditions (21) as the degree of freedom. As we can see from figure 1.2 the GMM objective function converges to the asymptotic distribution, showing that indeed, the finite sample moments converge to the theoretical moments I derived from the model. We also see that using the asymptotic distribution to test the validity of the model will lead to an over-rejection in finite sample.

Notes: This figure represents the cumulative distribution function (cdf) of the objective function in the GMM estimation. The cdf are computed by simulating 5000 data samples of length (in years) $n = 87$,$n = 870$,$n = 1440$ and $n = 4350$. We see that there is a convergence to the asymptotic distribution as the sample size increases.

Figure 1.2: Convergence of the objective function to the asymptotic distribution



1.5.2 Results

Table 1.6 and 1.7 present the results of the three groups of estimations that I have run : (i) Extended Model taking into account the possible correlations between volatility and cash flows shocks; (ii) Standard BY model without the correlations between volatility and cash flows shocks; (iii) Extended Model with EIS fixed to 1 (as develop along the paper) taking into account the possible correlations between volatility and cash flows shocks. In each group, I run the estimations with and without the term structure constraint to gauge how imposing the constraint changes the estimates. In all the case, the model is rejected if we consider the asymptotic distribution for the test; but using the simulated finite sample distribution to test the model, it is not rejected in all the case.

The estimated value of the pure discount rate is very closed to 1 in every cases with a half-life time higher than 50 years. The risk aversion estimate is higher with the constraint compared to the no-constraint estimate and the EIS estimate is greater than 1. The mean of the consumption growth is estimated around an annualized value of 2.4 %; it is higher under the constraint estimation than under the no-constraint estimation. The mean of the dividend growth is also positive but more variable across the different estimations. The loading of the dividend growth on the expected consumption growth is estimated to be greater than 1 in all the constrained cases. Thus the dividend load more on the long-run consumption risk compared to the consumption. The expected consumption growth is less persistent under the constrained case compared to the non-constrained estimate and the persistence of the volatility is higher in the constrained estimates compare to the non-constrained one. The constrained estimates emphasize the role of the volatility and reduce the effect of the long-run component of the consumption growth. The loading of the dividend on the short-run consumption shock is smaller under the constrained estimation compared to the non-constrained one. The loading of the dividend on the consumption growth volatility is negative in the constrained case and positive or closed to zero in the non-constrained case. These negative exposure of consumption and dividend growth to the volatility in the constrained case compensate for the lower exposure of dividend growth to the short-run consumption growth shocks and enables to match the observed equity risk premium.

The constrained estimation performed poorly on the predictability moment conditions. We can see from table 1.8 that the standardized errors are significantly different from zero. It also failed to capture the variance and the auto-correlation of the price-dividend ratio, and the variance of the market return. In the non-constrained estimation, the means of the price-dividend ratio, the mean of the risk free rate and the variance of the consumption growth are also not well matched.

The inference about the estimated parameters is not very precise since the 95 % confidence intervals for most of the estimated parameters obtained using the delta method are very wide. This is due to non-linearity of the moment conditions and the weak identification of some parameters in the model. The 95 % weak identification robust confidence intervals are also provided for the estimates in the extended LRR model. The projection method uses the Anderson-Rubin statistic, but it is computationally more demanding given the number of parameters in the model.

In summary, the parameter's estimates under the constraint are not very far from the one without the constraint, except for the persistence of the expected consumption growth, the persistence of the volatility and the loading of the dividend on the short-run consumption risk. The constrained estimates give more importance to the volatility channel and less importance to the long-run consumption growth channel to match the observed cash flows and asset prices data.

1.6 Simulations and discussion

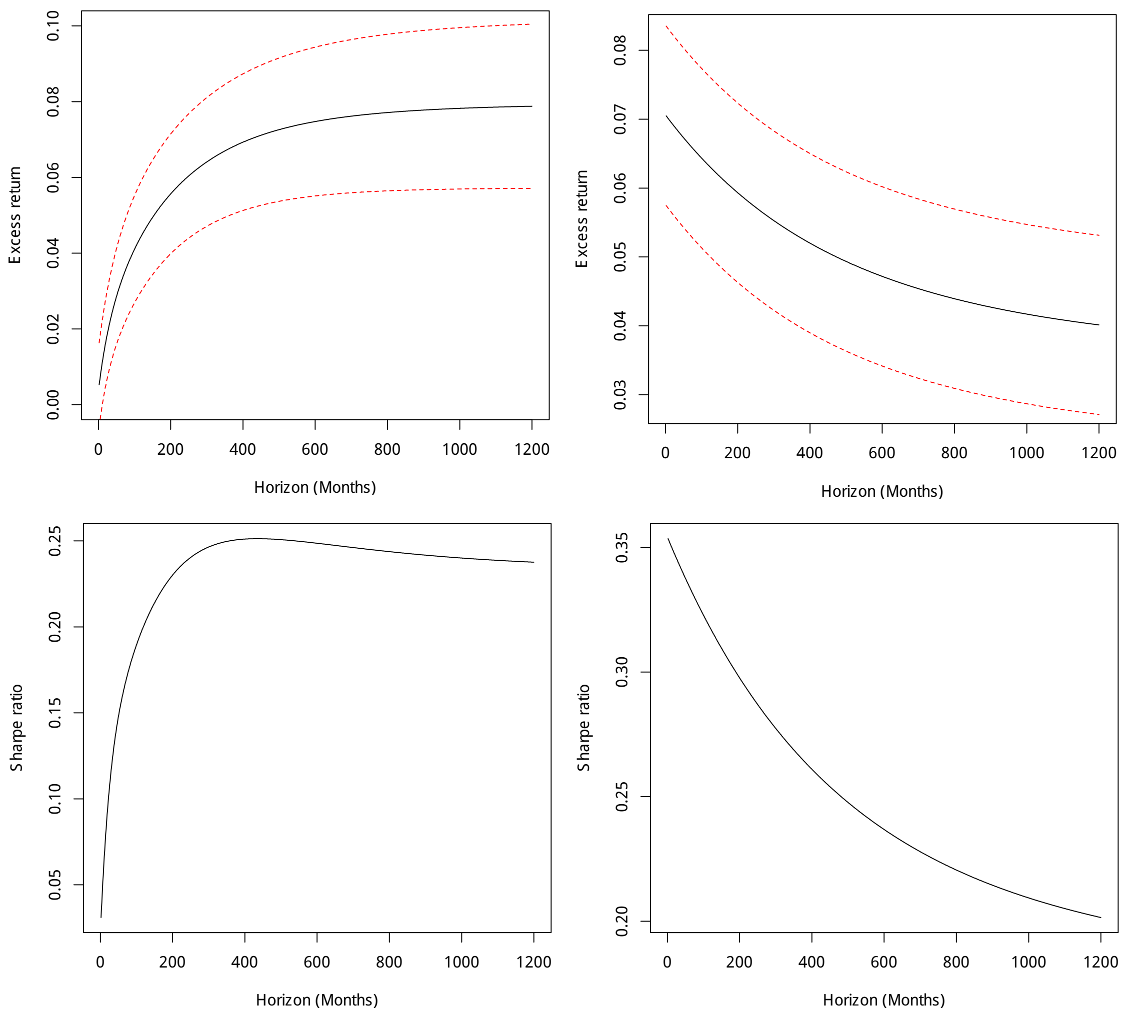
In this section, I use the estimated parameters with and without the term structure constraint, to simulate the data from the model and I compare the properties of the simulated data to the ones of the observed data. I first analyze the term-structures of the risk premium and Sharpe ratio. Then I move to the term-structure of the cash flows volatility and skewness.

1.6.1 The term-structure of the risk premium and Sharpe ratio

Figure 1.3 presents the simulated term structure of the equity risk premium (top panel) and the Sharpe ratio (bottom panel) implied by the model using the estimated parameters of the extended model without constraint (left panel) and extended model with the negative spread constraint (right panel). The blue dots represent the annualized mean and Sharpe ratio of the 6-month holding-period returns on the portfolios of dividend strips with maturities up to 12 months, 18 months and 24 months[19]. Similar to the standard LRR model, the risk premium term structure obtained for the estimates of the model without constraint is upward sloping. The same happens for the Sharpe ratio. On the contrary, in the case of the constrained estimates, both the risk premium and the Sharpe ratio have downward sloping term structures, implying that assets which pay-off in the distant future demand a lower return per unit of risk compared to assets which pay-off in the near future.

[19]The means and Sharpe ratio are computed from the data provided by van Binsbergen et al. [2012].

Figure 1.3: Implied term-structure of the risk premium and the Sharpe ratio

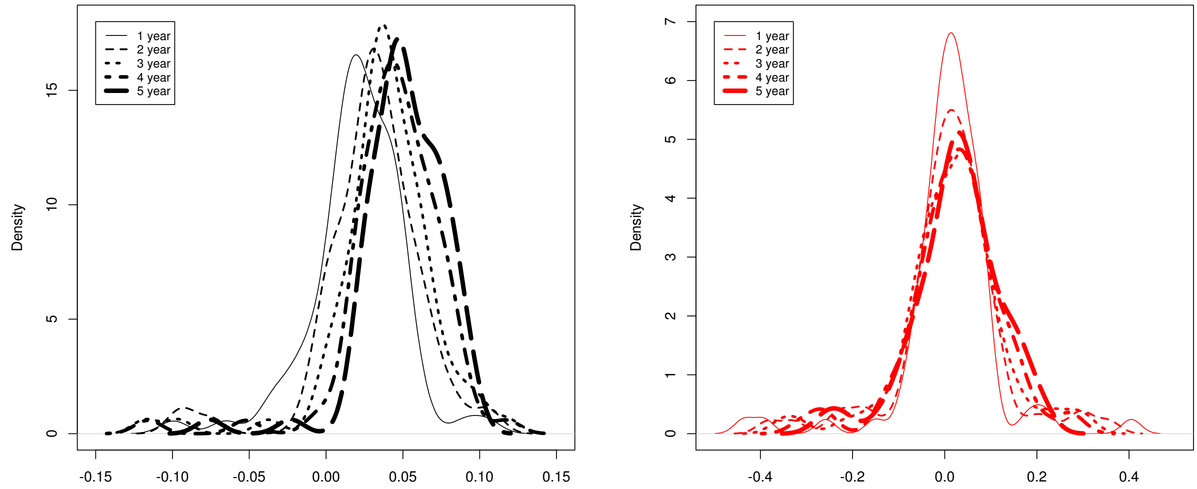


Notes: This graph shows the term structure of the equity mean excess return (upper panel) and the Sharpe ratio (bottom panel) implied by the model using the estimated parameters (Extended Model - NoConstr. for the left panel and Extended Model - Constr. for the right panel) for calibration. The dotted red lines on the right panel represent the 95 % confidence intervals. The left panel has been obtained using the estimates without the term structure constraint (Extended Model - NoConstr.) while the right panel uses the estimates with the term structure constraint (Extended Model - Constr.).

1.6.2 The timing of risk in the cash flow processes

The annual U.S data represented in figure 1.4 show that the distributions of consumption and dividends growths are more negatively skewed in the short-run than in the long-run. The shift of the distribution to the right as the time of aggregation increases tells us that the cash flows (consumption and dividends) are more risky in the short-run than in the long-run. Indeed, in the short-run the probability to observe a negative cash flow growth is higher than in the long-run and this can be the explanation for why investor will ask a higher risk premium for holding short-term assets than for long-term assets.

Figure 1.4: Distributions of rolling sum of consumption and dividends growths



This figure represents the distribution of the scaled rolling sum of consumption growth (left) and dividends growth (right). The sum is made over 1 year to 5 years. There is a shift of the distribution to the right as the aggregation period increases.

This empirical observation can be capture by the cash flow specification proposed in this paper. Indeed, the negative correlation between volatility and the cash flows will generate an asymmetry between good and bad times. During bad times, their is a negative shock on the cash flows and an increase of uncertainty; the higher volatility will thus amplify the negative shock on the cash flows. But during good time, there is a positive shock on the cash flows but the volatility is also low and thus, the positive cash flow shocks will be dampened. To see the evolution of the skewness of aggregate cash flow with the aggregation period in the proposed cash flow dynamics, let us define the time t - skewness over horizon τ for the cash flow realisations $(D_\tau)_{\tau=1\cdots T}$ as follows .

$$Skew_t(\tau) = \frac{\mathrm{E}_t\left[\left(\log\left(\frac{D_{t+\tau}}{D_t}\right) - \mathrm{E}_t \log\left(\frac{D_{t+\tau}}{D_t}\right)\right)^3\right]}{\mathrm{E}_t\left[\left(\log\left(\frac{D_{t+\tau}}{D_t}\right) - \mathrm{E}_t \log\left(\frac{D_{t+\tau}}{D_t}\right)\right)^2\right]^{\frac{3}{2}}} \tag{1.6.1}$$

Under the dynamics assumed in equations 1.3.2-1.3.5, assuming for simplicity that $\varphi_e = 0$, the time t - skewness over horizon τ of the dividend becomes:

$$Skew_t(\tau) = \frac{3\pi_\sigma\left[\left(\pi_c^2 + \varphi_d^2\right)\left(\tau - \frac{1-\nu^\tau}{1-\nu}\right)\frac{\sigma_w^2}{1-\nu}\right]}{\left[\tau\left(\left(\pi_c^2 + \varphi_d^2\right)\sigma_t^2 + \pi_\sigma^2\sigma_w^2\right)\right]^{\frac{3}{2}}} \tag{1.6.2}$$

We see that the skewness will be zero if there is no stochastic volatility ($\sigma_w = 0$) or no leverage effect ($\pi_\sigma = 0$). The sign of the skewness depends on the correlation between the innovations in the level of the cash flow and in its volatility[20]. The one period skewness converges to zero as the horizon increases.

This result on the term structure of skewness complements the one obtained by Gollier [2017] regarding the implications of stochastic volatility in the LRR model. Indeed, he showed that adding uncertainty to the variance of the consumption growth as in the standard BY model and making the shocks on the variance more persistent increases the average risk premium through an increasing term structure of the annualized kurtosis. Thus implying that the term structures of interest rates and risk premia will remain respectively decreasing and increasing. In our extended LRR model, the negative and increasing term structure of the skewness increases the average risk premium by magnifying the short-term risks and implies a decreasing term structure of the risk premium

Let us now look at the term structure of the cash flow volatility. The definition of the variance ratio[21] (VR) statistics used by Belo et al. [2015], Marfe [2016] is given by :

$$\mathrm{VR}_t(\tau) = \frac{\sigma_t^2(\tau)}{\sigma_t^2(1)} \tag{1.6.3}$$

Where

$$\sigma_t^2(\tau) = \frac{1}{\tau}\left[\log \mathrm{E}_t\left[D^2(\tau)/D^2(0)\right] - 2\log \mathrm{E}_t\left[D(\tau)/D(0)\right]\right] \tag{1.6.4}$$

The variance ratio allows to gauge the timing of the risk building (increase or decrease of the volatility) in the process by comparing this statistics to the value obtained in the benchmark case (which is the i.i.d and homoskedastic cash flow process with VR=1). When VR is below 1, it tells us that the bulk of the risk appears in the short term, while a VR above 1 shows that risk increases with the horizon and concentrates in the long-term.

Using the dynamics of the dividend growth process, we can derive an analytical formula of the dividend variance ratio as follows:

$$\mathrm{VR}_t(\tau) = \frac{1}{\tau}\frac{\left[H_{\tau-1}\sigma_t^2 + (1-\nu)\left(\sum_{i=0}^{\tau-2} H_i\right)\bar{\sigma}^2 + \left(2G_{\tau-1} - G'_{\tau-1}\right)\sigma_w^2\right]}{H_0\sigma_t^2 + \left(2G_0 - G'_0\right)\sigma_w^2} \tag{1.6.5}$$

[20]Notice that the numerator in the expression of the skewness in 1.6.2 corresponds to the term that drives the decreasing term structure of the variance ratio.

[21]The standard definition of the variance ratio is given by equation 1.10.22. When the dividend process is log normal with independent and identically distributed innovations, the two definitions coincide.

Where $\forall j \in \{1, 2, ...\}$

$$H_j = \nu H_{j-1} + \varphi_d^2 + \left[\phi\left(\frac{1-\rho^j}{1-\rho}\right)\varphi_e\right]^2 + \pi_c^2$$
$$G_j = G_{j-1} + \left[\pi_\sigma + H_{j-1}\right]^2$$
$$G_j^{'} = G_{j-1}^{'} + \left[\pi_\sigma + \frac{1}{2}H_{j-1}\right]^2$$

and $H_0 = \varphi_d^2 + \pi_c^2$, $G_0 = G_0^{'} = \pi_\sigma^2$.

To see the effect of the persistent component of the consumption growth process and thus the effect of serial correlation, we first shut down the stochastic volatility ($\nu = \sigma_w = 0$ and $\pi_\sigma = 0$). Then the VR statistic in 1.6.5 becomes :

$$\mathrm{VR}_t(\tau) = \frac{1}{\tau}\frac{\left[\left(\sum_{i=0}^{\tau-1} H_i\right)\right]}{H_0} = \frac{1}{\tau}\frac{\sum_{i=0}^{\tau-1}\left(\varphi_d^2 + \left[\phi\left(\frac{1-\rho^j}{1-\rho}\right)\varphi_e\right]^2 + \pi_c^2\right)}{\varphi_d^2 + \pi_c^2} \quad (1.6.6)$$

We can easily see that when $\varphi_e > 0$, the elements in the summation in the numerator of 1.6.6 start with the same value as the denominator and then increase with the horizon. The higher the persistence of the expected consumption growth (ρ), the higher the increments. So the serial correlation in the expected consumption growth induces that the risk is shifted toward the future, yielding an increasing term structure of dividend risk. When there is no serial correlation, the VR statistic is constant and the term structure of dividend risk is flat. This can be seen easily by setting $\varphi_e = 0$ or $\rho = 0$.

We now turn to analyze the effect of stochastic volatility by shutting down the persistent component of consumption growth ($\rho = \varphi_e = 0$)[22]. Thus the VR statistic in 1.6.5 becomes :

$$\mathrm{VR}_t(\tau) = \frac{1}{\tau}\frac{\left[H_{\tau-1}\sigma_t^2 + (1-\nu)\left(\sum_{i=0}^{\tau-2} H_i\right)\bar{\sigma}^2 + \left(\pi_\sigma^2 + \sum_{i=0}^{\tau-2}\left(\pi_\sigma^2 + 3\pi_\sigma H_i + \frac{7}{4}H_i^2\right)\right)\sigma_w^2\right]}{\left(\varphi_d^2 + \pi_c^2\right)\sigma_t^2 + \pi_\sigma^2\sigma_w^2} \quad (1.6.7)$$

where $\forall j \in \{1, 2, ...\}$

$$H_j = \left(\frac{1-\nu^j}{1-\nu}\right)\left(\varphi_d^2 + \pi_c^2\right)$$

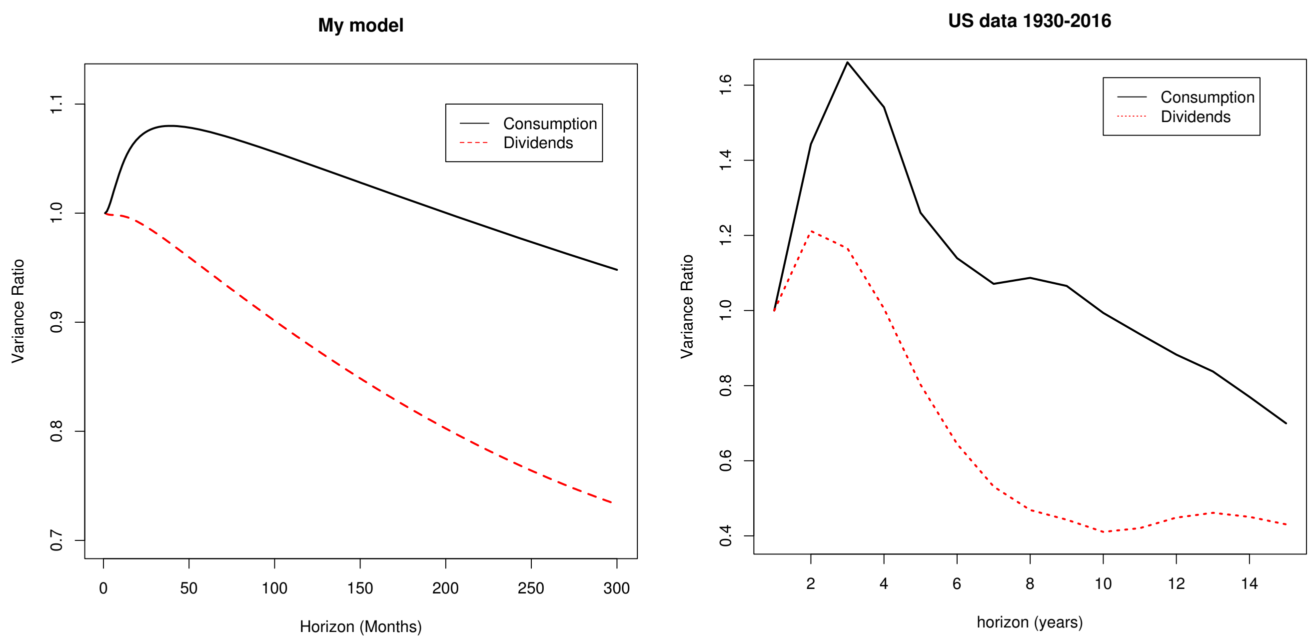
First notice that when there is no volatility persistence ($\nu = 0$), the part of the numerator that depends on σ_t^2 and $\bar{\sigma}^2$ is in expectation similar to the corresponding part in the denominator ($\left(\varphi_d^2 + \pi_c^2\right)\sigma_t^2$); but the part that depends on the volatility of volatility (σ_w) starts with the loading of the dividend growth on volatility shock (π_σ) as in the denominator and then builds on by adding the term $\left[\pi_\sigma^2 + 3\pi_\sigma\left(\varphi_d^2 + \pi_c^2\right) + \frac{7}{4}\left(\varphi_d^2 + \pi_c^2\right)^2\right]$. So, if the added term is negative then the VR will decrease as the horizon increases simply because for each supplementary period, the

[22]Notice that in this case, VR statistic is constant and the term structure of dividend risk is flat

added risk is below the one present in the first period. This happens for example when dividend reacts sufficiently negatively to positive volatility shocks, in this case when $\pi_\sigma < -\frac{7}{12}\left(\varphi_d^2 + \pi_c^2\right)$. When the persistence of the volatility is different from zero, the same reasoning leads to the conclusion that the variance ratio will be decreasing given that the loading of the dividend on the volatility (π_σ) is sufficiently negative. The more the persistence of the stochastic volatility process, the lower should π_σ be to obtain a decreasing VR statistics.

I use the parameters obtained in the estimation part to calibrate our model. I look at the implications for the cash flows dynamics by comparing the variance ratio implied by the model with the ones in the data. I used the parameter's estimates with the constraint to compute the variance ratios for consumption and dividend growth. As we can see from figure 1.5, the variance ratio of the consumption growth (left panel) is hump-shape in the model as in the data and it goes below 1 as the horizon increases. For the dividend growth, the variance ratio in the model also follows the same pattern as in the data and it falls more quickly than for consumption growth. This happens in the model because the loading of the dividend growth on volatility shock is more negative (higher in absolute value) than the loading of the consumption growth.

Figure 1.5: Cash flows variance ratios



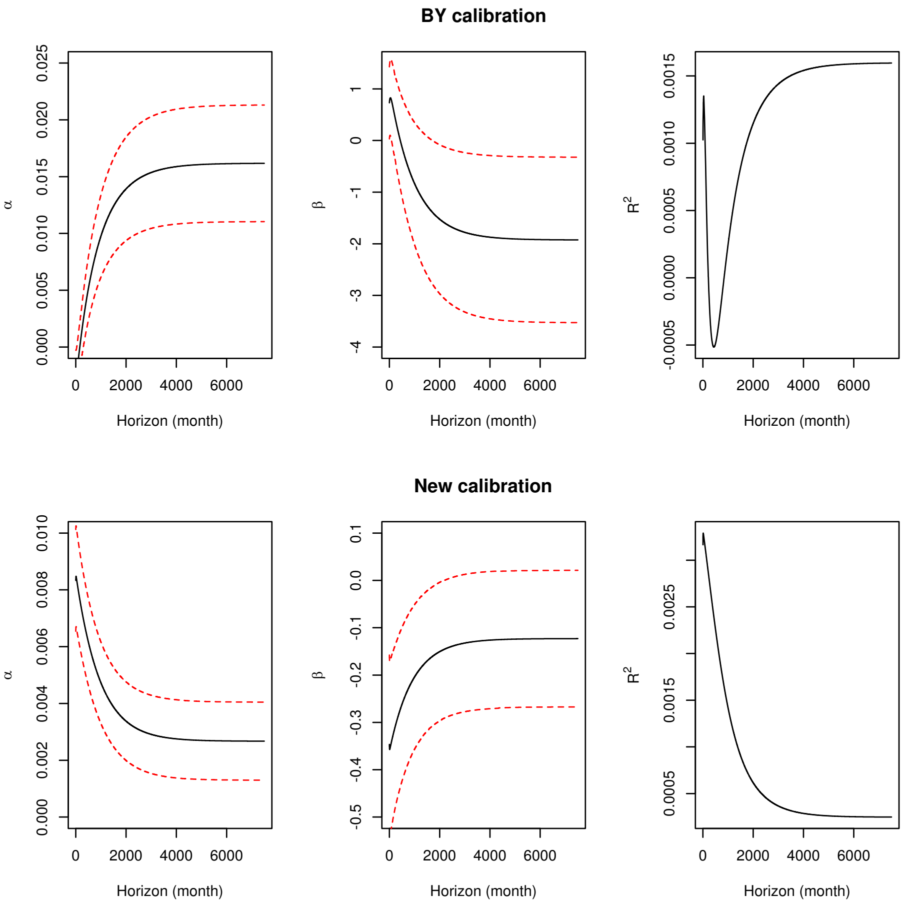
Notes: This figure shows the variance ratios of consumption and dividend growths in the model (left) and in the data (right).

1.6.3 Implications for the cross-section of returns

In the section 1.2 we have seen that returns on Value stocks load more negatively on the uncertainty measure (VIX) compared to returns on Growth stocks. We are now going to check if the same pattern happens in our model when we do a similar exercise with the simulated data. More specifically, we want to verify if the returns of short duration stocks load more negatively on the uncertainty measure compared to the returns of long-duration stocks. As we know already from figure 1.3 short-duration equity earns a higher risk premium compared to long duration equity and as we explained in the previous section, this pattern mainly comes from the increase with the maturity of the conditional variance loading in the log-price dividend ratio of dividend strips (see Fig. 1.10).

A regression of dividend strips excess returns on the consumption growth expected volatility shows different results for the standard LRR model and for our extended LRR model. Indeed, as we can see from Figure 1.6 in the case of the LRR model with the BKY calibration, long-duration dividend strips counter factually behaves like value stocks with a more negative loading on volatility than short-duration dividend strips. Thus if we believe Lettau and Wachter [2007, 2011] who associate the long-duration assets with growth stocks and the short-duration assets with value stocks, then the standard LRR model will imply a "growth premium". Contrary to that, in the same context, the extended LRR model implies the well known "value premium". Indeed as we can see from Figure 1.6, short-duration dividend strips have a higher unexplained excess return, they load more negatively on volatility and it has a higher explanatory power for them than for long-duration dividend strips. So as we expect, firms that weight more on short-duration dividend strips have a higher expected returns but also are more negatively exposed to uncertainty compared to firms that weight more on long-term dividend strips. In other terms, value stocks are more negatively exposed to volatility risk compared to growth stocks as observed in the data.

Figure 1.6: regression of dividend strip returns on Implied volatility in the LRR model and in the extended LRR model



Notes: This graph shows the intercept, the slope and the adjusted R-squared from the regression of excess return on dividends strips (at different maturities) on the implied volatility from the model using the BKY calibration (upper panel) and the estimated parameters of the Extended Model (bottom panel) . The regression equation is the following: $r_{d,t}^{(n)} - r_{f,t} = \alpha^{(n)} + \beta^{(n)} \left[E_t(\sigma_{t+1}^2)\right]^{\frac{1}{2}} + \epsilon_t$. The Top panel shows that in the standard LRR model, the alphas are increasing with the duration (more excess return left unexplained for long-duration assets than for short duration assets), the beta are decreasing (more negative exposure of long duration asset to volatility than short-duration asset). The bottom panel shows that in the extended LRR model the alphas are decreasing with the duration (more excess return left unexplained for short-duration assets than for long-duration assets), the beta are increasing (more negative exposure of short-duration assets to volatility than long-duration assets) and the adjusted R-squared are decreasing (more explanatory power of the volatility for short-duration assets than for long-duration assets).

1.7 What drives the slope of the term structure of equity returns?

1.7.1 Back to the simplest case: time separable utility function

The idea that the term structure of equity risk premium is downward slopping is counter-intuitive when we look at the standard time separable utility function. Indeed, with such preferences, the risk premium of a given dividend strip over n-period is proportional to the variance of the consumption growth over that period and that variance is expected to increase as the horizon increases, implying that the aggregate risk premium should be higher for a longer aggregation period. When the consumption growth is i.i.d, the term structure is flat and it is increasing with positively serially correlated consumption growth. To see why, I will apply the following restriction : $\gamma = \frac{1}{\psi}$ to the term spread equations 1.10.11 and 1.10.13. I will focus on three restrictions usually applied in the literature to the consumption growth process and see the implications for the equity risk premium and its term structure: (i) i.i.d normal consumption growth process; (ii) independently normally distributed consumption growth (with stochastic volatility); (iii) identically normally distributed consumption growth process (with persistent component). Finally, without loss of generality, I will restrict myself to the cases of assets whose loading on the expected consumption growth is given by $\phi = \gamma$.

In the first case (i), the risk premium is given by 1.7.1, which is the risk aversion coefficient times the variance of the consumption growth and the dividend loading on consumption growth shock. Thus, the term structure of equity risk premium is flat.

$$rp_t^{(n)} = \gamma \pi_c \bar{\sigma}^2 \tag{1.7.1}$$

In the second case, the consumption growth process incorporate a stochastic volatility that changes the conditional variance of consumption growth each period. So, their is an extra uncertainty that the investor perceives and which will affect the risk premium required to bear that risk. The risk premium on a n-period dividend strip (1.7.2) has an additional term that depends on the loading on the volatility shock in the consumption growth process and on the conditional variance of the volatility process. When the consumption growth does not react to shocks to the volatility ($\varphi_\sigma = 0$), the risk premium is time varying but has the same form as in the i.i.d normal consumption growth case. When $\varphi_\sigma \neq 0$, part of the risk premium that comes from the stochastic volatility depends on the loading of the dividend growth process on the volatility shock and also of the squares of the loadings on consumption growth and dividend growth shocks. But the spread between the risk premiums of two consecutive maturities dividend strips only depends on the risk aversion, on the loading of dividend growth on the consumption growth shocks and on the idiosyncratic dividend growth loading; it does not depend on the loading of the dividend growth on the volatility shock.

$$rp_t^{(n)} = \gamma\pi_c\sigma_t^2 + (\gamma\varphi_\sigma)\left(\pi_\sigma + \frac{1}{2}\left[(\pi_c-\gamma)^2+\varphi_d^2\right]\left(\frac{1-\nu^{n-1}}{1-\nu}\right)\right)\sigma_w^2 \tag{1.7.2}$$

The persistence of the volatility process emphasizes the contributions of both the idiosyncratic dividend growth shock and the loading of the dividend growth on consumption growth shock to the risk premium. The higher is ν, the bigger will be the contributions of consumption and dividend growth shocks to the level of the risk premium. The slope of the term structure also depends on the persistence of the volatility process; when there is no persistence in the volatility ($\nu = 0$), the term structure of the risk premium is flat. The sign of the term structure's slope only depends on the loading of consumption growth on volatility shock (φ_σ) given that all the other terms are positive $\left(\left[(\pi_c-\gamma)^2+\varphi_d^2\right]\right)$. When $\varphi_\sigma < 0$, which implies that the price of volatility risk is negative, the equity risk premium has a downward sloping term structure.

Finally, let us now assume that the consumption growth has a persistent component, thus introducing serial correlation in the process but a constant volatility. Then the risk premium on a n-period dividend strip is given by (1.7.1) and it has a flat term structure. This happens here because the expected consumption growth risk is not priced ($\lambda_x = 0$), thus the correlation in the expected consumption growth does not affect the risk premium term structure.

In summary, we see that in a Consumption based Capital Asset Pricing Model (CCAPM) with CRRA utility function, the term structure of equity risk premium could change depending of the sign of the correlation between the consumption growth and the volatility process. When the stochastic volatility process and the consumption growth process are independent, the term structure of the risk premium is flat even when there is some serial correlation introduced in the consumption growth process through its expected component. The term structure is increasing when the innovation in the volatility is positively correlated with innovations in consumption growth and it is decreasing when consumption growth and volatility evolve counter-cyclically as we observe in the data. So, in a discounted expected utility model, the sign of the correlation between consumption growth and the stochastic volatility determines the sign of the slope of the equity risk premium term structure.

1.7.2 Compatibility with the implied dividend strip data

From section 1.4.1, we show that a necessary condition for the downward sloping term structure of equity returns is that the loading of the log price-dividend ratio on the volatility should be positive and increasing. So, on average the log price-dividend ratio should be an increasing function of the volatility. But, is this feature compatible with the data ? To assess the compatibility between the model implications and the data, we regress the monthly log price-dividend ratio on various proxies of the time varying volatility of consumption growth representing the macro-uncertainty in our model. The proxies that we used are the VIX index, the macro-uncertainty measures computed by Jurado et al. [2015] and the realized volatility on the dividends growth. We use the data on the S&P 500 implied dividend swap prices from January 2005 to October 2016[23] and long term government bond yield data from the Federal Reserve Bank of St-Louis. The maturities of the dividend futures range from 1 years to 19 years. We computed the spot price of the dividend strips by discounting the dividend futures prices with the corresponding bond yield. Finally, we regress the log price-dividend ratios of the dividend strips and the market index on the different proxies of the volatility. The results[24] in Table 1.10 show that macro-uncertainty explains the log price-dividend ratios with a negative slope. So contrary to our model prediction, a positive shock on the volatility reduces the log price-dividend ratios of dividend futures and market index. This negative result is counterbalanced by a positive result concerning the correlation between the expected return on the market portfolio and its conditional variance : $\text{cov}\left(\text{E}_t r_{m,t+1}, \text{Var}_t r_{m,t+1}\right) = \Gamma_{2m}\left(\beta_{m,x}^2 + \beta_{m,c}^2 + \beta_{m,d}^2\right)\text{Var}(\sigma_t^2)\ \Gamma_{2m} = \left(k_{1m}\nu - 1\right)A_{2m} < 0$. Indeed, this correlation has been documented to be negative in the data (Glosten et al. [1993], Whitelaw [1994]) and the standard LRR model is not able to capture it (Bansal and Yaron [2004]). Notice that this correlation has the same formulation as in the standard LRR model, but what differs is the sign of the loading of the price dividend ratio on volatility. There is a tight link between the average term structure slope, the correlation of price dividend ratio and the volatility, and the correlation between the expected return and the conditional variance of the market return. All those quantities depend on one coefficient (the loading of the price dividend ratio on volatility :A_{2m}) in the LRR model but they have been documented to have different signs in the data. Using the dividend strip data on the S&P 500, Bansal et al. [2017] found that the slope of the dividend yield curve is time varying; positive during normal time, negative

[23] We thank Christian Mueller-Glissman for providing us with the data.

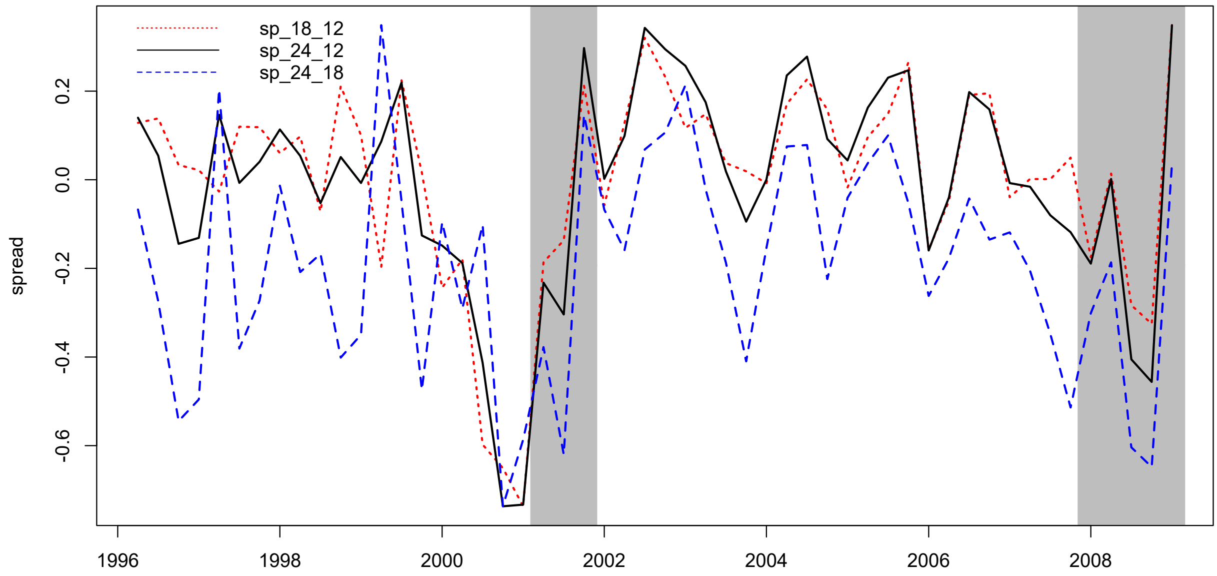
[24] For the sake of brievity, we only report the results for the macro uncertainty measures of Jurado et al. [2015]. But the results are qualitatively the same for the other measures of uncertainty we mentionned above.

during recessions and on average positive. Indeed, we can see from figure **??** that the downward slope of equity returns especially happened during the 2001 and 2008 recessions. Their result suggests that there should be two regimes with a time varying transition probability in the spirit of Whitelaw [2000] to model the term structure of equity returns. This could also help to break the tight link between the three quantities mentionned previously and it is left for future research.

1.7.3 Time variation in the slope of the term-structure

The term-structure of the equity premium is not just upward or downward sloping; it is time varying and the sign of the slope at a given time depends on the economic conditions. Bansal et al. [2017] have recently found that the slope of the equity term-structure is strongly positive during normal time and negative during recessions. On the contrary, Gormsen [2017] argued that slope of the equity term-structure is counter-cyclical; it is downward sloping in good times, but upward sloping in bad times. Using the data on dividend strip return provided by van Binsbergen et al. [2012], I computed the spread between the annualized 6 months holding period returns on dividend strips with 12 months, 18 months and 24 months of maturity. Figure 1.7 represent the time series of those spread. We can see that there is no clear cut on the sign of the spread during normal times versus recessions. On one hand, we can see that during the 2008 recession, the spread was negative, meaning that the term-structure was downward sloping but there are also many normal times were this happen too. On the other hand, we can see that during the 2001 recession the spread was positive for some time; meaning that the term-structure was upward sloping and this also happens during normal times.

Figure 1.7: Time variation in the term-structure slope



Notes: This figure shows the spread (difference in annualized returns) between : the dividend strips with 18 months to maturity and 12 months to maturity (sp_18_12), the dividend strips with 24 months to maturity and 12 months to maturity (sp_24_12) and the dividend strips with 24 months to maturity and 18 months to maturity (sp_24_18). The gray bars represent the NBER recessions.

In our model, the time variation of the term-structure slope is driven by the relative weights between the price of the volatility risk and the price of the expected consumption growth risk. We choose the uncertainty parameter (the volatility of the volatility, σ_w) to be constant, but this can be made time varying[25]. So the time variation of the term-structure in our model is similar to the one present in the standard LRR model since it isc driven by the same weighted price of the long run consumption risk; the difference being that the term structure slope can be positive or negative during certain times in our model while it is always positive in the standard LRR model. With a time varying uncertainty, the time variation in the term-structure slope could also come from the price of volatility risk driven by the time variation in the consumption growth uncertainty (volatility of the volatility) or from the price of the expected consumption growth through the time variation in the consumption growth volatility. During bad times, the volatility of the consumption growth is high (which gives more weight to positive part of the spread), but the raise in the level of the consumption growth volatility could also be accompanied by a raise in the uncertainty on the consumption growth (an increase of the volatility of the consumption growth volatility) which will decrease the negative part of the spread. At the end, the slope of the term-structure could be negative or positive depending on the dominant effect

[25]See for example Bollerslev et al. [2009], Tauchen [2011].

(level of the volatility vs variability of the volatility).

1.8 Other implications of the model extension

In this section we will show the implications of the model extension for the price of human capital, the cross-section of assets and the pricing of long-term investment project. In summary, our extended model which allows to capture the decreasing term-structure of the risk premium, implies that the price of the human capital negatively co-moves with consumption growth volatility and; a lower discount rate should be used to value long-term investment projects.

1.8.1 Implications for the price of human capital

One of the main implications of our model specification, while using the estimated parameters under constraint for calibration, is that the log of the price of the wealth portfolio to consumption ratio co-moves negatively with the conditional variance of the consumption growth while the log of the price of the market portfolio to dividend ratio co-moves positively with the same conditional variance. This is also one of the key differences in term of asset pricing implications between our model specification and the standard LRR model. Indeed, with our model specification and calibration, as we can see from Figure 1.10, the conditional variance loading in the log-price dividend ratio of dividend strips is positive and increases with the dividend strip's maturity, meaning that risky assets which pay-off far in the future positively react to an increase in the conditional variance; their prices increase as they become more safer compared to risky assets which pay-off in the near future. This declining pattern of the impact of cash flows uncertainty on the risk premium is in accordance with the view that while uncertainty might reduce short-run consumption, hiring or investment, it might also encourage research and innovation, thus improving economic prospect in the long-run and lowering the long-run risk premium (Bloom [2014]). The positivity of the conditional variance loadings of the log-price dividend ratio for dividend strip translates into a positive loading on the conditional variance of the log price dividend ratio for the aggregate market portfolio. In the standard LRR model the contrary effect happens; the dividend strips log-price dividend ratio loadings of the conditional variance are negative and decreases with the maturity and because of the volatility build-up, risky assets that pay-off far in the future are more exposed to the expected volatility risk, their prices drop more during bad time compared to risky assets which pay-off in the near future.

Considering that the wealth portfolio is made by human capital and other financial assets constituting the market portfolio, the price of the wealth portfolio is the sum of the human capital, which can be obtained as the present value of the stream of future labor incomes, and the price of the market portfolio[26]. The negative co-movement of the price of the wealth portfolio with the consumption growth volatility combines with the positive co-movement of the price of the market portfolio with the consumption growth volatility lead to two important predictions: Firstly, the price of the human capital will negatively co-move with consumption growth volatility and secondly, the human capital drives the wealth portfolio exposure to consumption growth volatility risk. The first prediction is confirmed by the observation that during high economic uncertainty periods, unemployment increases[27] and labor income drops, thus the price of human capital decreases. The second prediction is corroborated by the fact that labor income constitutes around 75 % of the consumption Santos and Veronesi [2006].

1.8.2 Economic policy implication: The pricing of long term projects

The pricing of long term projects (e.g: Investment in the fight against climate change or desertification) for which the expected cash flows might happen or continue to fall very far in the future, requires to know the term structure of the risk premium. Indeed, a slight modification of the discount rate used to compute the present value of the future expected benefits will have a huge impact on the outcome and might have different policy implications in a benefits and costs analysis. For illustration, suppose you are asked the following question: "How much are you willing to pay for an "average market" risky investment that is expected to pay-off 742 billions $ (1 % of the world GDP in 2015) in 100 years ?". By "average market" risky investment, we mean an investment which mimic the market index in terms of cash flows and returns. The answer will depend on the discount rate that will be used to compute the present value of the expected cash flow. Table 1.8.2 summarizes the results of this computation. If the CAPM model is applied, the discount rate that will be used is the historical average of the return on the S&P 500 which is around 7.5 % and the present value will be around 536 millions $. Instead, if the LRR model is used with the BKY calibration, the discount rate that will be used is 11 % resulting in a present

[26] The price of the wealth portfolio denoted by W_t can be expressed as : $W_t = \lim_{T\to\infty} \mathrm{E}_t\left(\sum_{i=1}^{T} M_{t+i}C_{t+i}\right) = \lim_{T\to\infty} \mathrm{E}_t\left(\sum_{i=1}^{T} M_{t+i}\left[D_{t+i} + L_{t+i}\right]\right) = P_t^M + P_t^H$ where P_t^M and P_t^H stand respectively for the price of market portfolio (including real estate and financial assets) and the price of human capital.

[27] In a BVAR model with macro uncertainty, unemployment, inflation and interest rate, Leduc and Liu [2016] found that a one standard-deviation positive shock on uncertainty (measured by the VIX) acts like a negative aggregate demand shock by increasing unemployment for about 2 years, by decreasing inflation for about 15 months and by decreasing the interest rate.

value of only 22 millions \$. Finally, if the extended LRR model is applied, a discount rate of 4 % will be used and the present value of 14 billions \$ will be obtained. Thus, the last computation will give the highest chance to the project to be implemented once the cost is evaluated.

Table 1.2: Present value by discount rate

Model	discount rate	present value	Term structure
CAPM	7.5 %	536 millions \$	flat
LRR	11 %	22 millions \$	increasing
ELRR	4 %	14 billions \$	decreasing

Notes: This table shows the present value of an expected cash flow of 742 billions \$ (1 % of the world GDP in 2015) that will occur in 100 years. The present value is computed assuming the equilibrium discount rate under different models (CAPM, the long-run Risk model and Extended LRR model we propose in this paper).

Our extension of the LRR model would be of particular importance for policy makers because it has more flexibility regarding the implied term structure of risk premium. Indeed, as the recent findings concerning the slope of the term structure are still considered by some researchers as fragile[28] it is useful to have a flexible model in hand that could be able to cope with any situation.

1.9 Conclusion

Recent empirical works in asset pricing have shown that the term structure of the risk premium, the term structure of the Sharpe ratio and the timing of risk in the cash flows are all downward sloping, meaning that risky assets which pay-off in the near future earn a higher expected return compared to risky assets which pay-off in the distant future. Reproducing these observations have been challenging for leading asset pricing models which on the contrary predict an increasing term structure of the risk premium. In this paper, I show that allowing for the negative correlation between cash flows and consumption growth volatility representing the macroeconomic uncertainty could enable to reverse the term structure of the risk premium in leading asset pricing models. The mechanism at play being that allowing for this negative correlation in the dynamics of cash flows enables to shift the risk structure toward the near future as it can be seen from the variance ratio statistic. Furthermore, the exposure of consumption growth to uncertainty shock adds another source of risk that is priced in the short-run. Risky assets

[28] The liquidity of the options used by van Binsbergen et al. [2012] seems to dry up as the maturity increases, implying more uncertainty on the price for dividend strips with longer maturity (see Cochrane [2017] for further critics on the statistical significance of their results). Furthermore, t-statistics of results obtained using the proprietary data on dividend swaps for the US in van Binsbergen and Koijen [2016] do not permit to reject the null hypothesis of a significant difference between the expected returns on short-term assets an on long-term assets.

which are exposed to the same (macro-uncertainty) shock earn a higher return in the short-run, therefore risky asset which pay-off in the distant future appear to be safer compared to asset which pay-off in the near future and are more exposed to the short-run volatility risk. The price of the former is then higher compared to the price of the a similar asset paying off in the near future. Hence a declining term structure of the risk premium can be achieved. I also obtain two testable predictions from our model : firstly, the price of the human capital negatively co-moves with consumption growth volatility and secondly, the human capital drives the wealth portfolio exposure to consumption growth volatility risk.

1.10 Appendix

1.10.1 Data description

For the estimation of the structural parameters of the model, we used 5 variables : real consumption growth rate, real dividend growth rate, log price dividend ratio, real market return and real risk free rate. The data on consumption are obtained from the Bureau of Economic Analysis (NIPA table 2.4.6) by summing real personal expenditures on non-durable goods and services. We used population data from FRED of the Federal Reserve Bank of St. Louis to compute the individual consumption expenditure. Then we used the price index from NIPA table 2.3.4 to compute the real consumption expenditures and we took the difference of log consumptions to obtain the real consumption growth rate. For the market return, we used the CRSP data on the S&P500 index value weighted return. The dividends were computed using the data on the level of the S&P500 index, the data on value weighted returns including dividends and the data on value weighted returns excluding dividends. The risk free rate is the one month T Bill rate obtained from Kenneth French data library. All the data spanned from 1930 to 2016 for the annual frequency and from 1947Q2 to 2016Q4 for the quarterly frequency.

Table 1.3: Summary statistics

	Consumption growth	Dividend Growth	Log(P/D)	Market Return	Risk-Free Rate
Mean	0.0183	0.0171	3.4107	0.0801	0.0064
St.dev.	0.0213	0.1131	0.4536	0.1927	0.0379
Skewness	-1.4601	-0.7208	0.2191	-0.5401	-0.0438
Kurtosis	4.9243	5.4932	-0.4654	0.7048	2.4715
Min.	-0.0803	-0.4266	2.3592	-0.4756	-0.1147
Max.	0.0731	0.4447	4.4429	0.6199	0.1352
nobs	87	87	87	87	87

This table shows the descriptive statistics for the consumption growth, dividend growth, log price dividend ratio, market return and the risk free rate. The database is at the annual frequency and covers the period from 1930 to 2016.

Table 1.4: Summary statistics

	Consumption growth	Dividend Growth	Log(P/D)	Market Return	Risk-Free Rate
Mean	0.004	0.0077	4.8756	0.0211	0.0024
St.dev.	0.0057	0.1402	0.4263	0.0805	0.0066
Skewness	-0.4036	0.0733	0.1577	-0.7552	-0.4427
Kurtosis	1.535	3.219	-0.3031	1.3022	2.1897
Min.	-0.0169	-0.5724	3.7258	-0.2950	-0.0279
Max.	0.0254	0.5450	5.9787	0.2115	0.02198
nobs	279	279	279	279	279

This table shows the descriptive statistics for the consumption growth, dividend growth, log price dividend ratio, market return and the risk free rate. The database is at the quarterly frequency and covers the period from 1947Q2 to 2016Q12.

Table 1.5: Estimation of the VAR model with logVIX and consumption growth

	Δc_t	$\log VIX_t$
Δc_{t-1}	0.152 (0.101)	-8.517 * (4.852)
$\log VIX_{t-1}$	-0.004 * (0.002)	0.775 *** (0.104)
Δc_{t-2}	0.253 ** (0.098)	2.933 (4.71)
$\log VIX_{t-2}$	0.005 * (0.003)	-0.072 (0.131)
Δc_{t-3}	0.326 *** (0.101)	5.851 (4.862)
$\log VIX_{t-3}$	-0.001 (0.002)	0.148 (0.103)
const.	0.001 (0.004)	0.433 * (0.211)
R^2	0.31	0.68

Notes: This table shows the results of the estimation of the VAR model with logVIX and consumption growth. Std. Errors of the estimates are given in brackets. The number of lag has been selected using the information criteria (AIC, HQ,FPE). The estimated correlation between the residual is -0.305 with [-0.5698, -0.0724] as 95% confidence interval obtained by block bootstrap.

Table 1.6: Estimations results

Par.	Constr.	Par.	Constr.	Par.	NoConstr.	Par.	NoConstr.
δ	0.9995	φ_e	0.190	δ	0.9993	φ_e	5.614e-02
[1]	[0.995,1.001]	[1]	[0.0464,0.407]	[1]	[0.995,1.0001]	[1]	[1.821e-02,6.720e-01]
[2]	[9.987e-01, 1.001]	[2]	[2.62e-04, 6.646e-01]	[2]	[9.994e-01 1.00]	[2]	[2.430e-02 ,1.520e-01]
[3]	[8.8e-01,1.119]	[3]	[-9.95,10.33]	[3]	[9.67e-01,1.031]	[3]	[-0.119,0.2314]
γ	9.969	$\bar{\sigma}$	4.158e-03	γ	3.302	$\bar{\sigma}$	2.896e-03
[1]	[5.83,20.704]	[1]	[6.822e-04,7.844e-03]	[1]	[2.238 ,50]	[1]	[4.446e-04,4.380e-03]
[2]	[8.193, 13.35]	[2]	[2.268e-03 , 1.331e-02]	[2]	[1.569, 5.236]	[2]	[3.375e-04, 1.284e-02]
[3]	[-752.65,772.58]	[3]	[-8.038e-02,8.869e-02]	[3]	[-4.90,11.51]	[3]	[-1.71e-03,7.45e-03]
ψ	1.178	ν	9.990e-01	ψ	1.724	ν	9.899e-01
[1]	[0.952,+∞]	[1]	[0.986,1]	[1]	[0.8144,+∞]	[1]	[0,0.9997]
[2]	[1.086, 1.56]	[2]	[9.817e-01, 9.991e-01]	[2]	[1.010,2.621]	[2]	[9.87e-01, 9.996e-01]
[3]	[-27.32,29.67]	[3]	[0.9256,1.072]	[3]	[-31.38,34.83]	[3]	[0.944,1.035]
μ_c	2.338e-03	σ_w	4.868e-06	μ_c	2.290e-03	σ_w	5.000e-06
[1]	[1.388e-03,3.227e-03]	[1]	[1.20e-06,1.415e-05]	[1]	[9.79e-04,2.860e-03]	[1]	[0,2.373e-04]
[2]	[1.538e-03, 3.25e-03]	[2]	[3.783e-06, 4.994e-06]	[2]	[1.242e-03, 2.762e-03]	[2]	[2.152e-06, 4.025e-05]
[3]	[2.059e-03,2.617e-03]	[3]	[-2.7065e-04,2.804e-04]	[3]	[1.997e-03,2.583e-03]	[3]	[-5.588e-05,6.588e-05]
μ_d	1.174e-03	π_c	0.173	μ_d	1.356e-03	π_c	1.170
[1]	[-2.780e-03,4.60e-03]	[1]	[-2.695, 2.695]	[1]	[-9.672e-05,5.886e-03]	[1]	[-1.415e+01,1.650e+01]
[2]	[-1.74e-03, 3.10e-03]	[2]	[1.449e-01, 3.286]	[2]	[-2.127e-03 ,2.516e-03]	[2]	[1.21e-01, 2.327]
[3]	[6.596e-05,2.283e-03]	[3]	[-91.14,91.49]	[3]	[2.246e-04,2.486e-03]	[3]	[-19.78,22.12]
ϕ	1.794	φ_σ	-377.981	ϕ	3.605	φ_σ	1.617e-03
[1]	[0.980,4.499]	[1]	[-4816.642,2269.73]	[1]	[0.781,9.551]	[1]	[-7436.145,3129.627]
[2]	[1.221, 6.855]	[2]	[-8.261e+02 ,6.517e-01]	[2]	[1.625, 7.325]	[2]	[-2.016e+02 , 2.016e+02]
[3]	[-38.82,42.41]	[3]	[-3.8252e+04,3.750e+04]	[3]	[3.94e-01,6.82]	[3]	[-3.917e+03,3.917e+03]
φ_d	2.942	π_σ	-3.016e+03	φ_d	5.351	π_σ	3.007
[1]	[-7.445,7.445]	[1]	[-1.611e+05,1472.09]	[1]	[-23.749,23.749]	[1]	[-2.351e+05,12828.74]
[2]	[5.045e-01, 6.311]	[2]	[-6032.854, 1.268]	[2]	[2.849, 8.145]	[2]	[-2.846e+03, -2.846e+03]
[3]	[-176.54,182.43]	[3]	[-1.27545e+05,1.21513e+05]	[3]	[-31.13,13.82]	[3]	[-8.599e+03,8.605e+03]
ρ	0.827	TJ_T	109.215	ρ	9.900e-01	TJ_T	39.516
[1]	[0.687,0.978]	[1]		[1]	[0.9825,0.9981]	[1]	
[2]	[5.464e-01, 9.912e-01]	[2]	[6.77, 172.1]	[2]	[0.986, 9.985e-01]	[2]	
[3]	[-3.14,4.80]	[3]		[3]	[9.52e-01,1.028]	[3]	

This table shows the results of the GMM estimation of the Extended LRR model. The first two columns correspond to the estimates with the negativity constraint on the slope and the last two columns are for the estimates without the constraint on the slope. We provide the 95 % confidence intervals obtained by three methods: [1] is the modified projection method following the procedure explained in 1.10.10, [2] is the parametric boostrap and [3] is the Delta method.

Table 1.7: Estimations results (restricted models)

	Standard BY Model		Extended Model (EIS=1)	
par	Constr.	NoConstr.	Constr.	NoConstr.
δ	0.9988 [0.925,1.072]	0.9983 [0.993,1.004]	0.99995 [9.985e-01,1.001]	1.000 [9.984e-01,1.002]
γ	10.00 [-45.550,65.550]	7.296 [-12.163,26.755]	10 [-3.162e+01,5.162e+01]	8.286 [-1.864e+01,3.522e+01]
ψ	1.463 [-41.987,44.912]	15.00 [-6.175e+02,6.475,e+02]	1 -	1 -
μ_c	2.480e-03 [2.237e-03,2.723e-03]	2.058e-03 [1.806e-03,2.310e-03]	0.0023 [2.045e-03,2.604e-03]	1.976e-03 [1.737e-03,2.216e-03]
μ_d	0.00 [-1.134e-03,1.134e-03]	3.040e-03 [2.027e-03,4.053e-03]	0.00164 [5.471e-04,2.737e-03]	2.289e-03 [8.875e-04,3.690e-03]
ϕ	0.00 [-16.429,16.429]	2.311 [-19.795,24.418]	5.065 [-2.361e+04,2.362e+04]	2.758 [-2.541,8.056]
φ_d	0.573 [-118.880,120.026]	2.236 [-2.121e+03,2.125e+03]	10.054 [-1.963e+02,2.164e+02]	1.957 [-1.165e+03,1.169e+03]
ρ	0.988 [9.433e-01,1.033]	7.947e-01 [-6.608,8.197]	0.821 [-8.219e+02,8.236e+02]	9.838e-01 [8.922e-01,1.075]
φ_e	2.721e-02 [-5.184e-02,1.063e-01]	0.3615 [-42.61,43.33]	0.0155 [-6.242e+01,6.245e+01]	2.098e-01 [-8.884e-01,1.308]
$\bar{\sigma}$	4.820e-03 [2.692e-03,6.947e-03]	2.796e-03 [-0.172,0.177]	0.00115 [-2.085e-02,2.314e-02]	9.942e-04 [-1.321e-03,3.309e-03]
ν	0.9987 [0.959,1.039]	0.997 [0.996,0.999]	0.9989 [9.972e-01,1.001]	9.214e-01 [7.355e-01,1.107]
σ_w	3.870e-06 [3.157e-06,4.583e-06]	4.609e-06 [-5.409e-04,5.501e-04]	1.2018e-06 [1.187e-06,1.217e-06]	1.703e-06 [1.195e-06,2.211e-06]
π_c	6.35 [-6.468,19.171]	10.934 [-1.124e+03,1.146e+03]	1.612 [-3.222e+02,3.255e+02]	5.958 [-4.230e+02,4.348e+02]
φ_σ	0 -	0	-3920.078 [-1.1141e+04,3.3018e+03]	1.107e+03 [-9.784e+03,1.200e+04]
π_σ	0 -	0 -	-2.854e-16 [-2.624e+04,2.624e+04]	-9.144e+03 [-8.037e+04,6.209e+04]
TJ_T	122.854	41.486	118.447	57.265

This table shows the results of the GMM estimations of the restricted models. The first two columns correspond to the standard LRR model and the last two columns are for the extended model but with EIS=1 as developed along the paper. The confidence intervals made by the Delta method are provided in brackets below the estimates.

Table 1.8: Standardized errors

	Standard BY Model		Extended Model (EIS=1)		Extended Model	
Moment	Constr.	NoConstr.	Constr.	NoConstr.	Constr.	NoConstr.
$\mathrm{E}(g_t^a)$	2.394	0.250	1.602	-0.167	1.673	1.427
$\mathrm{Var}(g_t^a)$	-1.716	-2.162	-2.362	-2.190	-1.678	-1.966
$\mathrm{ACV1}(g_t^a)$	0.960	-0.109	-0.647	1.507	0.211	1.802
$\mathrm{E}(g_t^a - \mathrm{E}(g_t^a))^3$	-1.329	-1.329	-1.380	-1.252	-1.451	-1.329
$\mathrm{E}(g_{d,t}^a)$	-1.932	1.788	0.077	0.868	-0.495	-0.273
$\mathrm{Var}(g_{d,t}^a)$	-0.309	-0.097	-1.598	-1.098	-1.152	-0.947
$\mathrm{ACV1}(g_{d,t}^a)$	1.073	1.388	-0.232	0.993	0.280	1.746
$\mathrm{E}(g_{d,t}^a - \mathrm{E}(g_{d,t}^a))^3$	0.367	0.367	0.367	0.2678	0.308	0.367
$\mathrm{CV}(g_t^a, g_{d,t}^a)$	0.336	0.140	-1.331	-0.923	-0.690	-0.311
$\mathrm{E}(pd_t^a)$	0.190	2.326	0.822	0.379	-0.946	2.111
$\mathrm{Var}(pd_t^a)$	-2.483	1.323	-1.329	-2.254	-2.579	-0.549
$\mathrm{ACV1}(pd_t^a)$	-2.379	1.336	-1.183	-2.268	-2.444	-0.614
$\mathrm{E}(r_{f,t}^a)$	2.081	1.228	3.193	3.485	2.021	2.638
$\mathrm{Var}(r_{f,t}^a)$	-1.760	-1.965	-2.140	-1.829	-1.679	-1.918
$\mathrm{ACV1}(r_{f,t}^a)$	-1.359	-1.71	-2.008	-1.531	-1.450	-1.649
$\mathrm{E}(r_{m,t}^a)$	-2.280	-0.665	-1.309	-0.771	-1.195	-1.765
$\mathrm{Var}(r_{m,t}^a)$	-3.500	0.121	-4.854	-0.904	-4.883	0.694
$\mathrm{ACV1}(r_{m,t}^a)$	0.287	0.258	0.317	-1.288	0.335	0.240
$\mathrm{CV}(r_{t+1}^{ex}, pd_t^a)$	2.225	-0.090	2.019	1.660	2.45	1.290
$\mathrm{CV}(r_{t+1,t+3}^{ex}, pd_t^a)$	2.315	-0.586	2.050	2.048	2.605	1.252
$\mathrm{CV}(r_{t+1,t+5}^{ex}, pd_t^a)$	2.468	-1.120	2.132	2.349	2.833	1.274

Notes: This table shows the standardized errors obtained by dividing the mean error from each moment condition by its Standard. Error from the HAC variance-covariance matrix.

Table 1.9: Returns regressions on VIX and 3 Fama-French factors

B/M portfolios deciles	Growth	Dec2	Dec3	Dec4	Dec5	Dec6	Dec7	Dec8	Dec9	Value	V-G
Panel A: $R_{i,t} - R_{f,t} = \alpha_i + \beta\Delta\text{VIX}_t + \varepsilon_{i,t}$											
α_i	0,52	0,55	0,60	0,50	0,50	0,57	0,35	0,44	0,67	0,63	0,11
	(2,29)	(2,75)	(3,23)	(2,66)	(2,58)	(2,91)	(1,24)	(1,83)	(2,71)	(1,96)	(0,34)
β(VIX)	-0,67	-0,65	-0,63	-0,71	-0,68	-0,61	-0,73	-0,65	-0,72	-0,97	-0,30
	(-10,74)	(-10,09)	(-8,42)	(-10,32)	(-9,56)	(-9,33)	(-12,84)	(-9,80)	(-8,65)	(-9,20)	(-4,34)
Adj.R^2	0,26	0,30	0,29	0,35	0,32	0,28	0,33	0,26	0,28	0,30	0,05
Panel B: $R_{i,t} - R_{f,t} = \alpha_i + \beta_1\text{Rm-Rf}_t + \beta_2\text{SMB}_t + \beta_3\text{HLM}_t + \beta_4\Delta\text{VIX}_t + \varepsilon_{i,t}$											
α_i	0,01	-0,01	-0,02	-0,12	-0,15	-0,07	-0,33	-0,36	-0,16	-0,39	-0,41
	(0,15)	(-0,11)	(-0,26)	(-1,17)	(-1,59)	(-0,72)	(-2,64)	(-3,33)	(-1,36)	(-3,00)	(-2,73)
β(Rm-Rf)	1,01	0,94	0,96	0,91	0,91	0,87	0,94	0,99	1,07	1,22	0,21
	(42,31)	(22,52)	(27,06)	(33,31)	(29,59)	(18,80)	(25,18)	(21,88)	(26,38)	(34,84)	(6,00)
β(SMB)	-0,15	-0,06	-0,05	-0,02	-0,04	-0,02	-0,04	0,23	0,19	0,44	0,59
	(-4,33)	(-1,68)	(-0,88)	(-0,34)	(-0,56)	(-0,36)	(-0,69)	(4,93)	(5,00)	(9,25)	(11,60)
β(HML)	-0,42	-0,07	0,15	0,29	0,43	0,48	0,50	0,69	0,65	0,90	1,32
	(-9,33)	(-0,81)	(2,12)	(3,81)	(7,15)	(4,80)	(5,49)	(9,62)	(15,99)	(14,46)	(17,86)
$\beta(\Delta$VIX)	0,01	-0,01	0,01	-0,10	-0,08	-0,04	-0,12	0,04	0,02	-0,10	-0,11
	(0,85)	(-0,54)	(0,30)	(-3,86)	(-2,31)	(-1,39)	(-1,87)	(0,94)	(0,54)	(-2,31)	(-2,45)
Adj.R^2	0,94	0,90	0,89	0,87	0,87	0,84	0,86	0,90	0,91	0,86	0,74

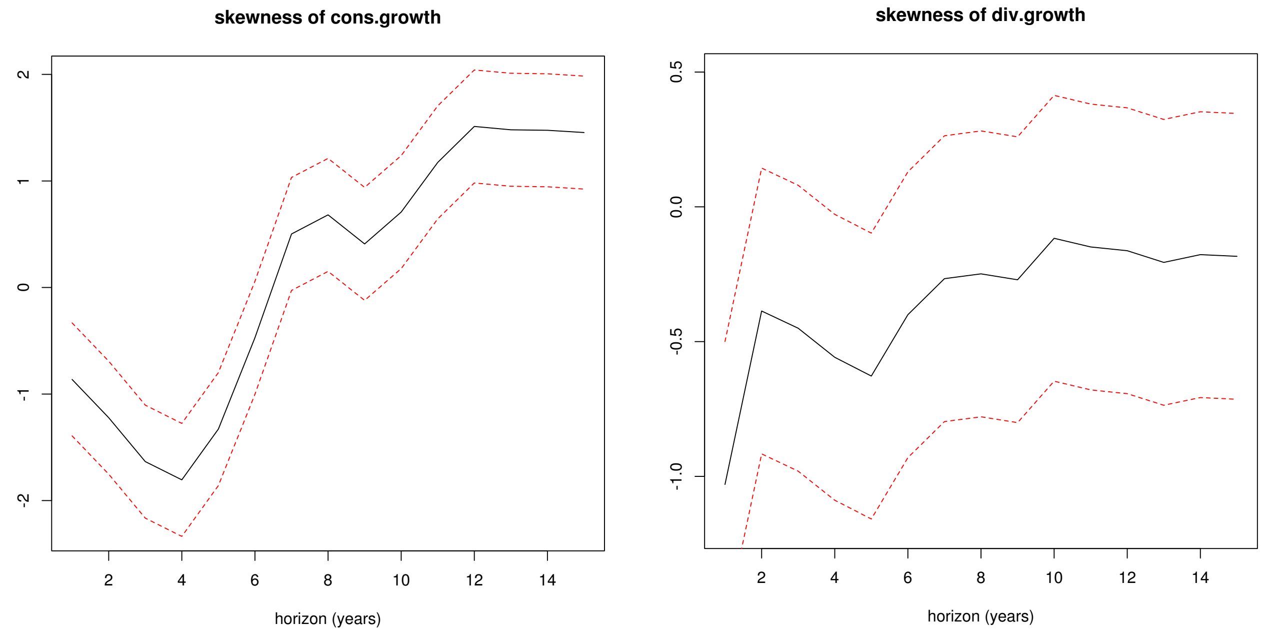
Notes: This table reports the coefficients, t-stat and adjusted R-square of the regression of the excess return of Book to Market sorted portfolio deciles on the first order difference of the VIX index (Panel A) and on the Fama-French 3 factors: excess market return, Small minus Big, High minus Low (Panel B). All the variables are expressed at the monthly frequency. t-statistics are in brackets, they have been computed using the Newey-West Heteroskedasticity and Autocorrelation Consistent estimator of the variance-covariance matrix of residuals with automatically selected number of lag (see Newey and West [1994]). The dependent variable is the value-weighted return (dividends excluded) on Fama-French portfolio deciles sorted by Book to Market ratios minus 1-Month Treasury Bill rate from Kenneth French data library.

Table 1.10: Regression of log-dividend price ratio and dividend yields on macro uncertainty

$dy_{i,t} = a_{1,i} + b_{1,i}h_{1,t} + \epsilon_{i,t}$											
	d/p	dy1	dy2	dy3	dy4	dy5	dy6	dy7	dy8	dy9	dy10
cst.	-4,514	-0,580	-0,397	-0,282	-0,230	-0,197	-0,172	-0,154	-0,140	-0,137	-0,130
s.e	0,171	0,080	0,058	0,038	0,028	0,022	0,019	0,017	0,015	0,014	0,013
slope	0,902	0,808	0,541	0,379	0,308	0,265	0,234	0,212	0,194	0,190	0,181
s.e	0,256	0,121	0,088	0,056	0,040	0,032	0,026	0,023	0,021	0,019	0,018
R-squared	0,440	0,764	0,706	0,666	0,683	0,679	0,670	0,663	0,662	0,711	0,722
N	142	131	142	142	142	142	142	142	142	132	120
$dy_{i,t} = a_{2,i} + b_{2,i}h_{2,t} + \epsilon_{i,t}$											
	d/p	dy1	dy2	dy3	dy4	dy5	dy6	dy7	dy8	dy9	dy10
cst.	-4,614	-0,671	-0,459	-0,325	-0,265	-0,227	-0,199	-0,178	-0,162	-0,157	-0,149
s.e	0,197	0,091	0,066	0,043	0,031	0,025	0,021	0,019	0,017	0,016	0,015
slope	0,872	0,782	0,525	0,367	0,299	0,257	0,227	0,205	0,188	0,183	0,173
s.e	0,244	0,114	0,083	0,053	0,038	0,030	0,025	0,022	0,020	0,018	0,017
R-squared	0,446	0,777	0,722	0,681	0,698	0,693	0,682	0,674	0,672	0,717	0,726
N	142	131	142	142	142	142	142	142	142	132	120
$dy_{i,t} = a_{3,i} + b_{3,i}h_{3,t} + \epsilon_{i,t}$											
	d/p	dy1	dy2	dy3	dy4	dy5	dy6	dy7	dy8	dy9	dy10
Cst.	-5,240	-1,255	-0,856	-0,607	-0,496	-0,426	-0,375	-0,337	-0,307	-0,295	-0,278
s.e	0,405	0,191	0,136	0,085	0,061	0,048	0,040	0,036	0,033	0,030	0,028
Slope	1,442	1,317	0,890	0,627	0,512	0,441	0,389	0,352	0,322	0,310	0,293
s.e	0,441	0,114	0,083	0,053	0,038	0,030	0,025	0,022	0,020	0,018	0,017
R-squared	0,427	0,773	0,725	0,695	0,717	0,714	0,704	0,696	0,692	0,727	0,733
N	142	131	142	142	142	142	142	142	142	132	120

This table show the regressions of the log-dividend price ratio and dividend yields on the measures of macro uncertainty (h1,h2 and h3) provided by Jurado et al. [2015]. The dividend yields are with maturities ranging from 1 year (dy1) to 10 years (dy10). h_1, h_2, h_3 are respectively the predicted macro uncertainty 1 month, 3 month and 12 months ahead. The standard errors have been computed using the Newey and West [1987] estimator with 8 lags.

Figure 1.8: Term structures of the skewness of consumption and dividend growths



This figure represents the evolution of the skewness of the temporally aggregated growth rate of consumption (left) and dividends (right) as a function of the horizon of aggregation (from 1 to 15 years). The skewness are negative in the short-run and increase as the horizon of aggregation increases.

Notes: This graph shows the variance ratio statistic for different horizons in the cases where the cash flow process is (i) i.i.d (top panel), (ii) it is simulated using the BKY calibration: meaning that there is serial auto-correlation and the volatility is stochastic but independent front the cash flow process (middle panel) and (iii) it is simulated using the BKY calibration but allowing the cash flow process to be negatively correlated with stochastic volatility process (bottom panel)

Notes: This graph shows the loading on the expected consumption growth of the dividend strips log price dividend ratio (left panel) and the loading on the consumption growth volatility (right panel) implied by the model using the estimated parameters (Extended Model - NoConstr. for the top panel and Extended Model - Constr. for the bottom panel) for calibration. The Top panel has been obtained using the estimates without the term structure constraint (Extended Model - NoConstr.) while the bottom panel uses the estimates with the term structure constraint (Extended Model - Constr.)

1.10.2 Model solution when EIS$\neq$1

The LRR model assumes a rational representative agent embedded with Epstein and Zin [1989] recursive utility function given by 1.10.1 who maximizes its continuation value subject to its inter-temporal budget constraint.

Figure 1.9: VR simulation of cash flow with and without correlation with stochastic volatility

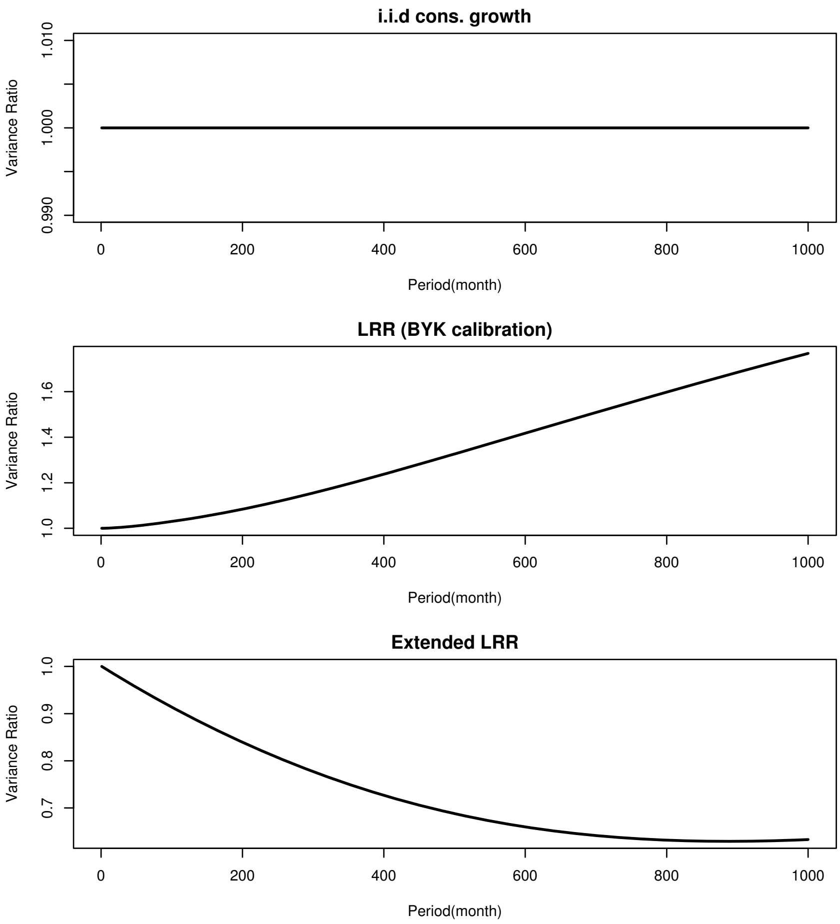
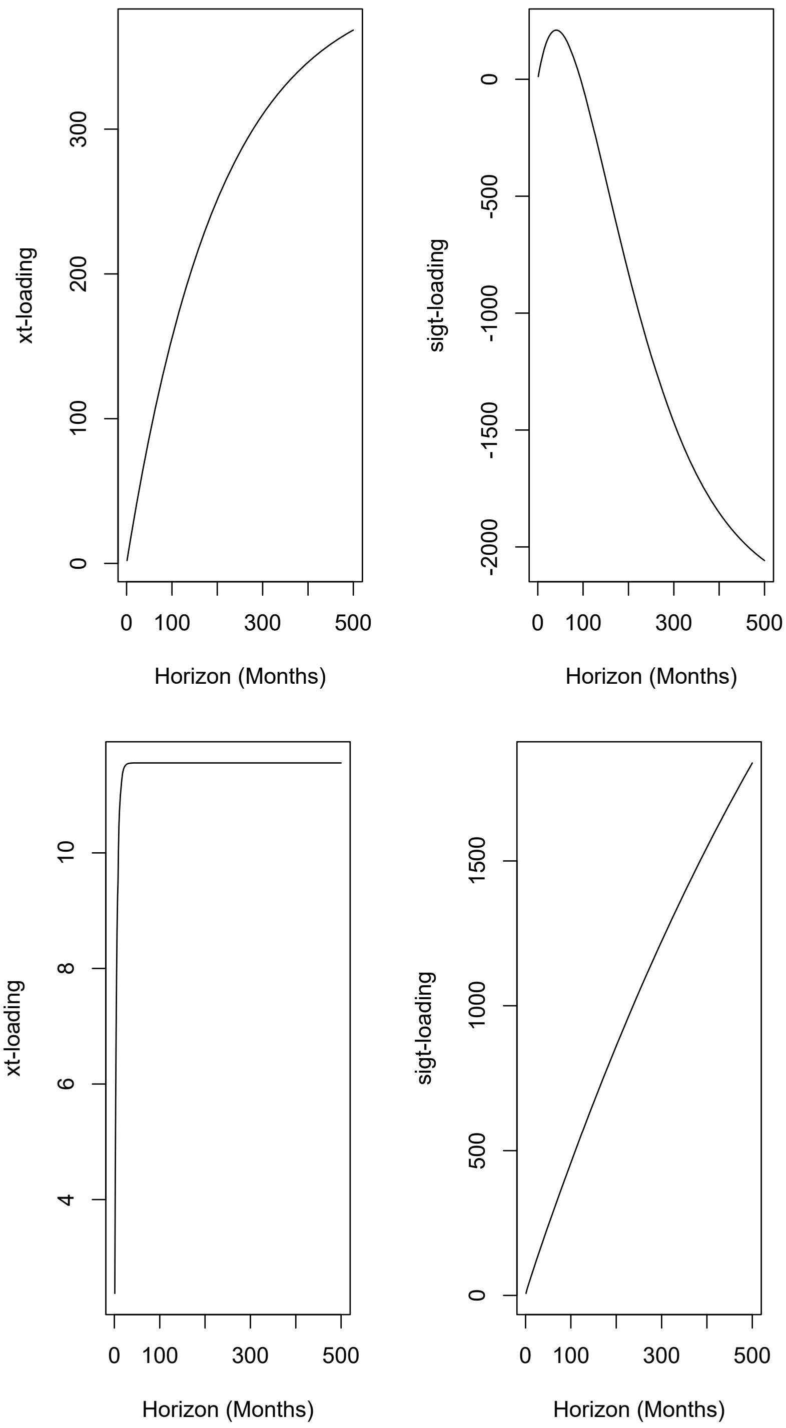


Figure 1.10: Loadings implied by the standard LRR model and our extended LRR model



$$V_t = \begin{cases} \left[(1-\delta)C_t^{1-\frac{1}{\psi}} + \delta\left(E_t\left(V_{t+1}^{1-\gamma}\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right)\right]^{\frac{1}{1-\frac{1}{\psi}}} & if \quad \psi \neq 1, \gamma \neq 1 \\ C_t^{1-\delta}\left(E_t\left(V_{t+1}^{1-\gamma}\right)^{\frac{\delta}{1-\gamma}}\right) & if \quad \psi = 1, \gamma \neq 1 \\ \left[(1-\delta)C_t^{1-\frac{1}{\psi}} + \delta exp\left(E_t\left(logV_{t+1}\right)\right)^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}} & if \quad \psi \neq 1, \gamma = 1 \\ C_t^{1-\delta} exp\left(E_t\left(logV_{t+1}\right)\right)^{\delta} & if \quad \psi = 1, \gamma = 1 \end{cases} \tag{1.10.1}$$

The SDF is given by :

$$\begin{aligned} m_{t+1} &= \theta \log\delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta-1)r_{c,t+1} \\ &= \theta \log\delta - \frac{\theta}{\psi}(\mu + x_t + \sigma_t\varepsilon_{c,t+1} + \varphi_\sigma\sigma_w w_{t+1}) + (\theta-1)r_{c,t+1} \\ &= \theta \log\delta - \frac{\theta}{\psi}\mu + (\theta-1)\left[k_0 + (k_1-1)A_0 + \mu_c + k_1A_2(1-\nu)\bar{\sigma}^2\right] + \left[-\frac{\theta}{\psi} + (\theta-1)\left((k_1\rho-1)A_1+1\right)\right]x_t \\ &+ \left[(\theta-1)(k_1\nu-1)A_2\right]\sigma_t^2 + \left[(\theta-1)\left(k_1A_1\varphi_e\right)\right]\sigma_t\varepsilon_{x,t+1} + \left[-\frac{\theta}{\psi} + (\theta-1)\right]\sigma_t\varepsilon_{c,t+1} \\ &+ \left[-\frac{\theta}{\psi}\varphi_\sigma + (\theta-1)\left(k_1A_2 + \varphi_\sigma\right)\right]\sigma_w w_{t+1} \\ &= \theta\log\delta - \gamma\mu_c + (\theta-1)\left[k_0 + (k_1-1)A_0 + \mu_c + k_1A_2(1-\nu)\bar{\sigma}^2\right] - \frac{1}{\psi}x_t + (\theta-1)(k_1\nu-1)A_2\sigma_t^2 \\ &+ \left[(\theta-1)k_1A_1\varphi_e\right]\sigma_t\varepsilon_{x,t+1} + \left[-\gamma\right]\sigma_t\varepsilon_{c,t+1} + \left[-\gamma\varphi_\sigma + (\theta-1)k_1A_2\right]\sigma_w w_{t+1} \end{aligned}$$

So,

$$m_{t+1} = a_{0m} + a_{1m}x_t + a_{2m}\sigma_t^2 + \lambda_c\sigma_t\varepsilon_{c,t+1} + \lambda_x\sigma_t\varepsilon_{x,t+1} + \lambda_w\sigma_w w_{t+1} \tag{1.10.2}$$

Where

$$\begin{aligned} a_{0m} &= \log\delta - \frac{1}{\psi}\mu_c + \frac{1}{2}\theta(1-\theta)\left[(1-\frac{1}{\psi})\varphi_\sigma + k_1A_2\right]^2\sigma_w^2 \\ a_{1m} &= -\frac{1}{\psi} \\ a_{2m} &= (\theta-1)(k_1\nu-1)A_2 \end{aligned}$$

and the prices of risk are given by :

$$-\lambda_c = \gamma \tag{1.10.3}$$

$$-\lambda_x = (1-\theta)k_1A_1\varphi_e \tag{1.10.4}$$

$$-\lambda_w = \gamma\varphi_\sigma + (1-\theta)k_1A_2 \tag{1.10.5}$$

1.10.3 is the price of the short-run consumption risk, 1.10.4 is the price of the long-run consumption risk and 1.10.5 is the price of the volatility risk).

We have that:

$$\begin{aligned} m_{t+1} + r_{c,t+1} &= \theta \log \delta + (1-\gamma)\mu_c + \theta \left[k_0 + (k_1 - 1)A_0 + k_1 A_2 (1-\nu)\bar{\sigma}^2\right] + \left[1 - \tfrac{1}{\psi} + (k_1\rho - 1)A_1\right] x_t \\ &+\theta(k_1\nu - 1)A_2\sigma_t^2 + \left[\theta k_1 A_1 \varphi_e\right] \sigma_t \varepsilon_{x,t+1} + \left[(1-\gamma)\right] \sigma_t \varepsilon_{c,t+1} \\ &+ \left[(1-\gamma)\varphi_\sigma + \theta k_1 A_2\right] \sigma_w \varepsilon_{w,t+1} \end{aligned}$$

So, $\mathrm{E}_t(\exp(m_{t+1} + r_{c,t+1})) = 1$ implies that:

$$A_0 = \tfrac{1}{1-k_1} \left[\log \delta + (1 - \tfrac{1}{\psi})\mu_c + k_0 + k_1 A_2(1-\nu)\bar{\sigma}^2 + \tfrac{1}{2}\theta \left((1-\tfrac{1}{\psi})\varphi_\sigma + k_1 A_2\right)^2 \sigma_w^2\right]$$

$$A_1 = \frac{1-\frac{1}{\psi}}{1-k_1\rho}$$

$$A_2 = \frac{(1-\gamma)(1-\frac{1}{\psi})}{2(1-k_1\nu)} \left[1 + \left(\frac{k_1\varphi_e}{1-k_1\rho}\right)^2\right]$$

Assuming that the log-price dividend ratio is an affine function of the state variables:

$$pd_t^{(n)} = A_0(n) + A_1(n)x_t + A_2(n)\sigma_t^2$$

As

$$\begin{aligned} m_{t+1} + \Delta d_{t+1} + pd_{t+1}^{(n-1)} &= \left[a_{0m} + \mu_d + A_0(n-1) + A_2(n-1)(1-\nu)\bar{\sigma}^2\right] + \left[a_{1m} + \phi + A_1(n-1)\rho\right] x_t \\ &+ \left[a_{2m} + A_2(n-1)\nu\right] \sigma_t^2 + \left[\lambda_c + \pi_c\right] \sigma_t \varepsilon_{c,t+1} + \varphi_d \sigma_t u_{t+1} \\ &+ \left[\lambda_x + \varphi_e A_1(n-1)\right] \sigma_t \varepsilon_{x,t+1} + \left[\lambda_w + \pi_\sigma + A_2(n-1)\right] \sigma_w \varepsilon_{w,t+1} \end{aligned}$$

The log-price dividend ratio of the n-periods dividend strip can be expressed as follow:

$$\begin{aligned} pd_t^{(n)} &= \log \mathrm{E}_t \left[\exp\left(m_{t+1} + \Delta d_{t+1} + pd_{t+1}^{(n-1)}\right)\right] \\ &= \left[a_{0m} + \mu_d + A_0(n-1) + A_2(n-1)(1-\nu)\bar{\sigma}^2\right] + \left[a_{1m} + \phi + A_1(n-1)\rho\right] x_t \\ &+ \left[a_{2m} + A_2(n-1)\nu + \tfrac{1}{2}\left(\lambda_c + \pi_c\right)^2 + \tfrac{1}{2}\left(\lambda_x + \varphi_e A_1(n-1)\right)^2 + \tfrac{1}{2}\varphi_d^2\right] \sigma_t^2 \\ &+ \left[\lambda_w + \pi_\sigma + A_2(n-1)\right]^2 \sigma_w^2 \end{aligned}$$

Implying that $n \geq 1$[29]:

$$A_0(n) = A_0(n-1) + a_{0m} + \mu_d + A_2(n-1)(1-\nu)\bar{\sigma}^2 + \tfrac{1}{2}\left[\lambda_w + \pi_\sigma + A_2(n-1)\right]^2 \sigma_w^2$$

$$A_1(n) = \phi + a_{1m} + A_1(n-1)\rho = \left(\phi - \tfrac{1}{\psi}\right)\left(\tfrac{1-\rho^n}{1-\rho}\right)$$

$$A_2(n) = a_{2m} + A_2(n-1)\nu + \tfrac{1}{2}\left(\lambda_c + \pi_c\right)^2 + \tfrac{1}{2}\left(\lambda_x + \varphi_e A_1(n-1)\right)^2 + \tfrac{1}{2}\varphi_d^2$$

Which is equivalent to:

$$\begin{aligned} A_0(n) &= A_0(n-1) + a_{0m} + \mu_d + A_2(n-1)(1-\nu)\bar{\sigma}^2 \\ &+ \tfrac{1}{2}\left[\pi_\sigma - \gamma\varphi_\sigma + (\theta-1)k_1 A_2 + A_2(n-1)\right]^2 \sigma_w^2 \\ A_1(n) &= \left(\phi - \tfrac{1}{\psi}\right)\left(\tfrac{1-\rho^n}{1-\rho}\right) \\ A_2(n) &= A_2(n-1)\nu + (\theta-1)\left(k_1\nu - 1\right)A_2 + \tfrac{1}{2}\left(\pi_c - \gamma\right)^2 \\ &+ \tfrac{1}{2}\left(\left((\theta-1)k_1 A_1 + A_1(n-1)\right)\varphi_e\right)^2 + \tfrac{1}{2}\varphi_d^2 \end{aligned} \tag{1.10.6}$$

For the zero coupon bond,

$$P_t^{(n)} = \exp\left(B_0(n) + B_1(n)x_t + B_2(n)\sigma_t^2\right)$$

where

$$\begin{aligned} B_0(n) &= B_0(n-1) + a_{0m} + B_2(n-1)(1-\nu)\bar{\sigma}^2 + \tfrac{1}{2}\left[\lambda_w + B_2(n-1)\right]^2 \sigma_w^2 \\ B_1(n) &= -\tfrac{1}{\psi}\left(\tfrac{1-\rho^n}{1-\rho}\right) \\ B_2(n) &= B_2(n-1)\nu + (\theta-1)(k_1\nu - 1)A_2 + \tfrac{1}{2}\lambda_c^2 + \tfrac{1}{2}\left(\lambda_x + B_1(n-1)\varphi_e\right)^2 \end{aligned} \tag{1.10.7}$$

with $B_2(1) = (-\frac{1}{2} + \gamma)$

The return on the risk free asset is given by:

$$r_{f,t} = -\mathrm{E}_t(m_{t+1}) - \frac{1}{2}\mathrm{Var}_t(m_{t+1}) = -a_{0m} - \frac{1}{2}\lambda_w^2\sigma_w^2 - a_{1m}x_t - \left(a_{2m} + \frac{1}{2}(\lambda_c^2 + \lambda_x^2)\right)\sigma_t^2$$

The return on the n-period dividend strip is given by :

[29] $A_0(0) = 0, A_1(0) = 0, A_2(0) = 0$

$$
\begin{aligned}
r_{d,t+1}^{(n)} &= \Delta d_{t+1} + pd_{t+1}^{(n-1)} - pd_t^{(n)} \\
&= \left[\mu_d + A_0(n-1) - A_0(n) + A_2(n-1)(1-\nu)\bar{\sigma}^2\right] + \left(\phi + A_1(n-1)\rho - A_1(n)\right) x_t \\
&+ \left(A_2(n-1)\nu - A_2(n)\right)\sigma_t^2 + \left(A_1(n-1)\varphi_e\right)\sigma_t \varepsilon_{x,t+1} + \pi_c \sigma_t \varepsilon_{c,t+1} \\
&+ \left(\pi_\sigma + A_2(n-1)\right)\sigma_w \varepsilon_{w,t+1} + \varphi_d \sigma_t \varepsilon_{d,t+1}
\end{aligned}
$$

So, the excess return is given by:

$$
\begin{aligned}
r_{d,t+1}^{(n)} - r_{f,t} &= \left[a_{0m} + \frac{1}{2}\lambda_w^2 \sigma_w^2 + \mu_d + \left(A_0(n-1) - A_0(n)\right) + A_2(n-1)(1-\nu)\bar{\sigma}^2\right] \\
&+ \left(\phi - \frac{1}{\psi} + A_1(n-1)\rho - A_1(n)\right) x_t + \left(A_2(n-1)\nu - A_2(n) + a_{2m} + \frac{1}{2}\left(\lambda_c^2 + \lambda_x^2\right)\right)\sigma_t^2 \\
&+ \left(A_1(n-1)\varphi_e\right)\sigma_t \varepsilon_{x,t+1} + \pi_c \sigma_t \varepsilon_{c,t+1} + \left(\pi_\sigma + A_2(n-1)\right)\sigma_w \varepsilon_{w,t+1} + \varphi_d \sigma_t \varepsilon_{d,t+1} \\
&= \frac{1}{2}\left[\lambda_w^2 \sigma_w^2 - \left(\pi_\sigma + \lambda_w + A_2(n-1)\right)^2 \sigma_w^2\right] \\
&+ \frac{1}{2}\left[\lambda_c^2 - (\pi_c + \lambda_c)^2 + \lambda_x^2 - \left(\lambda_x + A_1(n-1)\varphi_e\right)^2 - \varphi_d^2\right]\sigma_t^2 \\
&+ \left(A_1(n-1)\varphi_e\right)\sigma_t \varepsilon_{x,t+1} + \pi_c \sigma_t \varepsilon_{c,t+1} + \left(\pi_\sigma + A_2(n-1)\right)\sigma_w \varepsilon_{w,t+1} + \varphi_d \sigma_t \varepsilon_{d,t+1} \\
&= \frac{1}{2}\left[-\left(\pi_\sigma + \lambda_w + A_2(n-1)\right)^2 - 2\lambda_w\left(\pi_\sigma + A_2(n-1)\right)\right]\sigma_w^2 \\
&+ \frac{1}{2}\left[-\pi_\sigma^2 - 2\lambda_c \pi_c - \left(A_1(n-1)\varphi_e\right)^2 - 2\lambda_x\left(A_1(n-1)\varphi_e\right) - \varphi_d^2\right]\sigma_t^2 \\
&+ \left(A_1(n-1)\varphi_e\right)\sigma_t \varepsilon_{x,t+1} + \pi_c \sigma_t \varepsilon_{c,t+1} + \left(\pi_\sigma + A_2(n-1)\right)\sigma_w \varepsilon_{w,t+1} + \varphi_d \sigma_t \varepsilon_{d,t+1} \qquad (1.10.8)
\end{aligned}
$$

We then deduce the risk premium on the n-period dividend strip:

$$
\begin{aligned}
rp_t^{(n)} &= \mathrm{E}_t(r_{d,t+1}^{(n)} - r_{f,t}) + \frac{1}{2}\mathrm{Var}_t(r_{d,t+1}^{(n)} - r_{f,t}) \\
&= -\lambda_w\left(\pi_\sigma + A_2(n-1)\right)\sigma_w^2 + \left[-\lambda_c\left(\pi_c\right) - \lambda_x\left(A_1(n-1)\varphi_e\right)\right]\sigma_t^2 \qquad (1.10.9)
\end{aligned}
$$

So the risk premium is a weighted sum of the risk prices. Under preference for early resolution of uncertainty (which happens when $\gamma > 1$and $\psi > 1$), the price of volatility risk is negative while the prices of long-run and short-run consumption risks are both positive.

As $A_1(0) = 0$ and $A_2(0) = 0$, the risk premium on the 1-period dividend strip return is :

$$
\begin{aligned}
rp_t^{(1)} &= \mathrm{E}_t(r_{d,t+1}^{(1)} - r_{f,t}) + \frac{1}{2}\mathrm{Var}_t(r_{d,t+1}^{(1)} - r_{f,t}) \\
&= -\lambda_w \pi_\sigma \sigma_w^2 + \left[-\lambda_c \pi_c\right] \sigma_t^2 \qquad (1.10.10)
\end{aligned}
$$

compared to the standard BY model, there is one new term in the one month risk premium because of the cross correlation between cash flows shocks and the volatility shocks. If $\pi_\sigma < 0$, meaning that dividend growth reacts negatively to an increase of uncertainty, then the short-run risk premium should be higher compared to the case where the cross correlation is not taken into account.

The short term spread defines as the difference between n-periods dividend strip risk premium and $n-1$-period dividend strip risk premium is given by:

$$
\begin{aligned}
S_{n,t}^{(1)} &= rp_t^{(n)} - rp_t^{(n-1)} \\
&= -\lambda_w \left(\left[A_2(n-1) - A_2(n-2)\right]\right) \sigma_w^2 - \lambda_x \left[A_1(n-1) - A_1(n-2)\right] \varphi_e \sigma_t^2
\end{aligned}
$$

As

$$
A_2(n) - A_2(n-1) = A_2(1)\nu^{n-1} + \left(\phi - \frac{1}{\psi}\right)\left[\left(\frac{\nu^{n-1} - \rho^{n-1}}{\nu - \rho}\right) F + \left(\frac{\nu^{n-1} - \rho^{2(n-1)}}{\nu - \rho^2}\right) G\right]
$$

where

$$
\begin{aligned}
F &= \left[\left(\frac{\phi - \frac{1}{\psi}}{1-\rho}\right) + (\theta - 1)k_1 A_1\right] \varphi_e^2 \\
G &= -\frac{1}{2}\left(\frac{1+\rho}{1-\rho}\right)\left(\phi - \frac{1}{\psi}\right) \varphi_e^2
\end{aligned}
$$

Therefore

$$
\begin{aligned}
S_{n,t}^{(1)} &= -\lambda_w \left(A_2(1)\nu^{n-2} + \left(\phi - \frac{1}{\psi}\right)\left[\left(\frac{\nu^{n-2} - \rho^{n-2}}{\nu - \rho}\right) F + \left(\frac{\nu^{n-2} - \rho^{2(n-2)}}{\nu - \rho^2}\right) G\right]\right) \sigma_w^2 \qquad (1.10.11) \\
&\quad - (\lambda_x \varphi_e)\left(\phi - \frac{1}{\psi}\right) \rho^{n-2} \sigma_t^2
\end{aligned}
$$

The long term spread between n-periods dividend strip return and 1-period dividend strip return is given by:

$$S_t^{(n)} = rp_t^{(n)} - rp_t^{(1)} = \sum_{i=2}^{n} S_{i,t}^{(1)} \tag{1.10.12}$$

$$= -\lambda_w \left(A_2(n-1)\right)\sigma_w^2 + \left[-\lambda_x A_1(n-1)\varphi_e\right]\sigma_t^2$$

$$= -\lambda_w \left(A_2(1)\left(\frac{1-\nu^{n-1}}{1-\nu}\right) + \left(\phi - \frac{1}{\psi}\right)\left[\left(\frac{1-\nu^{n-1}}{1-\nu} - \frac{1-\rho^{n-1}}{1-\rho}\right)\frac{F}{\nu-\rho}\right.\right. \tag{1.10.13}$$

$$\left.\left. + \left(\frac{1-\nu^{n-1}}{1-\nu} - \frac{1-\rho^{2(n-1)}}{1-\rho^2}\right)\frac{G}{\nu-\rho^2}\right]\right)\sigma_w^2 - (\lambda_x\varphi_e)\left(\phi - \frac{1}{\psi}\right)\left(\frac{1-\rho^{n-1}}{1-\rho}\right)\sigma_t^2$$

The spread between n-periods dividend strip return and 1-period dividend strip return is given by:

$$S_t^{(n)} = rp_t^{(n)} - rp_t^{(1)} = -\lambda_w \left(A_2(n-1)\right)\sigma_w^2 + \left[-\lambda_x A_1(n-1)\varphi_e\right]\sigma_t^2 \tag{1.10.14}$$

$$S_t = \lim_{n\to\infty} S_t^{(n)} = -\lambda_w \left(A_2(\infty)\right)\sigma_w^2 + \left[-\lambda_x A_1(\infty)\varphi_e\right]\sigma_t^2 \tag{1.10.15}$$

Where

$$A_1(\infty) = \left(\frac{\phi - \frac{1}{\psi}}{1-\rho}\right) \tag{1.10.16}$$

$$A_2(\infty) = \frac{1}{(1-\nu)}\left[(\theta-1)(k_1\nu-1)A_2 + \frac{1}{2}\left(\pi_c-\gamma\right)^2 + \frac{1}{2}\left(\left((\theta-1)k_1A_1 + A_1(\infty)\right)\varphi_e\right)^2 + \frac{1}{2}\varphi_d^2\right] \tag{1.10.17}$$

1.10.3 Convergence of the price dividend ratio

The ratio of the price of the market portfolio on the aggregate dividend at time t can be expressed as:

$$\frac{P_t}{D_t} = \sum_{n=1}^{\infty}\frac{P_t^{(n)}}{D_t} = \sum_{n=1}^{\infty}\exp(A_0(n) + A_1(n)x_t + A_2(n)\sigma_t^2) \tag{1.10.18}$$

The necessary and sufficient conditions for this sum to converge are given below. Once they are satisfied, we can compute the return on the market portfolio as:

$$R_{t+1}^m = \frac{P_{t+1}+D_{t+1}}{P_t} = \frac{(P_{t+1}/D_{t+1})+1}{P_t/D_t}\frac{D_{t+1}}{D_t} \tag{1.10.19}$$

To prove the convergence of the aggregate equity price dividend ratio, I closely follow Lettau and Wachter [2007]. First a necessary (but not sufficient) condition for $pd_t^{(n)}$ to converge for all values of x_t and σ_t^2, is that $A_1(n)$ and $A_2(n)$ approach finite values as $n \to \infty$. We have that $A_1(n)$ converges if and only if

$$| \rho |< 1 \tag{1.10.20}$$

Assuming that 1.10.20 holds, $A_2(n)$ converges if and only if

$$| \nu |< 1 \tag{1.10.21}$$

Given 1.10.20 and 1.10.21, $\lim_{n\to\infty} A_1(n) = A_1(\infty)$ and $\lim_{n\to\infty} A_2(n) = A_2(\infty)$.

Let denote

$$A_0(\infty) = a_{0m} + \mu_d + A_2(\infty)(1-\nu)\bar{\sigma}^2 + \frac{1}{2}\left[\pi_\sigma - \gamma\varphi_\sigma + (\theta-1)k_1 A_2 + A_2(\infty)\right]^2 \sigma_w^2$$

It follows from the recursion in 1.10.6 for $A_0(n)$ that for $n > N$ with N sufficiently large,

$$A_0(n) \approx nA_0(\infty) + constant$$

$$\sum_{n=N}^{\infty} \exp\left(A_0(n) + A_1(n)x_t + A_2(n)\sigma_t^2\right) \approx \exp\left(constant + A_1(\infty)x_t + A_2(\infty)\sigma_t^2\right) \sum_{n=N}^{\infty} \exp\left(nA_0(\infty)\right)$$

Therefore the necessary and sufficient conditions for convergence of the aggregate price dividend ratio are 1.10.20, 1.10.21 and $A_0(\infty) < 0$.

1.10.4 Variance Ratio statistics

The standard definition of the variance ratio (VR1) is given by the ratio of the variance of the sum of cash flow growth up to the horizon τ to the variance of one period cash flow growth :

$$\mathrm{VR}_{1,t}(\tau) = \frac{\frac{1}{\tau}\mathrm{Var}_t\left[\log\left(\frac{D_{t+\tau}}{D_t}\right)\right]}{\mathrm{Var}_t\left[\log\left(\frac{D_{t+1}}{D_t}\right)\right]} \tag{1.10.22}$$

Under the dynamics assumed in 1.3.2-1.3.5, given the conditional normality, 1.6.4 is the conditional variance at time t of the sum of dividend growth over the horizon τ (from $t+1$ to $t+\tau$) divided by τ. The first definition of the variance ratio provides the following analytical formula for the dividend:

$$\mathrm{VR}_{1,t}(\tau)=\frac{1}{\tau}\frac{\left[\left(\pi_c^2+\varphi_d^2\right)\left(\frac{1-\nu^\tau}{1-\nu}\right)+\left(\frac{\phi\varphi_e}{1-\rho}\right)^2\left(\left(\frac{1-\nu^{\tau-1}}{1-\nu}\right)-2\rho\left(\frac{\rho^{\tau-1}-\nu^{\tau-1}}{\rho-\nu}\right)+\rho^2\left(\frac{\rho^{2(\tau-1)}-\nu^{\tau-1}}{\rho^2-\nu}\right)\right)\right]\sigma_t^2+\left[\left(\frac{\phi\varphi_e}{1-\rho}\right)^2\left(\tau-1-\frac{1-\nu^{\tau-1}}{1-\nu}\right)+\left(\pi_c^2+\varphi_d^2\right)\left(\tau-\frac{1-\nu^\tau}{1-\nu}\right)\right]\bar{\sigma}^2+\tau\pi_\sigma^2\sigma_w^2}{\left(\pi_c^2+\varphi_d^2\right)\sigma_t^2+\pi_\sigma^2\sigma_w^2} \tag{1.10.23}$$

Interestingly we can see that at the steady state ($\sigma_t^2=\bar{\sigma}^2$), VR1 equals 1 plus a positive term[30] that goes up and converges to a constant as the horizon increases. So this variance ratio is above 1 and is increasing due to the positive auto-correlation of the expected consumption growth, thus telling us that the risk (the volatility) seen from the present will be higher in the long-run.

The variance ratio statistics defined by 1.6.3 and 1.6.5 are computed using the following intermediary calculus :

[30]This term is given by: $\frac{1}{\tau}\left(\frac{\phi\varphi_e}{1-\rho}\right)^2\left(\tau-1-2\rho\left(\frac{\rho^{\tau-1}-\nu^{\tau-1}}{\rho-\nu}\right)+\rho^2\left(\frac{\rho^{2(\tau-1)}-\nu^{\tau-1}}{\rho^2-\nu}\right)\right)\bar{\sigma}^2/\left(\left(\pi_c^2+\varphi_d^2\right)\bar{\sigma}^2+\pi_\sigma^2\sigma_w^2\right)$

$$
\begin{aligned}
\log \mathrm{E}_t\left[D(\tau)/D(0)\right] &= \log \mathrm{E}_t\left[\exp\left(\sum_{i=1}^{\tau}\Delta d_{t+i}\right)\right]\\
&= \log \mathrm{E}_t \exp \sum_{i=1}^{\tau}\left(\mu_d + \phi x_{t+i-1} + \pi_c s_{t+i} + \pi_x v_{t+i} + \pi_\sigma w_{t+i} + \varphi_d u_{d,t+i}\right)\\
&= \log \mathrm{E}_t \exp \sum_{i=1}^{\tau}\left(\mu_d + \phi \rho^{i-1} x_t + \sum_{r=0}^{i-2}\left[\rho^r \varphi_e v_{t+i-r-1}\right] + \pi_c s_{t+i} + \pi_\sigma w_{t+i} + \varphi_d u_{d,t+i}\right)\\
&= \log \mathrm{E}_t \exp \left(\tau\mu_d + \phi\left[\frac{1-\rho^\tau}{1-\rho}\right]x_t + \sum_{i=1}^{\tau}\sum_{r=0}^{i-2}\left[\rho^r \varphi_e v_{t+i-r-1}\right] + \sum_{i=1}^{\tau}\pi_c s_{t+i} + \sum_{i=1}^{\tau}\pi_\sigma w_{t+i} + \sum_{i=1}^{\tau}\varphi_d u_{d,t+i}\right)\\
&= \tau\mu_d + \phi\left[\frac{1-\rho^\tau}{1-\rho}\right]x_t + \log \mathrm{E}_t \exp\left(\sum_{i=1}^{\tau}\pi_c s_{t+i} + \sum_{i=1}^{\tau}\left(\frac{1-\rho^{\tau-i}}{1-\rho}\right)\varphi_e v_{t+i} + \sum_{i=1}^{\tau}\pi_\sigma w_{t+i} + \sum_{i=1}^{\tau}\varphi_d u_{d,t+i}\right)\\
&= \tau\mu_d + \phi\left[\frac{1-\rho^\tau}{1-\rho}\right]x_t + \log \mathrm{E}_t \exp\left(\sum_{i=1}^{\tau-1}\left[\pi_c s_{t+i} + \left(\frac{1-\rho^{\tau-i}}{1-\rho}\right)\varphi_e v_{t+i}\right.\right.\\
&\left.\left. +\pi_\sigma w_{t+i} + \varphi_d u_{d,t+i}\right) \times \mathrm{E}_{t+\tau-1}\exp\left[\pi_c s_{t+\tau} + \pi_\sigma w_{t+\tau} + \varphi_d u_{d,t+\tau}\right]\right)\\
&= \tau\mu_d + \phi\left[\frac{1-\rho^\tau}{1-\rho}\right]x_t + \log \mathrm{E}_t\left(\exp\sum_{i=1}^{\tau-1}\left[\pi_c s_{t+i} + \left(\frac{1-\rho^{\tau-i}}{1-\rho}\right)\varphi_e v_{t+i}\right.\right.\\
&\left.\left. +\pi_\sigma w_{t+i} + \varphi_d u_{d,t+i}\right] \times \exp\left[\frac{1}{2}H_0\sigma_{t+\tau-1}^2 + \frac{1}{2}G_0\sigma_w^2\right]\right)\\
&= \tau\mu_d + \phi\left[\frac{1-\rho^\tau}{1-\rho}\right]x_t + \log \mathrm{E}_t\left(\exp\sum_{i=1}^{\tau-2}\left[\pi_c s_{t+i} + \left(\frac{1-\rho^{\tau-i}}{1-\rho}\right)\varphi_e v_{t+i}\right.\right.\\
&\left. +\pi_\sigma w_{t+i} + \varphi_d u_{d,t+i}\right] \times \mathrm{E}_{t+\tau-2}\exp\left[\frac{1}{2}H_0\nu\sigma_{t+\tau-2}^2 + \frac{1}{2}G_0\sigma_w^2 + \frac{1}{2}(1-\nu)H_0\bar{\sigma}^2\right.\\
&\left.\left.\pi_c s_{t+\tau-1} + \varphi_e v_{t+\tau-1} + \left(\pi_\sigma + \frac{1}{2}H_0\right)w_{t+\tau-1}\right]\right)\\
&= \tau\mu_d + \phi\left[\frac{1-\rho^\tau}{1-\rho}\right]x_t + \log \mathrm{E}_t\left(\exp\sum_{i=1}^{\tau-2}\left[\pi_c s_{t+i} + \left(\frac{1-\rho^{\tau-i}}{1-\rho}\right)\varphi_e v_{t+i}\right.\right.\\
&\left.\left. +\pi_\sigma w_{t+i} + \varphi_d u_{d,t+i}\right] \times \exp\left[\frac{1}{2}H_1\sigma_{t+\tau-2}^2 + \frac{1}{2}G_1\sigma_w^2 + \frac{1}{2}(1-\nu)\left[H_0+H_1\right]\bar{\sigma}^2\right]\right)\\
&= \tau\mu_d + \phi\left[\frac{1-\rho^\tau}{1-\rho}\right]x_t + \log \mathrm{E}_t\left(\exp\sum_{i=1}^{\tau-3}\left[\pi_c s_{t+i} + \left(\frac{1-\rho^{\tau-i}}{1-\rho}\right)\varphi_e v_{t+i}\right.\right.\\
&\left.\left. +\pi_\sigma w_{t+i} + \varphi_d u_{d,t+i}\right] \times \exp\left[\frac{1}{2}H_2\sigma_{t+\tau-3}^2 + \frac{1}{2}G_2\sigma_w^2 + \frac{1}{2}(1-\nu)\left[H_0+H_1+H_2\right]\bar{\sigma}^2\right]\right)\\
&\vdots\\
&= \tau\mu_d + \phi\left[\frac{1-\rho^\tau}{1-\rho}\right]x_t + \frac{1}{2}H_{\tau-1}\sigma_t^2 + \frac{1}{2}G_{\tau-1}\sigma_w^2 + \frac{1}{2}(1-\nu)\left[\sum_{i=0}^{\tau-2}H_i\right]\bar{\sigma}^2
\end{aligned}
$$

The same way, we obtain that:

$$
\log \mathrm{E}_t\left[D^2(\tau)/D^2(0)\right] = 2\tau\mu_d + 2\phi\left[\frac{1-\rho^\tau}{1-\rho}\right]x_t + 2H_{\tau-1}\sigma_t^2 + 2G_{\tau-1}^{'}\sigma_w^2 + 2(1-\nu)\left[\sum_{i=0}^{\tau-2}H_i\right]\bar{\sigma}^2
$$

Where $H_0 = \varphi_d^2 + \pi_x^2 + \pi_c^2$, $G_0 = G_0^{'} = \pi_\sigma^2$ and $\forall j \in \{1, 2, \ldots\}$,

$$H_j = \nu H_{j-1} + \varphi_d^2 + \left[\left(\frac{1-\rho^j}{1-\rho}\right)\varphi_e\right]^2 + \pi_c^2$$
$$G_j = G_{j-1} + \left[\pi_\sigma + \frac{1}{2}H_{j-1}\right]^2$$
$$G_j^{'} = G_{j-1}^{'} + [\pi_\sigma + H_{j-1}]^2$$

1.10.5 Term structure of interest rate

The Nominal yield curve

In order to derive the term structure of the nominal interest rate from the model, that can be more easily compared to the yield curve observed in the data, I need to specify the dynamics of the inflation rate (π_t) and to compute the nominal log stochastic discount factor ($m_{t+1}^{\$}$) as follows:

$$m_{t+1}^{\$} = m_{t+1} - \pi_{t+1} \tag{1.10.24}$$

Following Augustin and Tedongap [2016] and the references therein, I specify an exogenous dynamics for the inflation rate process similar to the consumption growth process with a time varying mean and volatility.

$$\pi_{t+1} = \mu_\pi + z_t + \nu_\pi \sigma_t \varepsilon_{c,t+1} + \pi_\pi \sigma_w \varepsilon_{w,t+1} + v_t \varepsilon_{\pi,t+1} \tag{1.10.25}$$
$$z_{t+1} = \phi_z z_t + \nu_z(\nu_\pi \sigma_t \varepsilon_{c,t+1} + v_t \varepsilon_{\pi,t+1})$$
$$v_{t+1}^2 = \phi_v v_t^2 + (1-\phi_v)\bar{v} + \nu_v \varepsilon_{v,t+1} \tag{1.10.26}$$
$$(\varepsilon_{\pi,t+1}, \varepsilon_{w,t+1}, \varepsilon_{c,t+1}, \varepsilon_{v,t+1}) \sim N.i.id(0, I)$$

Where$\nu_\pi$captures the effect of consumption growth shock on inflation, while $\pi_\pi$captures the effect of consumption growth volatility (macro uncertainty) shock on inflation. v_{t+1}^2 is the stochastic volatility process driving the variance of the inflation.

Assuming that the nominal price of the bond is an affine function of the state variable:

$$P_{n,t}^{\$} = \exp\left(B_0^{\$}(n) + B_x^{\$}(n)x_t + B_z^{\$}(n)z_t + B_\sigma^{\$}(n)\sigma_t^2 + B_v^{\$}(n)v_t^2\right) \tag{1.10.27}$$

and using the bond pricing equation 1.4.13, we can deduce the coefficient of the affine function as follows:

$$
\begin{aligned}
B_x^\$(n) &= a_{1m} + \rho B_x^\$(n-1) \\
B_z^\$(n) &= -1 + \phi_z B_z^\$(n-1) \\
B_\sigma^\$(n) &= a_{2m} + \nu B_\sigma^\$(n-1) + \frac{1}{2}\left(\lambda_c - \nu_\pi + \nu_z \nu_\pi B_z^\$(n-1)\right)^2 + \frac{1}{2}\left(\lambda_x + \varphi_e B_x^\$(n-1)\right)^2 \\
B_v^\$(n) &= \phi_v B_v^\$(n-1) + \frac{1}{2}\left(-1 + \nu_z B_z^\$(n-1)\right)^2 \\
B_0^\$(n) &= a_{0m} - \mu_\pi + B_0^\$(n-1) + (1-\nu) B_\sigma^\$(n-1)\bar{\sigma}^2 + (1-\phi_v) B_v^\$(n-1)\bar{v} \\
&\quad + \frac{1}{2}\left(\nu_v B_v^\$(n-1)\right)^2 + \frac{1}{2}\left(\lambda_w - \pi_\pi + B_\sigma^\$(n-1)\right)^2 \sigma_w^2
\end{aligned}
\tag{1.10.28}
$$

As the aim of the paper is the term structure of the equity risk premium, we will not pursue the calibration or the estimation of the parameters to match the inflation process and the term structure of the nominal interest rate; this is left for future research. The formulas determining the zero coupon nominal prices in equation (1.10.27) only shows that there is enough flexibility to match the nominal bond term structure without compromising the other achievements of the model.

1.10.6 Recovering the equity term-structure slope using dividend strip data

There are three difficulties that we need to address before implementing our method. Two are related to the frequency and the length of the data. Indeed, the data-set we have contains dividend strip prices from January 1996 to October 2009 at the monthly frequency; while the other data that we used previously in the GMM estimation with constraint are all at the annual frequency and span from 1930 to 2016. To overcome the data availability problem we follow Zhou and Zhu [2015] by using the short sample to run a estimate a reduced form model explaining the variable of interest by the variables for which we have long sample data and then extrapolate the short sample data of the variable of interest. The third difficulty is related to the maturities of the dividend strip available. Indeed, we have dividend strip prices for maturities of 6 months, 12 months, 18 months and 24 months. So we can not compute the one period holding return, instead we will compute the 6-month holding period return using the data and in our model. Let us denote by $F_t^{(i)}$ the price at (the beginning month) t of all the monthly dividends that will occur between t and $t+i$. We know that $F_t^{(i)} = \sum_{k=0}^{i} P_t^{(k)}$, where $P_t^{(k)}$ is the price at t of the k-month dividend strip. Then the 6-month holding period from a dividend contract maturing in n periods is given by:

$$
\begin{aligned}
r_{6,t+6}^{(n)} &= \log\left(\frac{F_{t+6}^{(n-6)} + \sum_{i=1}^{6} D_{t+i}}{F_t^{(n)}}\right) \\
&= \log\left(1 + \frac{F_{t+6}^{(n-6)}}{\sum_{i=0}^{5} D_{t+6-i}}\right) - \log\left(\frac{F_t^{(n)}}{\sum_{i=0}^{5} D_{t-i}}\right) + \log\left(\frac{\sum_{i=0}^{5} D_{t+6-i}}{\sum_{i=0}^{5} D_{t-i}}\right)
\end{aligned}
$$

Let us denote the sum of the 6 month dividends following the date t by $D_t^{0.5y} = \sum_{k=0}^{5} D_{t-k}$, then we can show that:

$$
\log\left(\frac{F_t^{(n)}}{D_t^{0.5y}}\right) = \log\left(\sum_{i=1}^{n} \frac{P_t^{(i)}}{\sum_{i=0}^{5} D_{t-i}}\right) \approx \log\left[\sum_{i=1}^{n} \exp\left(pd_t^{(i)} - \log 6 - \sum_{k=0}^{5}\left(\frac{k+1}{6} - 1\right)\Delta d_{t-j}\right)\right] \tag{1.10.29}
$$

Where $pd_t^{(i)}$ is the log-price dividend ratio of the$i$-period dividend strip[31].

So we can rewrite the expression of the 6 month-holding period return as follows:

$$
r_{6,t+6}^{(n)} = \log D_{t+6}^{0.5y} - \log D_t^{0.5y} + \log\left[1 + \exp\log\left(\frac{F_{t+6}^{(n-6)}}{D_{t+6}^{0.5y}}\right)\right] - \log\left(\frac{F_t^{(n)}}{D_t^{0.5y}}\right) \tag{1.10.30}
$$

There is no closed form solution for $r_{6,t+6}^{(n)}$ even though there are closed form for Δd_{12t} and $pd_t^{(i)}$ because of the log of the sum of exponential appearing in the expression of $\log\left(\frac{F_t^{(n)}}{D_t^{0.5y}}\right)$. So the theoretical counterpart of the 6 month holding period returns moments can only be simulated in our model.

1.10.7 Extensions: Habit formation model

In this section, we show that our main result stating that allowing for positive shocks on uncertainty to negatively affect cash flows helps to explain the decreasing term structure of equity risk premium, is also valid in the Cochrane and Campbell [1999] habit formation model. Indeed as shown in Appendix 1.10.7 when the conditional variance of the consumption growth is stochastic, the short term spread corresponding to the difference between two consecutive period dividend strips can be expressed in this model as a weighted sum of the price of the volatility risk and the price of the consumption growth risk, similar to expression obtained in the LRR model. Furthermore, the price of consumption growth risk is positive and given the negative effect of a positive volatility shock (an increase of uncertainty) on the consumption growth ($\varphi_\sigma < 0$), the price of volatility risk is negative (meaning that investor would like to hedge themselves again

[31] Equation (1.10.64) in 1.10.8 is used for the proof.

volatility risk). So, the sign of the short term spread that defines the slope of the term structure of the equity risk premium can be negative or positive depending on the dominant weighted price. Thus, if the weight on the price of the volatility risk is positive in the short term spread expression in (1.10.41) and it is such that the weighted price of volatility risk dominates the weighted price of consumption growth risk, then the term structure of the equity risk premium will be downward sloping. On the other hand if the weighted price of consumption growth risk dominates then the term structure will be upward sloping. This result is not surprising since from section 1.7.1, we know that allowing cash flows to be more affected during high uncertainty period could enable to obtain a declining term structure of equity risk premium in the simple discounted expected utility model and the Habit model shares the same formulation[32] of the stochastic discount factor with this DEU model with a CRRA utility function.

Following Wachter [2005], we assume that the representative consumer maximizes its life time utility given by:

$$\mathrm{E}\sum_{t=0}^{\infty}\delta^{t}\frac{(C_t - X_t)^{1-\gamma}-1}{1-\gamma} \tag{1.10.31}$$

Where X_t denotes the habit level which is define indirectly through the surplus consumption ratio: $S_t = \frac{C_t - X_t}{C_t}$

We assume the following dynamics for the consumption growth, stochastic volatility and the log-surplus consumption ratio ($s_t = \log(S_t)$):

$$\Delta c_{t+1} = \mu_c + \sigma_t \varepsilon_{c,t+1} + \varphi_\sigma \sigma_w \varepsilon_{w,t+1} \tag{1.10.32}$$

$$\sigma_{t+1}^2 = \nu\sigma_t^2 + (1-\nu)\bar{\sigma}^2 + \sigma_w \varepsilon_{w,t+1} \tag{1.10.33}$$

$$s_{t+1} = (1-\phi)\bar{s} + \phi s_t + \lambda(s_t)(\Delta c_{t+1} - \mathrm{E}_t\Delta c_{t+1}) \tag{1.10.34}$$

Where $\lambda(s_t)$ is the sensitivity function that drives how the innovations in the consumption growth affect the surplus level. It is supposed to be known at time t. For the moment let's

[32] This sdf prices the exposure to the innovations in the consumption growth. The only difference is that in the Habit model, the risk aversion is made state dependent due to the sensitivity function, while it is fixed in the simple DEU model.

assume that $\lambda(s_t) = \lambda$ is a constant[33]. The log-pricing kernel is given by :

$$\begin{aligned} m_{t+1} &= \log(\delta) - \gamma\mu_c + \gamma(1-\phi)(s_t - \bar{s}) - \gamma(1+\lambda)\sigma_t\varepsilon_{c,t+1} - \gamma(1+\lambda)\varphi_\sigma\sigma_w\varepsilon_{w,t+1} \\ &= \log(\delta) - \gamma\mu_c + \gamma(1-\phi)(s_t - \bar{s}) - \lambda_c\sigma_t\varepsilon_{c,t+1} - \lambda_w\sigma_w\varepsilon_{w,t+1} \end{aligned} \quad (1.10.35)$$

where $-\lambda_c = \gamma(1+\lambda)$ and $-\lambda_w = \gamma(1+\lambda)\varphi_\sigma$ are respectively the price of consumption growth risk and volatility risk.

The log-sdf implies the following risk free rate :

$$\begin{aligned} r_{f,t} &= -\mathrm{E}_t(m_{t+1}) - \mathrm{Var}_t(m_{t+1}) \\ &= \log\delta + \gamma\mu_c - \gamma(1-\phi)(s_t - \bar{s}) - \frac{1}{2}\gamma^2(1+\lambda)^2\sigma_t^2 - \frac{1}{2}\gamma^2(1+\lambda)^2\varphi_\sigma^2\sigma_w^2 \end{aligned} \quad (1.10.36)$$

The dividend growth process is defined by:

$$\Delta d_{t+1} = \mu_d + \pi_c\sigma_t\varepsilon_{c,t+1} + \varphi_d\sigma_t\varepsilon_{d,t+1} + \pi_\sigma\sigma_w\varepsilon_{w,t+1} \quad (1.10.37)$$

Assuming that the log-price dividend ratio of a n-period dividend strip is an affine function of the state variables:

$$pd_t^{(n)} = A_0(n) + A_1(n)(s_t - \bar{s}) + A_2(n)\sigma_t^2 \quad (1.10.38)$$

The coefficient of the affine function can be computed recursively thank to the law of one price linking the price at time t of a n-period dividend strip to the price at $t+1$ of a $(n-1)$-period dividend strip. We have $pd_t^{(n)} = \log \mathrm{E}_t\left[\exp\left(m_{t+1} + \Delta d_{t+1} + pd_{t+1}^{(n-1)}\right)\right]$, implying that:

$$A_0(n) = A_0(n-1) + \log(\delta) + \mu_d - \gamma\mu_c + A_2(n-1)(1-\nu)\bar{\sigma}^2 + \frac{1}{2}\left(\pi_\sigma + A_2(n-1) + \left(A_1(n-1) - \gamma(1+\lambda)\right)\varphi_\sigma\right)^2\sigma_w^2$$

$$A_1(n) = A_1(n-1)\phi + \gamma(1-\phi)$$

$$A_2(n) = \nu A_2(n-1) + \frac{1}{2}\left(\varphi_d^2 + (\pi_c + A_1(n-1)\lambda - \gamma(1+\lambda))^2\right) \quad (1.10.39)$$

The log-return on the n-period dividend strip is given by $r_{d,t+1}^{(n)} = \Delta d_{t+1} + pd_{t+1}^{(n-1)} - pd_t^{(n)}$, from which we derive its risk premium as follow:

[33] In order to obtain the risk free rate as a linear function of the log-surplus, we can define the sensitivity function as follow: $\lambda(s_t, \sigma_t^2) = \frac{1}{S_t}\sqrt{1 - 2(s_t - \bar{s})} - 1$ where $S_t = \sqrt{(\sigma_t^2 + \varphi_\sigma^2\sigma_w^2)\frac{\gamma}{1-\phi-\frac{b}{\gamma}}}$ and to be sure that $\lambda(s_t, \sigma_t^2)$ is always defined, we set $\lambda(s_t, \sigma_t^2)$ =0 when $s_t > s_{t,max} = \bar{s} + \frac{1}{2}(1 - S_t^2)$

$$
\begin{aligned}
rp_t^{(n)} &= \mathrm{E}_t(r_{d,t+1}^{(n)} - r_{f,t}) + \frac{1}{2}\mathrm{Var}_t(r_{d,t+1}^{(n)} - r_{f,t}) \\
&= \lambda_w \left(\pi_\sigma + A_2(n-1) + A_1(n-1)\lambda\varphi_\sigma\right)\sigma_w^2 + \lambda_c \left(\pi_c + A_1(n-1)\lambda\right)\sigma_t^2
\end{aligned} \tag{1.10.40}
$$

The short term spread defines as the difference between n-periods dividend strip risk premium and $n-1$-period dividend strip risk premium is given by:

$$
\begin{aligned}
S_{n,t}^{(1)} &= rp_t^{(n)} - rp_t^{(n-1)} \\
&= -\lambda_w \left(\left[A_2(n-1) - A_2(n-2)\right] + \left[A_1(n-1) - A_1(n-2)\right]\lambda\varphi_\sigma\right)\sigma_w^2 - \lambda_c \left(\left[A_1(n-1) - A_1(n-2)\right]\lambda\right)\sigma_t^2
\end{aligned} \tag{1.10.41}
$$

We see that the same reasoning that leads us in the case of the long-run risk model to the conclusion that the term structure of the risk premium can be downward or upward sloping depending on the weighted prices of consumption growth risk and volatility risk also applies here. Indeed, the price of volatility risk is still negative and the price of the consumption growth risk is positive. Thus if the weight on the price of the volatility risk is positive in the short term spread expression in (1.10.41) and it is such that the weighted price of volatility risk dominates the weighted price of consumption growth risk, then the term structure of the equity risk premium will be downward sloping. On the other hand if the weighted price of consumption growth risk dominates then the term structure will be upward sloping.

1.10.8 Temporal aggregation and moment conditions

As done by Bansal et al. [2007b], the h-period aggregated consumption growth rate can be approximated by a weighted average of monthly consumption growth, with the weight taking a $\Lambda - shape$:

$$
\Delta c_{t+1}^a = log\frac{\sum_{j=0}^{h-1} C_{h(t+1)-j}}{\sum_{j=0}^{h-1} C_{ht-j}} \approx \sum_{j=0}^{2h-2} \tau_j \Delta c_{h(t+1)-j}
$$

where the index t is used to count the aggregated time and $h(t-1)+1$ to ht are the corresponding month within the aggregate period t

$$
\tau_j = \frac{j+1}{h} \quad if \quad j < h \qquad and \qquad \tau_j = \frac{2h-j-1}{h} \quad if \quad j \geqslant h
$$

$$
\begin{aligned}
\Delta c^a_{t+1} &= \log \frac{\sum_{j=0}^{h-1} C_{h(t+1)-j}}{\sum_{j=0}^{h-1} C_{ht-j}} \approx \sum_{j=0}^{2h-2} \tau_j \Delta c_{h(t+1)-j} \\
&= h\mu_c + b_{h-1}x_{h(t-1)} + \sum_{j=0}^{h-1} a_j \varphi_e \sigma_{h(t+1)-j-2}\varepsilon_{x,h(t+1)-j-1} + \sum_{j=0}^{h-2} b_j \varphi_e \sigma_{ht-j-2}\varepsilon_{x,ht-j-1} \\
&+ \sum_{j=0}^{2(h-1)} \tau_j \sigma_{h(t+1)-j-1}\varepsilon_{c,h(t+1)-j} + \sum_{j=0}^{2(h-1)} \tau_j \varphi_x \sigma_{h(t+1)-j-1}\varepsilon_{x,h(t+1)-j} \\
&+ \sum_{j=0}^{2(h-1)} \tau_j \varphi_\sigma \sigma_w \varepsilon_{w,h(t+1)-j}
\end{aligned}
$$

Where

$$
\begin{aligned}
a_j &= \sum_{k=0}^{j} \left(\frac{k+1}{h}\right) \rho^{j-k} = \frac{\rho^{j+2} - (j+2)\rho + j + 1}{h(1-\rho)^2} && if\ j < h \\
b_j &= \sum_{k=0}^{j+h} \tau_k \rho^{j+h-k} = \frac{\rho^{j+h+2} - 2\rho^{j+1} + (h-2-j)(1-\rho) + 1}{h(1-\rho)^2} && if\ j \leq h-1
\end{aligned}
$$

Let us denote $v_{h(t+1)-j} = \sigma_{h(t+1)-j-1}\varepsilon_{x,h(t+1)-j}$, $s_{h(t+1)-j} = \sigma_{h(t+1)-j-1}\varepsilon_{c,h(t+1)-j}$ and $w_{h(t+1)-j} = \sigma_w \varepsilon_{w,h(t+1)-j}$.

Then,

$$
\Delta c^a_{t+1} = h\mu_c + b_{h-1}x_{h(t-1)} + \sum_{j=0}^{h} \hat{a}_j v_{h(t+1)-j} + \sum_{j=0}^{h-2} \hat{b}_j v_{ht-j-1} + \sum_{j=0}^{h} c_j s_{h(t+1)-j} + \sum_{j=0}^{h-2} d_j s_{ht-j-1} \tag{1.10.42}
$$

$$
+ \sum_{j=0}^{h} f_j w_{h(t+1)-j} + \sum_{j=0}^{h-2} g_j w_{ht-j-1}
$$

Where

$$
\begin{cases}
\hat{a}_j = \varphi_e a_{j-1} & c_j = \tau_j & f_j = \varphi_\sigma \tau_j & 0 \leq j \leq h \\
\hat{b}_j = \varphi_e b_j & d_j = \tau_{h+j+1} & g_j = \varphi_\sigma \tau_{h+j+1} & 0 \leq j \leq h-2
\end{cases}
$$

Therefore,

$$
\mathrm{E}(\Delta c^a_{t+1}) = \mathrm{E}(\sum_{j=0}^{2(h-1)} \tau_j \mu_c) = h\mu_c \tag{1.10.43}
$$

$$
\begin{aligned}
\mathrm{Var}(\Delta c^a_{t+1}) &= b^2_{h-1}\mathrm{Var}\left(x_{h(t-1)}\right) + \left(\sum_{j=0}^{h} \hat{a}^2_j + \sum_{j=0}^{h-2} \hat{b}^2_j\right) \varphi^2_e \bar{\sigma}^2 + \left(\sum_{j=0}^{h} c^2_j + \sum_{j=0}^{h-2} d^2_j\right) \bar{\sigma}^2 \\
&+ \left(\sum_{j=0}^{h} f^2_j + \sum_{j=0}^{h-2} g^2_j\right) \sigma^2_w
\end{aligned} \tag{1.10.44}
$$

$$
\begin{aligned}
\mathrm{Cov}(\Delta c_t^a, \Delta c_{t+1}^a) = \mathrm{Cov}\Bigg(& h\mu_c + b_{h-1}x_{h(t-2)} + \sum_{j=0}^{h} \hat{a}_j v_{ht-j} + \sum_{j=0}^{h-2} \hat{b}_j v_{h(t-1)-j-1} + \sum_{j=0}^{h} c_j s_{ht-j} \\
& + \sum_{j=0}^{h-2} d_j s_{h(t-1)-j-1} + \sum_{j=0}^{h} f_j w_{ht-j} + \sum_{j=0}^{h-2} g_j w_{h(t-1)-j-1}, h\mu_c + b_{h-1}x_{h(t-1)} \\
& + \sum_{j=0}^{h} \hat{a}_j v_{h(t+1)-j} + \sum_{j=0}^{h-2} \hat{b}_j v_{ht-j-1} + \sum_{j=0}^{h} c_j s_{h(t+1)-j} + \sum_{j=0}^{h-2} d_j s_{ht-j-1} \\
& + \sum_{j=0}^{h} f_j w_{h(t+1)-j} + \sum_{j=0}^{h-2} g_j w_{ht-j-1} \Bigg) \\
& = \rho^h b_{h-1}^2 \mathrm{Var}\left(x_{h(t-2)}\right) + \mathrm{Cov}\Bigg(\sum_{j=0}^{h} \hat{a}_j v_{ht-j} + \sum_{j=0}^{h-2} \hat{b}_j v_{h(t-1)-j-1} + \sum_{j=0}^{h} c_j s_{ht-j} \\
& + \sum_{j=0}^{h-2} d_j s_{h(t-1)-j-1} + \sum_{j=0}^{h} f_j w_{ht-j} + \sum_{j=0}^{h-2} g_j w_{h(t-1)-j-1}, \sum_{j=0}^{h} \hat{a}_j v_{h(t+1)-j} \\
& + b_{h-1}\left(\sum_{k=0}^{h-1} \rho^k \left(\varphi_e v_{h(t-1)-k}\right)\right) + \sum_{j=0}^{h-2} \hat{b}_j v_{ht-j-1} \\
& + \sum_{j=0}^{h} c_j s_{h(t+1)-j} + \sum_{j=0}^{h-2} d_j s_{ht-j-1} + \sum_{j=0}^{h} f_j w_{h(t+1)-j} + \sum_{j=0}^{h-2} g_j w_{ht-j-1} \Bigg)
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Cov}(\Delta c_t^a, \Delta c_{t+1}^a) = & \rho^h b_{h-1}^2 \mathrm{Var}\left(x_{h(t-2)}\right) + \sum_{j=0}^{h-2} \left(\hat{a}_{j+1} + \rho b_{h-2}\rho^{j+1}\varphi_e\right) \hat{b}_j \bar{\sigma}^2 + \\
& + \sum_{j=0}^{h-2} \left(c_{j+1}\right) d_j \bar{\sigma}^2 + \sum_{j=0}^{h-2} \left(f_{j+1}\right) g_j \sigma_w^2 \\
& + \hat{a}_h \left(\hat{a}_0 + b_{h-1}\varphi_e\right) \bar{\sigma}^2 + c_h \left(c_0\right) \bar{\sigma}^2 + f_h \left(f_0\right) \sigma_w^2
\end{aligned} \tag{1.10.45}
$$

Where $\mathrm{Var}(x_t) = \frac{\varphi_e^2 \bar{\sigma}^2}{1-\rho^2}$

$$
\begin{aligned}
\mathrm{E}\left[\left(\Delta c_{t+1}^a - \mathrm{E}(\Delta c_{t+1}^a)\right)^3\right] = & 3\mathrm{E}\left\{ \left[\left(\sum_{j=0}^{h} \hat{a}_j v_{h(t+1)-j} + \sum_{j=0}^{h-2} \hat{b}_j v_{ht-j-1} \right)^2 + \left(\sum_{j=0}^{h} c_j s_{h(t+1)-j} + \sum_{j=0}^{h-2} d_j s_{ht-j-1} \right)^2 \right] \right. \\
& \left. \times \left(\sum_{j=0}^{h} f_j w_{h(t+1)-j} + \sum_{j=0}^{h-2} g_j w_{ht-j-1} \right) \right\} \\
& = 3\left(\sum_{j=0}^{h-1} \sum_{k=j+1}^{h} \left(\hat{a}_j^2 + c_j^2\right) f_k \nu^{k-j-1} + \sum_{j=0}^{h} \sum_{k=0}^{h-2} \left(\hat{a}_j^2 + c_j^2\right) g_k \nu^{h-j+k} + \sum_{j=0}^{h-3} \sum_{k=j+1}^{h-2} \left(\hat{b}_j^2 + d_j^2\right) g_k \nu^{k-j-1} \right) \sigma_w^2
\end{aligned}
$$

- **Annual dividend growth**

Let us denote $u_{h(t+1)-j} = \sigma_{h(t+1)-j-1}\varepsilon_{d,h(t+1)-j}$. Using the same formulation as for the consumption growth rate, the annual dividend growth rate can be expressed in term of monthly dividend growth rates as follow:

$$\Delta d^a_{t+1} = log\frac{\sum_{j=0}^{h-1} D_{h(t+1)-j}}{\sum_{j=0}^{h-1} D_{ht-j}} \approx \sum_{j=0}^{2(h-1)} \tau_j \Delta d_{h(t+1)-j}$$
$$= h\mu_d + \phi b_{h-1} x_{h(t-1)} + \phi \sum_{j=0}^{h-1} a_j \varphi_e v_{h(t+1)-j-1} + \phi \sum_{j=0}^{h-2} b_j \varphi_e v_{ht-j-1} + \pi_c \sum_{j=0}^{2(h-1)} \tau_j s_{h(t+1)-j}$$
$$+ \pi_x \sum_{j=0}^{2(h-1)} \tau_j v_{h(t+1)-j} + \pi_\sigma \sum_{j=0}^{2(h-1)} \tau_j w_{h(t+1)-j} + \sum_{j=0}^{2(h-1)} \tau_j \varphi_d u_{h(t+1)-j}$$

$$\begin{aligned}\Delta d^a_{t+1} = {} & h\mu_d + \phi b_{h-1} x_{h(t-1)} + \sum_{j=0}^{h} \left[\phi a_{j-1} \varphi_e\right] v_{h(t+1)-j} \\ & + \sum_{j=0}^{h-2} \left[\phi b_j \varphi_e\right] v_{ht-j-1} + \sum_{j=0}^{h} \left[\pi_c \tau_j\right] s_{h(t+1)-j} + \sum_{j=0}^{h-2} \left[\pi_c \tau_{h+j+1}\right] s_{ht-j-1} \\ & + \sum_{j=0}^{h} \left[\pi_\sigma \tau_j\right] w_{h(t+1)-j} + \sum_{j=0}^{h-2} \left[\pi_\sigma \tau_{h+j+1}\right] w_{ht-j-1} + \sum_{j=0}^{2(h-1)} \tau_j \varphi_d u_{h(t+1)-j}\end{aligned} \tag{1.10.46}$$

Therefore,

$$\mathrm{E}(\Delta d^a_{t+1}) = h\mu_d \tag{1.10.47}$$

$$\begin{aligned}\mathrm{Var}(\Delta d^a_{t+1}) = {} & \left[\phi b_{h-1}\right]^2 \mathrm{Var}\left(x_{h(t-1)}\right) + \sum_{j=0}^{h} \left[\phi a_{j-1} \varphi_e\right]^2 \bar{\sigma}^2 + \sum_{j=0}^{h} \left[\pi_c \tau_j\right]^2 \bar{\sigma}^2 \\ & + \sum_{j=0}^{h} \left[\pi_\sigma \tau_j\right]^2 \sigma_w^2 + \sum_{j=0}^{h-2} \left[\phi b_j \varphi_e\right]^2 \bar{\sigma}^2 + \sum_{j=0}^{h-2} \left[\pi_c \tau_{h+j+1}\right]^2 \bar{\sigma}^2 + \sum_{j=0}^{h-2} \left[\pi_\sigma \tau_{h+j+1}\right]^2 \sigma_w^2 \\ & + \left(\sum_{j=0}^{2(h-1)} \tau_j^2\right) \varphi_d^2 \bar{\sigma}^2\end{aligned} \tag{1.10.48}$$

$$\begin{aligned}\mathrm{Cov}(\Delta d^a_t, \Delta d^a_{t+1}) = \mathrm{Cov}\Bigg(& h\mu_d + \phi b_{h-1} x_{h(t-2)} + \sum_{j=0}^{h} \left[\phi a_{j-1} \varphi_e\right] v_{ht-j} \\ & + \sum_{j=0}^{h-2} \left[\phi b_j \varphi_e\right] v_{h(t-1)-j-1} + \sum_{j=0}^{h} \left[\pi_c \tau_j\right] s_{ht-j} + \sum_{j=0}^{h-2} \left[\pi_c \tau_{h+j+1}\right] s_{h(t-1)-j-1} \\ & + \sum_{j=0}^{h} \left[\pi_\sigma \tau_j\right] w_{ht-j} + \sum_{j=0}^{h-2} \left[\pi_\sigma \tau_{h+j+1}\right] w_{h(t-1)-j-1} + \sum_{j=0}^{2(h-1)} \tau_j \varphi_d u_{ht-j}, h\mu_d \\ & + \phi b_{h-1} x_{h(t-1)} + \sum_{j=0}^{h} \left[\phi a_{j-1} \varphi_e\right] v_{h(t+1)-j} + \sum_{j=0}^{h-2} \left[\phi b_j \varphi_e\right] v_{ht-j-1} \\ & + \sum_{j=0}^{h} \left[\pi_c \tau_j\right] s_{h(t+1)-j} + \sum_{j=0}^{h-2} \left[\pi_c \tau_{h+j+1}\right] s_{ht-j-1} + \sum_{j=0}^{h} \left[\pi_\sigma \tau_j\right] w_{h(t+1)-j} \\ & + \sum_{j=0}^{h-2} \left[\pi_\sigma \tau_{h+j+1}\right] w_{ht-j-1} + \sum_{j=0}^{2(h-1)} \tau_j \varphi_d u_{h(t+1)-j} \Bigg)\end{aligned}$$

$$\operatorname{Cov}(\Delta d_t^a, \Delta d_{t+1}^a) = \rho^h \left[\phi b_{h-1}\right]^2 \operatorname{Var}\left(x_{h(t-2)}\right) + \sum_{j=0}^{h} \left[\phi a_{j-1} \varphi_e\right] \left[\phi b_{j-1} \varphi_e\right] \bar{\sigma}^2 \tag{1.10.49}$$

$$+ \sum_{j=0}^{h} \left[\pi_c \tau_j\right] \left[\pi_c \tau_{h+j}\right] \bar{\sigma}^2 + \sum_{j=0}^{h} \left[\pi_\sigma \tau_j\right] \left[\pi_\sigma \tau_{h+j}\right] \sigma_w^2$$

$$+ \sum_{j=0}^{h-2} \left[\phi b_j \varphi_e\right] \left[\rho^{j+1} \phi b_{h-1} \varphi_e\right] \bar{\sigma}^2 + \varphi_d^2 \left(\sum_{j=0}^{h-2} \tau_j \tau_{h+j}\right) \bar{\sigma}^2$$

$$\operatorname{E}\left[\left(\Delta d_{t+1}^a - \operatorname{E}(\Delta d_{t+1}^a)\right)^3\right] = 3\operatorname{E}\left\{\left[\left(\sum_{j=0}^{h} \left[\phi a_{j-1} \varphi_e\right] v_{h(t+1)-j} + \sum_{j=0}^{h-2} \left[\phi b_j \varphi_e\right] v_{ht-j-1}\right)^2 + \left(\sum_{j=0}^{h} \left[\pi_c \tau_j\right] s_{h(t+1)-j} + \sum_{j=0}^{h-2} \left[\pi_c \tau_{h+j+1}\right] s_{ht-j-1}\right)^2\right.\right.$$

$$\left.\left.\left(\sum_{j=0}^{2(h-1)} \tau_j \varphi_d u_{h(t+1)-j}\right)^2\right] \times \left(\sum_{j=0}^{h} \left[\pi_\sigma \tau_j\right] w_{h(t+1)-j} + \sum_{j=0}^{h-2} \left[\pi_\sigma \tau_{h+j+1}\right] w_{ht-j-1}\right)\right\}$$

$$= 3\operatorname{E}\left\{\left[\left(\sum_{j=0}^{h} \left[\phi a_{j-1} \varphi_e\right]^2 v_{h(t+1)-j}^2 + \sum_{j=0}^{h-2} \left[\phi b_j \varphi_e\right]^2 v_{ht-j-1}^2\right) + \left(\sum_{j=0}^{h} \left[\pi_c \tau_j\right]^2 s_{h(t+1)-j}^2 + \sum_{j=0}^{h-2} \left[\pi_c \tau_{h+j+1}\right]^2 s_{ht-j-1}^2\right.\right.\right.$$

$$\left.\left.\left(\sum_{j=0}^{2(h-1)} \left[\tau_j \varphi_d\right]^2 u_{h(t+1)-j}^2\right)\right] \times \left(\sum_{j=0}^{h} \left[\pi_\sigma \tau_j\right] w_{h(t+1)-j} + \sum_{j=0}^{h-2} \left[\pi_\sigma \tau_{h+j+1}\right] w_{ht-j-1}\right)\right\}$$

$$= 3\left\{\sum_{j=0}^{h-1} \sum_{k=j+1}^{h} \left(\left[\phi a_{j-1} \varphi_e\right]^2 + \left[\pi_c \tau_j\right]^2 + \left[\tau_j \varphi_d\right]^2\right) \left[\pi_\sigma \tau_k\right] \nu^{k-j-1}\right.$$

$$+ \sum_{j=0}^{h} \sum_{k=0}^{h-2} \left(\left[\phi a_{j-1} \varphi_e\right]^2 + \left[\pi_c \tau_j\right]^2 + \left[\tau_j \varphi_d\right]^2\right) \left[\pi_\sigma \tau_{h+k+1}\right] \nu^{h+k-j}$$

$$\left. + \sum_{j=0}^{h-3} \sum_{k=j+1}^{h-2} \left(\left[\phi b_j \varphi_e\right]^2 + \left[\pi_c \tau_{h+j+1}\right]^2 + \left[\varphi_d \tau_{h+j+1}\right]^2\right) \left[\pi_\sigma \tau_{h+k+1}\right] \nu^{k-j-1}\right\} \sigma_w^2$$

$$\operatorname{Cov}(\Delta c_{t+1}^a, \Delta d_{t+1}^a) = \operatorname{Cov}\left(h\mu_c + b_{h-1} x_{h(t-1)} + \sum_{j=0}^{h} \hat{a}_j v_{h(t+1)-j} + \sum_{j=0}^{h-2} \hat{b}_j v_{ht-j-1} + \sum_{j=0}^{h} c_j s_{h(t+1)-j}\right.$$

$$+ \sum_{j=0}^{h-2} d_j s_{ht-j-1} + \sum_{j=0}^{h} f_j w_{h(t+1)-j} + \sum_{j=0}^{h-2} g_j w_{ht-j-1}, h\mu_d + \phi b_{h-1} x_{h(t-1)}$$

$$+ \sum_{j=0}^{h} \left[\phi a_{j-1} \varphi_e\right] v_{h(t+1)-j} + \sum_{j=0}^{h-2} \left[\phi b_j \varphi_e\right] v_{ht-j-1} + \sum_{j=0}^{h} \left[\pi_c \tau_j\right] s_{h(t+1)-j}$$

$$+ \sum_{j=0}^{h-2} \left[\pi_c \tau_{h+j+1}\right] s_{ht-j-1} + \sum_{j=0}^{h} \left[\pi_\sigma \tau_j\right] w_{h(t+1)-j} + \sum_{j=0}^{h-2} \left[\pi_\sigma \tau_{h+j+1}\right] w_{ht-j-1}$$

$$\left. + \sum_{j=0}^{2(h-1)} \tau_j \varphi_d u_{h(t+1)-j}\right)$$

$$\mathrm{Cov}(\Delta c^a_{t+1}, \Delta d^a_{t+1}) = \phi b^2_{h-1}\mathrm{Var}\left(x_{h(t-2)}\right) + \sum_{j=0}^{h}\left[\phi a_{j-1}\varphi_e\right]\hat{a}_j\bar{\sigma}^2 + \sum_{j=0}^{h}\left[\pi_c\tau_j\right]c_j\bar{\sigma}^2 \tag{1.10.50}$$

$$+ \sum_{j=0}^{h}\left[\pi_\sigma\tau_j\right]f_j\sigma^2_w + \sum_{j=0}^{h-2}\left[\pi_c\tau_{h+j+1}\right]d_j\bar{\sigma}^2 + \sum_{j=0}^{h-2}\left[\phi b_j\varphi_e\right]\hat{b}_j\bar{\sigma}^2$$

$$+ \sum_{j=0}^{h-2}\left[\pi_\sigma\tau_{h+j+1}\right]g_j\sigma^2_w$$

- **Annual market return**

Let's denote

$$\Gamma_0 = k_{0m} + (k_{1m}-1)A_{0m} + k_{1m}A_{2m}(1-\nu)\bar{\sigma}^2 + \mu_d \tag{1.10.51}$$

$$\Gamma_{1m} = A_{1m}(k_{1m}\rho - 1) + \phi; \quad \Gamma_{2m} = (k_{1m}\nu - 1)A_{2m}$$

$$\beta_{m,c} = \pi_c;\ \beta_{m,x} = k_{1m}A_{1m}\varphi_e;\ \beta_{m,w} = \pi_\sigma + k_{1m}A_{2m};\ \beta_{m,d} = \varphi_d$$

The monthly return on the market portfolio is given by:

$$\begin{aligned} r_{m,t+1} &= k_{0m} + k_{1m}z_{m,t+1} - z_{m,t} + \Delta d_{t+1} \\ &= \Gamma_{0m} + \Gamma_{1m}x_t + \Gamma_{2m}\sigma^2_t + \beta_{m,c}s_{t+1} + \beta_{m,x}v_{t+1} + \beta_{m,w}w_{t+1} + \beta_{m,d}u_{t+1} \end{aligned}$$

So the aggregate return on the market portfolio can be obtained as:
$\forall j \geq 1$,

$$\begin{aligned} r^a_{m,t+j} &= \sum_{k=0}^{h-1} r_{m,h(t+j)-k} = \sum_{k=0}^{h-1}\Big[\Gamma_{0m} + \Gamma_{1m}x_{h(t+j)-k-1} + \Gamma_{2m}\sigma^2_{h(t+j)-k-1} \\ &+\beta_{m,c}s_{h(t+j)-k} + \beta_{m,x}v_{h(t+j)-k} + \beta_{m,w}w_{h(t+j)-k} + \beta_{m,d}u_{h(t+j)-k}\Big] \\ &= h\Gamma_{0m} + \Gamma_{1m}\left[\sum_{k=0}^{h-1}\rho^{h(j+1)-k-1}x_{h(t-1)} + \sum_{k=0}^{h-1}\sum_{r=0}^{h(j+1)-2-k}\varphi_e\rho^r v_{h(t+j)-k-r-1}\right] \\ &+ \Gamma_{2m}\left[\sum_{k=0}^{h-1}(1-\nu)\left(\sum_{r=0}^{h(j+1)-2-k}\nu^r\right)\bar{\sigma}^2 + \sum_{k=0}^{h-1}\nu^{h(j+1)-1-k}\sigma^2_{h(t-1)}\right. \\ &\left.+\sum_{k=0}^{h-1}\sum_{r=0}^{h(j+1)-2-k}\nu^r w_{h(t+j)-k-r-1}\right)\Bigg] + \sum_{k=0}^{h-1}\beta_{m,c}s_{h(t+j)-k} \\ &+ \sum_{k=0}^{h-1}\beta_{m,x}v_{h(t+j)-k} + \sum_{k=0}^{h-1}\beta_{m,w}w_{h(t+j)-k} + \sum_{k=0}^{h-1}\beta_{m,d}u_{h(t+j)-k} \end{aligned}$$

$$r^a_{m,t+j} = h\Gamma_{0m} + \Gamma_{2m}\left(h - \nu^{hj}\left(\frac{1-\nu^h}{1-\nu}\right)\right)\bar{\sigma}^2 + \Gamma_{1m}\rho^{hj}\left(\frac{1-\rho^h}{1-\rho}\right)x_{h(t-1)} + \Gamma_{2m}\nu^{hj}\left(\frac{1-\nu^h}{1-\nu}\right)\sigma^2_{h(t-1)} \tag{1.10.52}$$

$$+ \sum_{k=0}^{h-1}\left[\beta_{m,x} + \varphi_e\Gamma_{1m}\left(\frac{1-\rho^k}{1-\rho}\right)\right]v_{h(t+j)-k} + \varphi_e\Gamma_{1m}\sum_{r=0}^{hj-1}\rho^r\left(\frac{1-\rho^h}{1-\rho}\right)v_{h(t+j-1)-r}$$
$$+ \sum_{k=0}^{h-1}\left[\beta_{m,w} + \Gamma_{2m}\left(\frac{1-\nu^k}{1-\nu}\right)\right]w_{h(t+j)-k} + \Gamma_{2m}\sum_{r=0}^{hj-1}\nu^r\left(\frac{1-\nu^h}{1-\nu}\right)w_{h(t+j-1)-r}$$
$$+ \sum_{k=0}^{h-1}\beta_{m,c}s_{h(t+j)-k} + \sum_{k=0}^{h-1}\beta_{m,d}u_{h(t+j)-k} \tag{1.10.53}$$

And

$$r^a_{m,t} = h\Gamma_{0m} + \Gamma_{2m}\left(h - \left(\frac{1-\nu^h}{1-\nu}\right)\right)\bar{\sigma}^2 + \Gamma_{1m}\left(\frac{1-\rho^h}{1-\rho}\right)x_{h(t-1)} + \Gamma_{2m}\left(\frac{1-\nu^h}{1-\nu}\right)\sigma^2_{h(t-1)} \tag{1.10.54}$$
$$+ \sum_{k=0}^{h-1}\beta_{m,c}s_{ht-k} + \sum_{k=0}^{h-1}\left[\beta_{m,x} + \varphi_e\Gamma_{1m}\left(\frac{1-\rho^k}{1-\rho}\right)\right]v_{ht-k} + \sum_{k=0}^{h-1}\beta_{m,d}u_{h(t+j)-k}$$
$$+ \sum_{k=0}^{h-1}\left[\beta_{m,w} + \Gamma_{2m}\left(\frac{1-\nu^k}{1-\nu}\right)\right]w_{ht-k}$$

Therefore,

$$\mathrm{E}(r^a_{m,t}) = h\Gamma_0 + h\Gamma_{2m}\sigma^2 \tag{1.10.55}$$

$$\mathrm{Var}(r^a_{m,t}) = \left(\Gamma_{1m}\left(\frac{1-\rho^h}{1-\rho}\right)\right)^2\mathrm{Var}(x_t) + \left(\Gamma_{2m}\left(\frac{1-\nu^h}{1-\nu}\right)\right)^2\mathrm{Var}(\sigma^2_t) + h\beta^2_{m,d}\bar{\sigma}^2 \tag{1.10.56}$$
$$+ 2\Gamma_{1m}\Gamma_{2m}\left(\frac{1-\rho^h}{1-\rho}\right)\left(\frac{1-\nu^h}{1-\nu}\right)\mathrm{Cov}(x_t,\sigma^2_t) + \sum_{k=0}^{h-1}\beta^2_{m,c}\bar{\sigma}^2$$
$$+ \sum_{k=0}^{h-1}\left[\beta_{m,x} + \varphi_e\Gamma_{1m}\left(\frac{1-\rho^k}{1-\rho}\right)\right]^2\bar{\sigma}^2 + \sum_{k=0}^{h-1}\left[\beta_{m,w} + \Gamma_{2m}\left(\frac{1-\nu^k}{1-\nu}\right)\right]^2\sigma^2_w$$

Where $\mathrm{Var}(\sigma^2_t) = \frac{\sigma^2_w}{1-\nu^2}$ and $\mathrm{Cov}(x_t, \sigma^2_t) = 0$

$$\mathrm{Cov}(r^a_{m,t}, r^a_{m,t+1}) = \rho^h\left[\Gamma_{1m}\left(\frac{1-\rho^h}{1-\rho}\right)\right]^2\mathrm{Var}(x_t) + \nu^h\left[\Gamma_{2m}\left(\frac{1-\nu^h}{1-\nu}\right)\right]^2\mathrm{Var}(\sigma^2_t) \tag{1.10.57}$$
$$+ \sum_{k=0}^{h-1}\left[\beta_{m,x} + \varphi_e\Gamma_{1m}\left(\frac{1-\rho^k}{1-\rho}\right)\right]\varphi_e\Gamma_{1m}\left(\frac{1-\rho^h}{1-\rho}\right)\rho^k\bar{\sigma}^2$$
$$+ \sum_{k=0}^{h-1}\left[\beta_{m,w} + \Gamma_{2m}\left(\frac{1-\nu^k}{1-\nu}\right)\right]\left[\Gamma_{2m}\left(\frac{1-\nu^h}{1-\nu}\right)\nu^k\right]\sigma^2_w$$
$$= \rho\left(\frac{\varphi^2_e\bar{\sigma}^2}{1-\rho^2}\right)\left(\frac{1-\rho^h}{1-\rho}\right)^2\Gamma^2_{1m} + \nu\left(\frac{\sigma^2_w}{1-\nu^2}\right)\left(\frac{1-\nu^h}{1-\nu}\right)^2\Gamma^2_{2m} \tag{1.10.58}$$
$$+ \beta_{m,x}\Gamma_{1m}\varphi_e\bar{\sigma}^2\left(\frac{1-\rho^h}{1-\rho}\right)^2 + \beta_{m,w}\Gamma_{2m}\sigma_w\left(\frac{1-\nu^h}{1-\nu}\right)^2$$

- **Annual risk free rate**

$\forall j \geq 1$,

$$
\begin{aligned}
r^a_{f,t+j} &= \sum_{r=0}^{h-1} r_{f,h(t+j)-r} = \sum_{r=0}^{h-1} \left(A_{0f} + A_{1f} x_{h(t+j)-r} + A_{2f}\sigma^2_{h(t+j)-r}\right) \\
&= hA_{0f} + \sum_{r=0}^{h-1} A_{1f} \left(\rho^{h(j+1)-r} x_{h(t-1)} + \sum_{k=0}^{h(j+1)-r-1} \rho^k \varphi_e v_{h(t+j)-r-k} \right. \\
&\quad \left. + A_{2f} \left(\nu^{h(j+1)-r} \sigma^2_{h(t-1)} + \sum_{k=0}^{h(j+1)-r-1} \left(\nu^k w_{h(t+j)-r-k} + (1-\nu)\bar{\sigma}^2 \nu^k \right) \right) \right)
\end{aligned}
$$

$$
\begin{aligned}
r^a_{f,t+j} &= hA_{0f} + A_{2f} \left[h - \nu^{hj+1} \left(\frac{1-\nu^h}{1-\nu} \right) \right] \bar{\sigma}^2 + A_{1f}\rho^{hj+1} \left(\frac{1-\rho^h}{1-\rho} \right) x_{h(t-1)} + A_{2f}\nu^{hj+1} \left(\frac{1-\nu^h}{1-\nu} \right) \sigma^2_{h(t-1)} \qquad (1.10.59) \\
&+ A_{1f}\varphi_e \left[\sum_{r=0}^{h-1} \left(\frac{1-\rho^{r+1}}{1-\rho} \right) v_{h(t+j)-r} + \sum_{r=0}^{hj-1} \rho^{r+1} \left(\frac{1-\rho^h}{1-\rho} \right) v_{h(t+j-1)-r} \right] \\
&+ A_{2f} \left[\sum_{r=0}^{h-1} \left(\frac{1-\nu^{r+1}}{1-\nu} \right) w_{h(t+j)-r} + \sum_{r=0}^{hj-1} \nu^{r+1} \left(\frac{1-\nu^h}{1-\nu} \right) w_{h(t+j-1)-r} \right]
\end{aligned}
$$

and

$$
\begin{aligned}
r^a_{f,t} &= hA_{0f} + A_{2f} \left[h - \nu \left(\frac{1-\nu^h}{1-\nu} \right) \right] \bar{\sigma}^2 + A_{1f}\rho \left(\frac{1-\rho^h}{1-\rho} \right) x_{h(t-1)} + A_{2f}\nu \left(\frac{1-\nu^h}{1-\nu} \right) \sigma^2_{h(t-1)} \qquad (1.10.60) \\
&+ A_{1f}\varphi_e \sum_{r=0}^{h-1} \left(\frac{1-\rho^{r+1}}{1-\rho} \right) v_{ht-r} + \sum_{r=0}^{h-1} \left(A_{2f} \left(\frac{1-\nu^{r+1}}{1-\nu} \right) \right) w_{ht-r}
\end{aligned}
$$

Therefore,

$$
\mathrm{E}(r^a_{f,t}) = hA_{0f} + hA_{2f} \qquad (1.10.61)
$$

$$
\begin{aligned}
\mathrm{Var}(r^a_{f,t}) &= \left[A_{1f}\rho \left(\frac{1-\rho^h}{1-\rho} \right) \right]^2 \mathrm{Var}(x_t) + \left[A_{2f}\nu \left(\frac{1-\nu^h}{1-\nu} \right) \right]^2 \mathrm{Var}(\sigma^2_t) \qquad (1.10.62) \\
&+ \sum_{r=0}^{h-1} \left[A_{1f}\varphi_e \left(\frac{1-\rho^{r+1}}{1-\rho} \right) \right]^2 \bar{\sigma}^2 + \sum_{r=0}^{h-1} \left[A_{2f} \left(\frac{1-\nu^{r+1}}{1-\nu} \right) \right]^2 \sigma^2_w
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Cov}(r^a_{f,t}, r^a_{f,t+1}) &= \rho^h \left[A_{1f}\rho \left(\frac{1-\rho^h}{1-\rho} \right) \right]^2 \mathrm{Var}(x_t) + \nu^h \left[A_{2f}\nu \left(\frac{1-\nu^h}{1-\nu} \right) \right]^2 \mathrm{Var}(\sigma^2_t) \qquad (1.10.63) \\
&+ (A_{1f}\varphi_e)^2 \sum_{r=0}^{h-1} \left[\rho^{r+1} \left(\frac{1-\rho^{r+1}}{1-\rho} \right) \left(\frac{1-\rho^h}{1-\rho} \right) \right]^2 \bar{\sigma}^2 \\
&+ \sum_{r=0}^{h-1} \left[\left(\frac{1-\nu^{r+1}}{1-\nu} \right) A_{2f} \right] \left[A_{2f}\nu^{r+1} \left(\frac{1-\nu^h}{1-\nu} \right) \right] \sigma^2_w
\end{aligned}
$$

- **Forward annual excess return**

Using 1.10.52 and 1.10.59, the aggregate excess return on the market portfolio can be expressed as follow $j \geq 1$:

$$
\begin{aligned}
r^a_{ex,t+j} &= r^a_{m,t+j} - r^a_{f,t+j} \\
&= h\left(\Gamma_{0m} - A_{0f}\right) + h\left(\Gamma_{2m} - A_{2f}\right)\bar{\sigma}^2 - \left(\Gamma_{2m} - \nu A_{2f}\right)\left(\nu^{hj}\left(\frac{1-\nu^h}{1-\nu}\right)\right)\bar{\sigma}^2 \\
&+ \left(\Gamma_{1m} - \rho A_{1f}\right)\rho^{hj}\left(\frac{1-\rho^h}{1-\rho}\right)x_{h(t-1)} + \left(\Gamma_{2m} - \nu A_{2f}\right)\nu^{hj}\left(\frac{1-\nu^h}{1-\nu}\right)\sigma^2_{h(t-1)} \\
&+ \sum_{k=0}^{h-1}\beta_{m,c}s_{h(t+j)-k} + \sum_{k=0}^{h-1}\left[\beta_{m,x} + \varphi_e\left(\Gamma_{1m}\left(\frac{1-\rho^k}{1-\rho}\right) - A_{1f}\left(\frac{1-\rho^{k+1}}{1-\rho}\right)\right)\right]v_{h(t+j)-k} \\
&+ \varphi_e\left(\frac{1-\rho^h}{1-\rho}\right)\sum_{r=0}^{hj-1}\rho^r\left[\Gamma_{1m} - A_{1f}\rho\right]v_{h(t+j-1)-r} + \left(\frac{1-\nu^h}{1-\nu}\right)\sum_{k=0}^{hj-1}\nu^k\left[\Gamma_{2m} - A_{2f}\nu\right]w_{h(t+j-1)-k} \\
&+ \sum_{k=0}^{h-1}\beta_{m,d}u_{h(t+j)-k} + \sum_{k=0}^{h-1}\left[\beta_{m,w} + \Gamma_{2m}\left(\frac{1-\nu^k}{1-\nu}\right) - \left(A_{2f}\left(\frac{1-\nu^{k+1}}{1-\nu}\right)\right)\right]w_{h(t+j)-k}
\end{aligned}
$$

In particular,

$$
\begin{aligned}
r^a_{ex,t} &= r^a_{m,t} - r^a_{f,t} \\
&= h\left(\Gamma_{0m} - A_{0f}\right) + h\left(\Gamma_{2m} - A_{2f}\right) - \left(\Gamma_{2m} - \nu A_{2f}\right)\left(\left(\frac{1-\nu^h}{1-\nu}\right)\right)\bar{\sigma}^2 \\
&+ \left(\Gamma_{1m} - \rho A_{1f}\right)\left(\frac{1-\rho^h}{1-\rho}\right)x_{h(t-1)} + \left(\Gamma_{2m} - \nu A_{2f}\right)\left(\frac{1-\nu^h}{1-\nu}\right)\sigma^2_{h(t-1)} \\
&+ \sum_{k=0}^{h-1}\beta_{m,c}s_{ht-k} + \sum_{k=0}^{h-1}\left[\beta_{m,x} + \varphi_e\left(\Gamma_{1m}\left(\frac{1-\rho^k}{1-\rho}\right) - A_{1f}\left(\frac{1-\rho^{k+1}}{1-\rho}\right)\right)\right]v_{ht-k} \\
&+ \sum_{k=0}^{h-1}\beta_{m,d}u_{ht-k} + \sum_{k=0}^{h-1}\left[\beta_{m,w} + \Gamma_{2m}\left(\frac{1-\nu^k}{1-\nu}\right) - \left(A_{2f}\left(\frac{1-\nu^{k+1}}{1-\nu}\right)\right)\right]w_{ht-k}
\end{aligned}
$$

- **Annual price-dividend ratio**

The aggregate log price dividend ratio at time t is the logarithm of the ratio of the price at the end of the period divided by the sum of monthly dividends over the aggregation period.

$$
\begin{aligned}
p_t^a - d_t^a &= \log P_{ht} - \log \sum_{j=0}^{h-1} D_{ht-j} \\
&= \log P_{ht} - \log D_{ht} + \sum_{j=0}^{h-1} (\log D_{ht-j} - \log D_{ht-j-1}) - \log h - \sum_{j=0}^{h-1} \frac{j+1}{h} \triangle d_{ht-j} \\
&= z_{m,ht} + \sum_{j=0}^{h-1} \triangle d_{ht-j} - \log h - \sum_{j=0}^{h-1} \frac{j+1}{h} \triangle d_{ht-j} \\
&= z_{m,ht} - \log h - \sum_{j=0}^{h-1} \left(\frac{j+1}{h} - 1 \right) \triangle d_{ht-j} \qquad (1.10.64)
\end{aligned}
$$

$$
\begin{aligned}
p_t^a - d_t^a &= A_{0m} + A_{1m} x_{ht} + A_{2m} \sigma_{ht}^2 - \log h - \sum_{j=0}^{h-1} \left(\frac{j+1}{h} - 1 \right) (\mu_d + \phi x_{ht-j-1} + \pi_c s_{ht-j} \\
&+ \varphi_d u_{ht-j} + \pi_\sigma w_{ht-j}) \\
&= \left[A_{0m} - \log h - \frac{1}{2}(h-1)\mu_d + A_{2m}(1-\nu^h)\bar{\sigma}^2 \right] + \left[A_{1m}\rho^h - \sum_{j=0}^{h-1} \left(\frac{j+1}{h} - 1 \right) \phi \rho^{h-j-1} \right] x_{h(t-1)} \\
&+ A_{2m}\nu^h \sigma_{h(t-1)}^2 + \sum_{j=0}^{h-1} \left[\left(A_{1m}\rho^j - \phi a_{j-1}^{'} \right) \varphi_e \right] v_{ht-j} + \sum_{j=0}^{h-1} \left[- \left(\frac{j+1}{h} - 1 \right) \pi_c \right] s_{ht-j} - \\
&\sum_{j=0}^{h-1} \left(\frac{j+1}{h} - 1 \right) \varphi_d u_{ht-j} + \sum_{j=0}^{h-1} \left[A_{2m}\nu^j - \left(\frac{j+1}{h} - 1 \right) \pi_\sigma \right] w_{ht-j}
\end{aligned}
$$

and

$$
\begin{aligned}
p_{t+1}^a - d_{t+1}^a &= A_{0m} + A_{1m} x_{h(t+1)} + A_{2m} \sigma_{h(t+1)}^2 - \log h - \sum_{j=0}^{h-1} \left(\frac{j+1}{h} - 1 \right) (\mu_d + \phi x_{h(t+1)-j-1} \\
&+ \pi_c s_{h(t+1)-j} + \varphi_d u_{h(t+1)-j} + \pi_\sigma w_{h(t+1)-j}) \\
&= \left[A_{0m} - \log h - \frac{1}{2}(h-1)\mu_d + A_{2m}(1-\nu^{2h})\bar{\sigma}^2 \right] + \rho^h \left[A_{1m}\rho^h - \phi a_{h-1}^{'} \right] x_{h(t-1)} + A_{2m}\nu^{2h}\sigma_{h(t-1)}^2 \\
&+ \sum_{j=0}^{h-1} \rho^j \left(A_{1m}\rho^h - \phi a_{h-1}^{'} \right) \varphi_e v_{ht-j} + \sum_{j=0}^{h-1} \left[\left(A_{1m}\rho^j - \phi a_{j-1}^{'} \right) \varphi_e \right] v_{h(t+1)-j} \\
&+ \sum_{j=0}^{h-1} + \sum_{j=0}^{h-1} \left[- \left(\frac{j+1}{h} - 1 \right) \pi_c \right] s_{h(t+1)-j} + \sum_{j=0}^{h-1} \left[A_{2m}\nu^{h+j} \right] w_{ht-j} - \sum_{j=0}^{h-1} \left(\frac{j+1}{h} - 1 \right) \varphi_d u_{ht-j} \\
&+ \sum_{j=0}^{h-1} \left[A_{2m}\nu^j - \left(\frac{j+1}{h} - 1 \right) \pi_\sigma \right] w_{h(t+1)-j}
\end{aligned}
$$

Where

$$
a_j^{'} = a_j - \left(\frac{1-\rho^{(j+1)}}{1-\rho} \right) \qquad for \quad j \in \{-1, 0, ..., h-1\}
$$

Therefore,

$$
\mathrm{E}(p_t^a - d_t^a) = A_{0m} - \log h - \frac{1}{2}(h-1)\mu_d + A_{2m}(1-\nu^h)\bar{\sigma}^2 \qquad (1.10.65)
$$

$$\mathrm{Var}(p_t^a - d_t^a) = \left[A_{1m}\rho^h - \phi a_{h-1}^{'}\right]^2 \mathrm{Var}(x_t) + \left[A_{2m}\nu^h\right]^2 \mathrm{Var}(\sigma_t^2) + 2\left[A_{1m}\rho^h - \phi a_{h-1}^{'}\right]\left[A_{2m}\nu^h\right] \mathrm{Cov}(x_t, \sigma_t^2) \tag{1.10.66}$$

$$+ \sum_{j=0}^{h-1}\left[\left(A_{1m}\rho^j - \phi a_{j-1}^{'}\right)\varphi_e\right]^2 \bar{\sigma}^2 + \sum_{j=0}^{h-1}\left[-\left(\frac{j+1}{h} - 1\right)\pi_c\right]^2 \bar{\sigma}^2$$

$$+ \sum_{j=0}^{h-1}\left[A_{2m}\nu^j - \left(\frac{j+1}{h} - 1\right)\pi_\sigma\right]^2 \sigma_w^2 + \sum_{j=0}^{h-1}\left(\frac{j+1}{h} - 1\right)^2 \varphi_d^2\bar{\sigma}^2$$

$$\mathrm{Cov}(p_t^a - d_t^a, p_{t+1}^a - d_{t+1}^a) = \rho^h\left[A_{1m}\rho^h - \phi a_{h-1}^{'}\right]^2 \mathrm{Var}(x_t) + \left(\rho^h + \nu^h\right)\left[A_{1m}\rho^h - \phi a_{h-1}^{'}\right]\left[A_{2m}\nu^h\right] \mathrm{Cov}(x_t, \sigma_t^2) \tag{1.10.67}$$

$$+ \nu^h\left[A_{2m}\nu^h\right]^2 \mathrm{Var}(\sigma_t^2) + \sum_{j=0}^{h-1}\rho^j\left[\left(A_{1m}\rho^j - \phi a_{j-1}^{'}\right)\varphi_e\right]\left(A_{1m}\rho^h - \phi a_{h-1}^{'}\right)\varphi_e\bar{\sigma}^2$$

$$+ \sum_{j=0}^{h-1}\left[A_{2m}\nu^j - \left(\frac{j+1}{h} - 1\right)\pi_\sigma\right]\left[A_{2m}\nu^{h+j}\right]\sigma_w^2$$

1.10.9 Theoretical moments for the predictive regression

- **Prediction of future excess return by the log price-dividend ratio**

$\forall j \geq 1$,

$$\mathrm{Cov}(r_{ex,t+j}^a, p_t^a - d_t^a) = (\Gamma_{1m} - \rho A_{1f})\,\rho^{hj}\left(\frac{1-\rho^h}{1-\rho}\right)\left[A_{1m}\rho^h - \phi a_{h-1}^{'}\right]\mathrm{Var}(x_t)$$

$$+ (\Gamma_{2m} - \nu A_{2f})\,\nu^{h(j+1)}\left(\frac{1-\nu^h}{1-\nu}\right)A_{2m}\mathrm{Var}(\sigma_t^2)$$

$$+ \sum_{r=0}^{h-1}\rho^{r+h(j-1)}\left(\frac{1-\rho^h}{1-\rho}\right)\left[\Gamma_{1m} - A_{1f}\rho\right]\left(A_{1m}\rho^j - \phi a_{j-1}^{'}\right)\varphi_e^2\bar{\sigma}^2$$

$$+ \sum_{r=0}^{h-1}(\Gamma_{2m} - \nu A_{2f})\,\nu^{r+h(j-1)}\left(\frac{1-\nu^h}{1-\nu}\right)\left[A_{2m}\nu^r - \left(\frac{r+1}{h} - 1\right)\pi_\sigma\right]\sigma_w^2$$

- Useful transformations

$$x_{h(t+j)-k} = \rho^{h(j+1)-k}x_{h(t-1)} + \sum_{r=0}^{h(j+1)-k-1}\left(\rho^r\varphi_e v_{h(t+j)-k-r}\right) \tag{1.10.68}$$

$$\sum_{r=0}^{h-1}\sum_{k=0}^{h(j+1)-r-1}\rho^k s_{h(t+j)-r-k} = \sum_{r=0}^{h-1}\left(\frac{1-\rho^{r+1}}{1-\rho}\right)s_{h(t+j)-r} + \sum_{r=0}^{hj-1}\rho^{r+1}\left(\frac{1-\rho^h}{1-\rho}\right)s_{h(t+j-1)-r} \tag{1.10.69}$$

1.10.10 Confidence Intervals of the GMM estimates

The standard approach applied in GMM to build the confidence intervals of the estimates is the Delta method. It is based on the fact that under some regularity conditions, the GMM estimator is asymptotically normally distributed. Among the regularity conditions, we need that (i) the GMM estimator converges to the true value and (ii) the jacobian matrix of the moment conditions with respect to the vector of parameters is full rank. But in a weak identification context, the estimator might not converge to the true value and second the jacobian matrix is (or close to) rank deficient. So applying the Delta method might be misleading. Fortunatelly there are some methods to build confidence intervals that are robust to the weak identification. One of them is the Anderson-Rubin projection method Dufour [2003], Stock and Wright [2000]. The method relies on the convergence of the GMM objective function (even in the weak identification case) to a chi-squared distribution with the number of moment conditions as degree of freedom. The (1-α)% confidence set will then consist of collecting all the parameter's vectors for which the value of the GMM objective function will be lower than the (1-α)% quantile of the chi-squared distribution with the number of moment conditions as degree of freedom. The confidence intervals for each parameter are then obtained by projecting the confidence set on the corresponding axe.

We apply a modification of the projection method to build the confidence intervals in our estimation. The modification has been made to take into account the slow convergence in distribution of the GMM objective function to its Chi-squared limit[34]. We simulate a finite sample distribution of the GMM objective function and we use the quantile of the simulated distribution (instead of the asymptotic Chi-squared distribution) for the projecting method. The simulation was made by parametric boostrap; the procedure consisted of generating simulated samples with the same size as the observed data and to re-estimate the vector of parameters on the simulated samples. At the end of the procedure, we obtain the distributions of the parameters estimates and a distribution of the GMM objective function.

[34] We run some simulations showing that our GMM objective function converges to the chi-squared distribution but very slowly (the sample size need to be huge to get the distribution of the GMM objective function closed to the corresponding chi-squared asymptotic distribution). See figure 1.2.

2 Long Run Risk Model with Regime Shifts

The Long Run Risk model provided by Bansal and Yaron [2004] has been very successful in solving many asset pricing puzzles and matching the macro finance data. The key idea in the model is that the representative consumer might be aware and is afraid of the presence of a slow moving component in the consumption growth whose shocks have a long lasting impact on its well being. Since the dividends paid on the financial market positively load on that expected consumption component, it is risky for the consumer to hold financial assets because their value might drop when its consumption is at the bottom and its marginal utility is high.

During normal times, the LRR model is also able to explain the term spread, meaning the difference between the dividend yield or the return on zero coupon equities with different maturities. Indeed, the dividend yield is an increasing function of the volatility and the slope coefficient increases as the maturity of the zero coupon equity increases, such that long term zero coupon equity prices drop more during high volatility period compared to short term zero coupon equity prices. So the dividend yields display an increasing term structure during normal times. But during bad times as it occurred in the 2008 financial crises, there is a change in the slope of the yield curve which now becomes negative; the prices of short term assets decrease more relative to the prices of long term assets and the yield curve slopes down. The standard LRR model is not able capture the negative slope of the yield curve (van Binsbergen et al. [2012], van Binsbergen and Koijen [2016]) but as shown by Tinang [2018] an extension of the LRR that accounts for the direct negative effect of a positive shock of consumption volatility on the cash flows can explain the downward slope of the yield curve. Finally, combining the standard LRR with its extension in a regime switching model could be the best way to capture the two developments in the yield curve slope that appear on figure 2.2. The regime switching models have widely been used to capture various situations in the development of the bond yield curves. They can be used to capture stochastic drifts and /or volatilites, to incorporate business cycles, to study contagion effects or to evaluate the effects of monetary policies Gourieroux et al. [2014]. But less emphasize has been put on the modeling of equities for which the pay-offs are random.

We develop a regimes shift consumption based asset pricing model with a representative agent having recursive preferences. The economy is subject to the same conditional Gaussian risk factors available in the standard LRR model such as the short-run consumption growth risk, the long run consumption growth risk, the dividend growth risk and the uncertainty risk. On top of those usual risks, it now faces the regime shift risk driven by an homogeneous Markov switching process. The agent is assumed to have full information about the risk factors that are priced in each state and also to know the transition probabilities of moving from one state to another. Each regime is characterized by a specific cash flow dynamics process. In the calibration of the model, we restrict the consumption growth process to be the same in all regimes and only the dividend dynamics changes with regimes. The "Downward Sloping term structure" regime is characterized by a dividend growth process which is negatively correlated with the volatility but not directly correlated to consumption growth idiosyncratic shock[1]. The "Upward Sloping term structure" regime is characterized by a dividend growth not directly affected by volatility shocks but which loads on the expected consumption growth and is positively correlated to short run consumption growth shocks. By mixing the asset pricing properties of these two regimes through the transition probability matrix, the designed model can deliver various shapes for the term structure of equity returns such as downward sloping, hump shape or upward sloping term structures.

The remaining of the chapter is organized as follows: Section 1 presents the model, its solution and the asset pricing formulas for bonds and equities. Section 2 presents the calibration of the model with various transition probability matrices and look at its implications for the term structure of equity returns. Section 3 concludes the chapter.

2.1 The Model

Following Bansal and Yaron [2004], we assume a representative consumer with Epstein and Zin [1989] recursive utility function choosing each period how much to consume and to invest on financial market in order to maximize its lifetime expected utility. Its utility function is given by:

$$V_t = \left[(1-\delta)C_t^{1-\frac{1}{\psi}} + \delta\left(E_t\left(V_{t+1}^{1-\gamma}\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\right)\right]^{\frac{1}{1-\frac{1}{\psi}}} \tag{2.1.1}$$

[1]The dividend growth is still correlated to consumption growth but only through the consumption growth volatility

Where δ is the pure discount factor, γ is the relative risk aversion coefficient and ψ is the Elasticity of Inter-temporal Substitution (EIS). This preference specification allows to disentangle the EIS from the risk aversion coefficient and to break the tight link imposed between them by the time additive preference where $\gamma = \frac{1}{\psi}$.

There are N regimess in the economy displayed each period t by a state variable $s_t \in \{1, ..., N\}$. The cash flows dynamics follow the specification provided by Tinang [2018] but the coefficients are regime dependent :

$$\Delta c_{t+1} = \mu_{c,s_t} + x_t + \sigma_t \varepsilon_{c,t+1} + \varphi_{\sigma,s_t} \sigma_{w,s_t} \varepsilon_{w,t+1} \tag{2.1.2}$$

$$x_{t+1} = \rho_{s_t} x_t + \varphi_{e,s_t} \sigma_t \varepsilon_{x,t+1} \tag{2.1.3}$$

$$\sigma^2_{t+1} = \nu_{s_t} \sigma^2_t + (1 - \nu_{s_t}) \bar{\sigma}^2_{s_t} + \sigma_{w,s_t} \varepsilon_{w,t+1} \tag{2.1.4}$$

$$\Delta d_{t+1} = \mu_{d,s_t} + \phi_{s_t} x_t + \pi_{c,s_t} \sigma_t \varepsilon_{c,t+1} + \pi_{\sigma,s_t} \sigma_{w,s_t} \varepsilon_{w,t+1} + \varphi_{d,s_t} \sigma_t \varepsilon_{d,t+1} \tag{2.1.5}$$

$$(\varepsilon_{c,t+1}, \varepsilon_{x,t+1}, \varepsilon_{d,t+1}, \varepsilon_{w,t+1}) \sim N.i.id(0, I)$$

There are three state variables in this specification : s_t which governs the regime, x_t which drives the long run consumption risk and σ_t which represents the time varying uncertainty faced by the consumer. Equation 2.1.2 describes the consumption growth ($\Delta c_{t+1} = \log C_{t+1} - \log C_t$) at time $t+1$, which depends on the expected consumption growth at time t, the short run consumption growth shock $\varepsilon_{c,t+1}$, the volatility shock $\varepsilon_{w,t+1}$. The coefficients of the consumption growth process ($\mu_{c,s_{t+1}}, \varphi_{\sigma,s_{t+1}}$ and $\sigma_{w,s_{t+1}}$) depend on the regime at time t, which is assumed to be known by the consumer and possibly unknown to the econometrician. Equation 2.1.3 represents the dynamics of the expected consumption growth which evolves like AR(1) process with heteroskedastic error term. The auto-regressive coefficient ρ_{s_t} and the variance scaling parameter φ_{e,s_t} are regime dependent. Equation 2.1.4 describes the stochastic volatility[2] driving the time varying uncertainty in the model, which also evolves like a AR(1) process with heteroskedastic error term and regime dependent coefficients. Equation 2.1.5 presents the dividend growth process which depends on the expected consumption growth, the short run consumption growth shock, the volatility shock and the idiosyncratic dividend growth shock. We assume that the coefficients in the dividend growth dynamics are also regime dependent.

We also assume that the regime follows an homogeneous Markov switching process with a

[2] In this specification, the squared of the volatility can go negative since the errors are normally distributed and thus unbounded. But the coefficients are chosen such that this will happen in very few cases and assuming a truncated normal distribution for the error term does not change the qualitative implications of the model.

transition matrix given by:

$$P = (p_{ij})_{i,j=1,\dots,N}$$

where $p_{ij} = P(s_{t+1=j}|s_t = i)$, $\forall i = 1..N, \sum_{j=1}^{N} p_{ij} = 1$. Furthermore, the regime process (s_t) is independent from all the other shocks in the model.

Let us denote by $R_{c,t+1} = \frac{W_{t+1}+C_{t+1}}{W_t}$ the gross return on the wealth portfolio that delivers the stream of consumption goods as dividends and $r_{c,t+1} = \log(R_{c,t+1})$ the log-return. The log stochastic discount factor in this model can be expressed as follows:

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

We assume that the logarithm of the wealth portfolio price to consumption ratio is an affine function of the states variables:

$$z_t = A_{0,s_t} + A_{1,s_t} x_t + A_{2,s_t}\sigma_t^2$$

Using the log-linear approximation of the returns : $r_{c,t+1} = \Delta c_{t+1} + k_{0,s_{t+1}} + k_{1,s_{t+1}} z_{t+1} - z_t$ where $k_{0,s_{t+1}}$and $k_{1,s_{t+1}}$are the states specific log-linearization constants given by

$$\begin{cases} k_{0,s_{t+1}} = & \log(1 + e^{\bar{z}_{s_{t+1}}}) - \bar{z}_{s_{t+1}} k_{1,s_{t+1}} \\ k_{1,s_{t+1}} = & e^{\bar{z}_{s_{t+1}}}/(1 + e^{\bar{z}_{s_{t+1}}}) \end{cases}$$

. The log-linearization can also be done around the mean of the price-dividend ratio but the approximation would be less precised compared to a state specific log linearization around the mean value of the log price dividend ratio in the state. The log stochastic discount factor can be expressed as follows:

$$\begin{aligned} m_{t+1} = {} & \theta \log \delta + (\theta - \frac{\theta}{\psi} - 1)\mu_{c,s_t} + (\theta - 1) k_{0,s_{t+1}} + (\theta - 1)(k_{1,s_{t+1}} A_{0,s_{t+1}} - A_{0,s_t}) \\ & + (\theta - 1) k_{1,s_{t+1}} A_{2,s_{t+1}} (1 - \nu_{s_t}) \bar{\sigma}_{s_t}^2 + \left[-\gamma + (\theta - 1)(k_{1,s_{t+1}} \rho_{s_t} A_{1,s_{t+1}} - A_{1,s_t})\right] x_t \quad (2.1.6) \\ & + (\theta - 1)(k_{1,s_{t+1}} A_{2,s_{t+1}} \nu_{s_t} - A_{2,s_t})\sigma_t^2 + \left[-\gamma \varphi_{\sigma,s_t} + (\theta - 1) k_{1,s_{t+1}} A_{2,s_{t+1}}\right] \sigma_{w,s_t} \varepsilon_{w,t+1} \\ & + (\theta - 1)(k_{1,s_{t+1}} A_{1,s_{t+1}} \varphi_{e,s_t} \sigma_t \varepsilon_{x,t+1} - \gamma \sigma_t \eta_{t+1} \end{aligned}$$

The one period sdf from time t to $t+1$ depends on the coefficients driving the dynamics of the cash flows at t and $t+1$; and compared to the one regime model where those two dynamics are forcefully the same, in the regime switching model, they can be different depend on the state prevailing at those dates. The sdf in the regime switching model will have a higher variability compared to a one regime model with the same sources of risk. Thus the Hansen-Jaganathan bound could be higher in the regime switching model, which means a higher chance to solve the equity premium puzzle but at the same time the risk free rate could be too much volatile.

We first proceed with the pricing of the wealth portfolio. We have :

$$
\begin{aligned}
m_{t+1} + r_{c,t+1} = \; & \theta \log \delta + (1-\gamma)\mu_{c,s_t} + \theta k_{0,s_{t+1}} + \theta(k_{1,s_{t+1}} A_{0,s_{t+1}} - A_{0,s_t}) \\
& + \theta k_{1,s_{t+1}} A_{2,s_{t+1}} \left(1-\nu_{s_t}\right) \bar{\sigma}_{s_t}^2 + \left[(1-\gamma) + \theta(k_{1,s_{t+1}} \rho_{s_t} A_{1,s_{t+1}} - A_{1,s_t})\right] x_t \\
& + \theta \left(k_{1,s_{t+1}} A_{2,s_{t+1}} \nu_{s_t} - A_{2,s_t}\right) \sigma_t^2 + \left[(1-\gamma)\,\varphi_{\sigma,s_t} + \theta k_{1,s_{t+1}} A_{2,s_{t+1}}\right] \sigma_{w,s_t} \varepsilon_{w,t+1} \\
& + \theta k_{1,s_{t+1}} A_{1,s_{t+1}} \varphi_{e,s_t} \sigma_t e_{t+1} + (1-\gamma)\,\sigma_t \eta_{t+1}
\end{aligned}
$$

We use the Euler equation $\mathrm{E}_t \left[\exp\left(m_{t+1} + r_{c,t+1}\right)\right] = 1$, the law of iterated expectation and the log-linear approximation $\exp(y) - 1 \approx y$ to determine the unknown coefficients A_{0,s_t},$A_{1,s_t}$and A_{2,s_t} as follows :

$\forall i \in \{1, ..., N\}$

$$
\begin{aligned}
\Rightarrow & \sum_{s_{t+1}=1}^{N} p_{is_{t+1}} \left((1-\gamma) + \theta(k_{1,s_{t+1}} \rho_i A_{1,s_{t+1}} - A_{1,i})\right) = 0 \\
& \sum_{s_{t+1}=1}^{N} p_{is_{t+1}} \left(\theta \left(k_{1,s_{t+1}} A_{2,s_{t+1}} \nu_i - A_{2,i}\right) + \frac{1}{2} \left(\theta k_{1,s_{t+1}} A_{1,s_{t+1}} \varphi_{e,i}\right)^2 + \frac{1}{2} \left(1-\gamma\right)^2 \right) = 0 \\
& \sum_{s_{t+1}=1}^{N} p_{is_{t+1}} \left\{\theta \log \delta + (1-\gamma)\mu_{c,i} + \theta k_{0,s_{t+1}} + \theta(k_{1,s_{t+1}} A_{0,s_{t+1}} - A_{0,i}) + \theta k_{1,s_{t+1}} A_{2,s_{t+1}} \left(1-\nu_i\right) \bar{\sigma}_i^2 \right. \\
& \left. + \frac{1}{2} \left[(1-\gamma)\,\varphi_{\sigma,i} + \theta k_{1,s_{t+1}} A_{2,s_{t+1}}\right]^2 \sigma_{w,i}^2 \right\} = 0
\end{aligned}
$$

$$\Rightarrow A_1 = \left[I_N - \left(\rho 1_N^{'}\right) \odot PD\right]^{-1} (1 - \frac{1}{\psi})1_N$$

$$A_2 = \frac{1}{2}\theta \left[I_N - \left(\nu 1_N^{'}\right) \odot PD\right]^{-1} \left(P(1 - \frac{1}{\psi})1_N \odot (1 - \frac{1}{\psi})1_N + \left(\left(\varphi_e 1_N^{'}\right) \odot PD\right) \odot \left(\left(\varphi_e 1_N^{'}\right) \odot PD\right)\right)$$

$$A_0 = [I_N - PD]^{-1} \left(\log \delta 1_N + (1 - \frac{1}{\psi})\mu_c + Pk_0 + \left((1-\nu) \odot \bar{\sigma}^2 1_N^{'}\right) \odot PD\right.$$
$$+ (1-\gamma)\left(\sigma_w \odot \varphi_\sigma \odot \sigma_w 1_N^{'}\right) \odot PDA_2 + \frac{\theta}{2}\left(\sigma_w 1_N^{'} \odot PDA_2\right) \odot \left(\sigma_w 1_N^{'} \odot PDA_2\right)$$
$$\left. + \frac{1}{2}(1-\gamma)\left(1 - \frac{1}{\psi}\right)(\varphi_\sigma \odot \sigma_w) \odot (\varphi_\sigma \odot \sigma_w)\right)$$

where $\odot$ denotes the element by element product of vectors,$A_i = (A_{i,1}, ..., A_{i,N}) \in \mathbb{R}^N$ for $i = 0, 1, 2$. $k_0, k_1, \rho, \nu \in \mathbb{R}^N$, $D = Diag\,(k_{1,1}, ..., k_{1,N}) = Diag\,(k_1) \in \mathbb{R}^N \times \mathbb{R}^N$, $1_N = (1...1)^{'}$, $\nu 1_N^{'} \in \mathbb{R}^N \times \mathbb{R}^N$, $\sigma_w 1_N^{'} \in \mathbb{R}^N \times \mathbb{R}^N$, $\rho 1_N^{'} \in \mathbb{R}^N \times \mathbb{R}^N$ and $\varphi_e 1_N^{'} \in \mathbb{R}^N \times \mathbb{R}^N$.

The coefficients $(A_{i,j})$,$i = 0, 1, 2$; $j = 1, ..., N$ are linear combinations of the "standard" coefficients obtained in the special case of a unique regime. Indeed, if we restrict the transition matrix to identity and we assume that the values of the cash flows'parameters remain the same in different states, then we recover the well known coefficient formulas in the one regime model.

The return on the risk free asset is given by:

$$\begin{aligned}
r_{f,t} &= -\log \mathrm{E}\left(\exp(m_{t+1})|s_t = i\right) \\
&= -\log \sum_{j=1...N} P_{ij} \mathrm{E} \exp\left(\theta \log \delta - \gamma \mu_{c,j} + (\theta - 1)\, k_{1j} A_{0j} + (\theta - 1)\, k_{1j} A_{2j} \left(1 - \nu_j\right) \bar{\sigma}_j^2 + (1-\theta)\, A_{0i}\right. \\
&+ \left[-\gamma + (\theta - 1)\left(k_{1j}\rho_j A_{1j} - A_{1i}\right)\right] x_t + \left((\theta - 1)\left[k_{1j} A_{2j}\nu_j - A_{2i}\right] + 0.5\left[\gamma^2 + ((\theta-1)k_{1j}A_{1j}\varphi_{ej})^2\right]\right)\sigma_t^2 \\
&\left. +0.5\left[-\gamma\varphi_{\sigma j} + (\theta - 1)\, k_{1j} A_{2j}\right]^2 \sigma_{wj}^2\right) \\
&= -\log\left\{P \mathrm{E} \exp\left(\theta \log \delta - \gamma \odot \mu_c + (\theta - 1)\,(D - 1_{N\times N}) A_0 + (\theta - 1)\, D A_2 \odot (1 - \nu) \odot \bar{\sigma}^2\right.\right. \\
&+ 0.5\left[-\gamma\varphi_\sigma + (\theta - 1)\, D^\nu A_2\right]^2 \odot \sigma_w^2 + \left[-\gamma + (\theta - 1)\left(D^\rho - 1_{N\times N}\right) A_1\right] x_t \\
&\left.\left. + \left((\theta - 1)\,(D^\nu - 1_{N\times N}) A_2 + 0.5\left(\gamma^2 + [(\theta - 1) D A_1 \odot \varphi_e]^2\right)\right)\sigma_t^2\right)\right\}
\end{aligned}$$

For the market portfolio, we use the Euler equation $\mathrm{E}_t\left[\exp\left(m_{t+1} + r_{m,t+1}\right)\right] = 1$, the law of iterated expectation and the log-linear approximation $\exp(y) - 1 \approx y$ to determine the unknown coefficients A_{0m,s_t},$A_{1m,s_t}$and A_{2m,s_t} as follows :

$$\begin{aligned}
m_{t+1}+r_{m,t+1} = {} & \theta\log\delta+\mu_{d,s_{t+1}}-\gamma\mu_{c,s_{t+1}}+(\theta-1)k_{0,s_{t+1}}+(\theta-1)(k_{1,s_{t+1}}A_{0,s_{t+1}}-A_{0,s_t}) \\
& +k_{0m,s_{t+1}}+k_{1m,s_{t+1}}A_{0m,s_{t+1}}-A_{0m,s_t}+k_{1m,s_{t+1}}A_{2m,s_{t+1}}\left(1-\nu_{s_{t+1}}\right)\bar{\sigma}^2_{s_{t+1}} \\
& +(\theta-1)k_{1,s_{t+1}}A_{2,s_{t+1}}\left(1-\nu_{s_{t+1}}\right)\bar{\sigma}^2_{s_{t+1}}+ \\
& +\left[\left(\phi_{s_{t+1}}-\gamma\right)+\rho_{s_{t+1}}k_{1m,s_{t+1}}A_{1m,s_{t+1}}-A_{1m,s_t}+(\theta-1)(k_{1,s_{t+1}}\rho_{s_{t+1}}A_{1,s_{t+1}}-A_{1,s_t})\right]x_t \\
& +\left[k_{1m,s_{t+1}}A_{2m,s_{t+1}}\nu_{s_{t+1}}-A_{2m,s_t}+(\theta-1)\left(k_{1,s_{t+1}}A_{2,s_{t+1}}\nu_{s_{t+1}}-A_{2,s_t}\right)\right]\sigma_t^2+ \\
& +\left[\pi_{\sigma,s_{t+1}}-\gamma\varphi_{\sigma,s_{t+1}}+k_{1m,s_{t+1}}A_{2m,s_{t+1}}+(\theta-1)k_{1,s_{t+1}}A_{2,s_{t+1}}\right]\sigma_{w,s_{t+1}}\varepsilon_{w,t+1} \\
& +\left[(\theta-1)k_{1,s_{t+1}}A_{1,s_{t+1}}+k_{1m,s_{t+1}}A_{1m,s_{t+1}}\right]\varphi_{e,s_{t+1}}\sigma_t\varepsilon_{x,t+1}+\left(\pi_{c,s_{t+1}}-\gamma\right)\sigma_t\varepsilon_{c,t+1} \\
& +\varphi_{d,s_{t+1}}\sigma_t\varepsilon_{d,t+1}
\end{aligned}$$

So, $\forall i\in\{1,...,N]$

$$\begin{aligned}
& \mathrm{E}_t\left(\exp(m_{t+1}+r_{m,t+1})\right)-1 && =0 \\
\Rightarrow & \sum_{j=1}^{N}p_{ij}\left\{\theta\log\delta+\mu_{d,j}-\gamma\mu_{c,j}+(\theta-1)k_{0,j}+(\theta-1)(k_{1,j}A_{0,j}-A_{0,i})\right. \\
& +k_{0m,j}+k_{1m,j}A_{0m,j}-A_{0m,i}+((\theta-1)k_{1,j}A_{2,j}+k_{1m,j}A_{2m,j})\left(1-\nu_j\right)\bar{\sigma}_j^2 \\
& +\frac{1}{2}\left[\pi_{\sigma,j}-\gamma\varphi_{\sigma,j}+k_{1m,j}A_{2m,j}+(\theta-1)k_{1,j}A_{2,j}\right]^2\sigma^2_{w,j} \\
& +\left[(\phi_j-\gamma)+\rho_jk_{1m,j}A_{1m,j}-A_{1m,i}+(\theta-1)(k_{1,j}\rho_jA_{1,j}-A_{1,i})\right]x_t \\
& +\left[k_{1m,j}A_{2m,j}\nu_j-A_{2m,i}+(\theta-1)\left(k_{1,j}A_{2,j}\nu_j-A_{2,i}\right)\right. \\
& \left.\left.+\frac{1}{2}\left((\pi_{c,j}-\gamma)^2+((\theta-1)k_{1,j}A_{1,j}+k_{1m,j}A_{1m,j})^2\varphi^2_{e,j}+\varphi^2_{d,j}\right)\right]\sigma_t^2\right\} && =0
\end{aligned}$$

Thus the unknown coefficients in the log price dividend ratio function are determined as follows :

$$A_{1m} = -\left[D_m^\rho P D_m^\rho - D_m^\rho\right]^{-1} D_m^\rho \left[P(\phi - \gamma 1_N) + (\theta - 1)\left(PD^\rho - I\right) A_1\right]$$

$$A_{2m} = -\left[D_m^\nu P D_m^\nu - D_m^\nu\right]^{-1} D_m^\nu \Bigg[(\theta - 1)\left(PD^\nu - I\right) A_2 + \frac{1}{2} P(\pi_c - \gamma 1_N) \odot (\pi_c - \gamma 1_N)$$
$$+\frac{1}{2} P\left((\theta - 1) D A_1 + D_m A_{1m}\right) \odot \varphi_e \odot \left((\theta - 1) D A_1 + D_m A_{1m}\right) \odot \varphi_e + \frac{1}{2} P \varphi_d \odot \varphi_d\Bigg]$$

$$A_{0m} = -\left[D_m P D_m - D_m\right]^{-1} D_m \left[(\theta - 1)\left(PD - I\right) A_0\right.$$
$$+ P\left(\theta \log \delta 1_N + (\theta - 1) k_0 + \mu_d - \gamma \mu_c + k_{0m} + \left((\theta - 1) D A_2 + D_m A_{2m}\right) \odot (1 - \nu) \odot \bar{\sigma}^2\right)$$
$$+ \frac{1}{2} P\left(\left[D_m A_{2m} + (\theta - 1) D A_2\right] \odot \sigma_w \odot \left[D_m A_{2m} + (\theta - 1) D A_2\right] \odot \sigma_w\right)$$
$$\left.+\frac{1}{2} P\left((\pi_c - \gamma \varphi_\sigma) \odot \sigma_w \odot (\pi_c - \gamma \varphi_\sigma) \odot \sigma_w\right)\right]$$

Where $A_{im} = (A_{im,1}, ..., A_{im,N}) \in \mathbb{R}^N$ for $i = 0, 1, 2$. $k_{0m}, k_{1m} \in \mathbb{R}^N$, $D_m = Diag\left(k_{1m,1}, ..., k_{1m,N}\right) = Diag\left(k_{1m}\right) \in \mathbb{R}^N \times \mathbb{R}^N$, $D_m^\rho = Diag\left(k_{1m,1}\rho_1, ..., k_{1m,N}\rho_N\right) = Diag\left(k_{1m} \odot \rho\right) \in \mathbb{R}^N \times \mathbb{R}^N$ and $D_m^\nu = Diag\left(k_{1m,1}\nu_1, ..., k_{1m,N}\nu_N\right) = Diag\left(k_{1m} \odot \nu\right) \in \mathbb{R}^N \times \mathbb{R}^N$

2.1.1 Bond yields

The price at time t of the n-years zero coupon bond $P_t^{B,(n)}$ is determined recursively as follows:

$$P_t^{B,(n)} = \mathrm{E}_t\left(M_{t+1} P_{t+1}^{B,(n-1)}\right) \tag{2.1.7}$$

with the boundary condition $P_t^{B,(0)} = 1$.

Given the conditional normality of the state variables and the sdf in each regime, the real price at t of a zero coupon bond maturing n-period later ($P_t^{B,(n)}$) is an affine function of the state variables in each regime.

$$P_t^{B,(n)}\left(s_t\right) = \exp\left(B_{0,s_t}^{(n)} + B_{1,s_t}^{(n)} x_t + B_{2,s_t}^{(n)} \sigma_t^2\right) \tag{2.1.8}$$

Equation 2.1.7 implies that : $\forall i \in \{1, ..., N]$

$$\sum_{j=1}^{N} p_{ij} \left(B_{1,j}^{(n-1)} \rho_j - B_{1,i}^{(n)} - \gamma + (\theta - 1)(k_{1,j}\rho_j A_{1,j} - A_{1,i}) \right) = 0$$

$$\sum_{s_{t+1}=1}^{N} p_{is_{t+1}} \left\{ B_{2,j}^{(n-1)} \nu_j - B_{2,i}^{(n)} + (\theta - 1)(k_{1,j} A_{2,j} \nu_j - A_{2,i}) + \frac{1}{2}\gamma^2 \right.$$
$$\left. + \frac{1}{2} \left(B_{1,j}^{(n-1)} + (\theta - 1)(k_{1,j} A_{1,j} \right)^2 \varphi_{e,,j}^2 \right\} = 0$$

$$\sum_{j=1}^{N} p_{ij} \left\{ \theta \log \delta + (\theta - \frac{\theta}{\psi} - 1)\mu_{c,,j} + (\theta - 1)k_{0,j} + (\theta - 1)(k_{1,j} A_{0,j} - A_{0,i}) \right.$$
$$+ \left(B_{2,j}^{(n-1)} + (\theta - 1)k_{1,j} A_{2,j} \right) (1 - \nu_j)\bar{\sigma}_j^2 + B_{0,j}^{(n-1)} - B_{0,i}^{(n)}$$
$$\left. + \frac{1}{2} \left[B_{2,j}^{(n-1)} - \gamma \varphi_{\sigma,j} + (\theta - 1)k_{1,j} A_{2,j} \right]^2 \sigma_{w,j}^2 \right\} = 0$$

So the unknown coefficients $B_0^{(n)}$,$B_1^{(n)}$and $B_2^{(n)}$ can be solved recursively as follows :

$$\Rightarrow B_1^{(n)} = Diag(\rho) P B_1^{(n-1)} + \left[-\gamma P 1_N + (\theta - 1)\left(P D^{\rho} - I \right) A_1 \right]$$

$$B_2^{(n)} = Diag(v) P B_2^{(n-1)} + (\theta - 1)\left(P D^{\nu} - I \right) A_2 + \frac{1}{2} P \left[\gamma 1_N \odot \gamma 1_N \right.$$
$$\left. + \left((\theta - 1) D A_1 + B_1^{(n-1)} \right) \odot \varphi_e \odot \left((\theta - 1) D A_1 + B_1^{(n-1)} \right) \odot \varphi_e \right] \Big)$$

$$B_0^{(n)} = P B_0^{(n-1)} + P\left(\theta \log \delta 1_N - \gamma \mu_c + (\theta - 1)k_0 \right) + (\theta - 1)(PD - I)A_2$$
$$+ P \left(B_2^{(n-1)} + (\theta - 1)\left(D - D^{\nu} \right) A_2 \right) \odot \bar{\sigma}^2$$
$$+ \frac{1}{2} P \left[(\pi_\sigma - \gamma \varphi_\sigma) + (\theta - 1) k_1 \odot A_2 + A_2^{(n-1)} \right] \odot \sigma_w \odot \left[(\pi_\sigma - \gamma \varphi_\sigma) + (\theta - 1) k_1 \odot A_2 + A_2^{(n-1)} \right] \odot \sigma_w \Big)$$

2.1.2 Dividend yields

The price at time t of the n-years zero coupon dividend $P_t^{(n)}$ is determined recursively as follows:

$$P_t^{(n)} = \mathrm{E}_t \left(M_{t+1} P_{t+1}^{(n-1)} \right) \tag{2.1.9}$$

with the boundary condition $P_t^{(0)} = D_t$.

Equation 2.1.9 implies that: $\forall i \in \{1, ..., N\}$

$$\sum_{s_{t+1}=1}^{N} p_{is_{t+1}} \left(A_{1,s_{t+1}}^{(n-1)} \rho_{s_{t+1}} - A_{1,i}^{(n)} + \phi_{s_{t+1}} - \gamma + (\theta - 1)(k_{1,s_{t+1}} \rho_{s_{t+1}} A_{1,s_{t+1}} - A_{1,i}) \right) = 0$$

$$\sum_{s_{t+1}=1}^{N} p_{is_{t+1}} \Big\{ (\theta - 1) \left(k_{1,s_{t+1}} A_{2,s_{t+1}} \nu_{s_{t+1}} - A_{2,s_t} \right) + A_{2,s_{t+1}}^{(n-1)} \nu_{s_{t+1}} - A_{2,i}^{(n)}$$
$$+ \frac{1}{2} \left(\left[\pi_{c,s_{t+1}} - \gamma \right]^2 + \left((\theta - 1) k_{1,s_{t+1}} A_{1,s_{t+1}} + A_{1,s_{t+1}}^{(n-1)} \right)^2 \varphi_{e,s_{t+1}}^2 + \varphi_{d,s_{t+1}}^2 \right) \Big\} = 0$$

$$\sum_{s_{t+1}=1}^{N} p_{is_{t+1}} \big\{ \theta \log \delta + \mu_{d,s_{t+1}} - \gamma \mu_{c,s_{t+1}} + (\theta - 1) k_{0,s_{t+1}} + (\theta - 1)(k_{1,s_{t+1}} A_{0,s_{t+1}} - A_{0,i})$$
$$+ A_{0,s_{t+1}}^{(n-1)} - A_{0,i}^{(n)} + \left[A_{2,s_{t+1}}^{(n-1)} \left(1 - \nu_{s_{t+1}} \right) + (\theta - 1) k_{1,s_{t+1}} A_{2,s_{t+1}} \left(1 - \nu_{s_{t+1}} \right) \right] \bar{\sigma}_{s_{t+1}}^2$$
$$+ \frac{1}{2} \left[\pi_{\sigma,s_{t+1}} - \gamma \varphi_{\sigma,s_{t+1}} + (\theta - 1) k_{1,s_{t+1}} A_{2,s_{t+1}} + A_{2,s_{t+1}}^{(n-1)} \right]^2 \sigma_{w,s_{t+1}}^2 \big\} = 0$$

So the unknown coefficients can be solved recursively as follows:

$$A_1^{(n)} = Diag(\rho) P A_1^{(n-1)} + \left[P \left(\phi - \gamma 1_N \right) + (\theta - 1) \left(P D^{\rho} - I \right) A_1 \right]$$

$$A_2^{(n)} = Diag(v) P A_2^{(n-1)} + (\theta - 1) \left(P D^{\nu} - I \right) A_2 + \frac{1}{2} P \left[(\pi_c - \gamma 1_N) \odot (\pi_c - \gamma 1_N) \right.$$
$$\left. + \left((\theta - 1) D A_1 + A_1^{(n-1)} \right) \odot \varphi_e \odot \left((\theta - 1) D A_1 + A_1^{(n-1)} \right) \odot \varphi_e + \varphi_d \odot \varphi_d \right] \Big)$$

$$A_0^{(n)} = P A_0^{(n-1)} + P \left(\theta \log \delta 1_N + \mu_d - \gamma \mu_c + (\theta - 1) k_0 \right) + (\theta - 1)(PD - I) A_2$$
$$+ P \left(A_2^{(n-1)} + (\theta - 1) k_1 \odot A_2 \right) \odot (1_N - \nu) \odot \bar{\sigma}^2$$
$$+ \frac{1}{2} P \left[(\pi_\sigma - \gamma \varphi_\sigma) + (\theta - 1) k_1 \odot A_2 + A_2^{(n-1)} \right] \odot \sigma_w \odot \left[(\pi_\sigma - \gamma \varphi_\sigma) + (\theta - 1) k_1 \odot A_2 + A_2^{(n-1)} \right] \odot \sigma_w \Big)$$

2.2 Model calibrations

As we saw from the closed form formulas obtained previously, the state space model allows to mixed the cash flows and asset pricing properties inherited from single state models. In this section, we evaluate this mixing feature of the state model. For that, we calibrate a two states regime switching model using two vectors of parameters that yield different term structure implications in single state models : One of them implies an increasing term structure of risk premium (upward sloping state) and the other implies a decreasing term structure of risk premium (downward sloping state). The two vectors of parameters used for calibrations are described as follows :

Table 2.1: Two states model's calibration

Parameter	state 1 (Upward)	state 2 (Downward)
μ_d	0.0043	0
ϕ	3.079	0
φ_d	5.741	0
π_c	2.168e-01	-8.581e-02
π_σ	0	-7405.06
Preference and consumption growth parameters		
δ	γ	ψ
0.998	9.503	13.535
μ_c	ρ	φ_e
0.0017	9.963e-01	1.459e-02
$\bar{\sigma}$	ν	σ_w
3.080e-03	9.91e-01	3.984e-06
φ_σ		
-1.679e-02		

The preference parameters and consumption growth parameters are the same for the two states and only the dividend growth parameters differ in each state. More specifically, in the upward sloping state the dividend growth heavily loads on the expected consumption growth ($\phi = 3.079$) and it is conditionally independent from the volatility. On the contrary, in the downward sloping state the dividend growth reacts negatively to an increase in the volatility ($\varphi_\sigma = -7405.06$) and does not load on the expected consumption growth. As explained in the Chapter 2, the main driver of the negative term structure average slope is the negative correlation between dividend growth and consumption growth volatility.

2.2.1 Implied basic moments

The one regime model that yields an increasing term structure of equity returns (upward sloping regime) is successful in capturing many moments in the observed data. The observed means of the consumption growth, the dividend growth, the risk free rate, the market return and the log price dividend ratio are well within the 95 % confidence intervals of the simulated counterparts.

So looking at those means, we could not reject (with 95 % confidence) the hypothesis that the observed data have been generated by the model calibrated at the upward state vector of parameters. The upward sloping model is also successful in matching the second order moments; it only fails to match the volatilities of the dividend growth and the risk free rate. The downward sloping regime model is less successful in matching the observed data moments. In particular, the mean of the log price dividend ratio is above the observed one and its volatility is very low compared to the observed one. The one period auto-correlation is also smaller in the downward sloping model compared to the observed data. The two regimes model with a 50-50 percent transition matrix has inherited the good properties of the upward sloping regime model and can also correct some bad properties of the downward sloping model. For example, the 95 % confidence interval of the log price dividend ratio mean contains the observed data mean ; its variability and auto-correlation have increased. However, the two regimes model also failed where the two one state models do, for example concerning the volatilities of the dividend growth and the risk free rate.

2.2.2 Implied term structure of risk premium

The term structure of the risk premium in the two regimes model is also a mixture of the term structures in the single regime models. Figure 2.1 represents the term structures of excess returns in the two regimes model with different transition matrices. The upper panel represents the one regime models : the Upward sloping regime on the left and the Downward sloping regime on the right. It is obtained by using the identity as the transition matrix such that by starting in a given state, we stay in that state forever. In the Upward sloping regime, the excess return on the zero coupon equity increases with the maturity of the cash flow : It is almost zero in the very short term and increases up to 12 % after 42 years. On the other side, in the Downward sloping regime, the excess return starts around 5 % in the very short term and decreases. So, the Upward sloping regime has more ability to solve the equity premium puzzle compared to the Downward sloping regime. The bottom panel represents the term structure in the two regimes models : On the left the conditional probability of staying in the upward sloping state is 15 % while the conditional probability of remaining in the downward sloping state is 90 %. On the bottom right picture, these two probabilities are respectively 50 %. We see on the bottom left picture that when the probability of staying in the Downward sloping regime is very large compared to the probability of staying in the Upward sloping regime, the term structure slopes

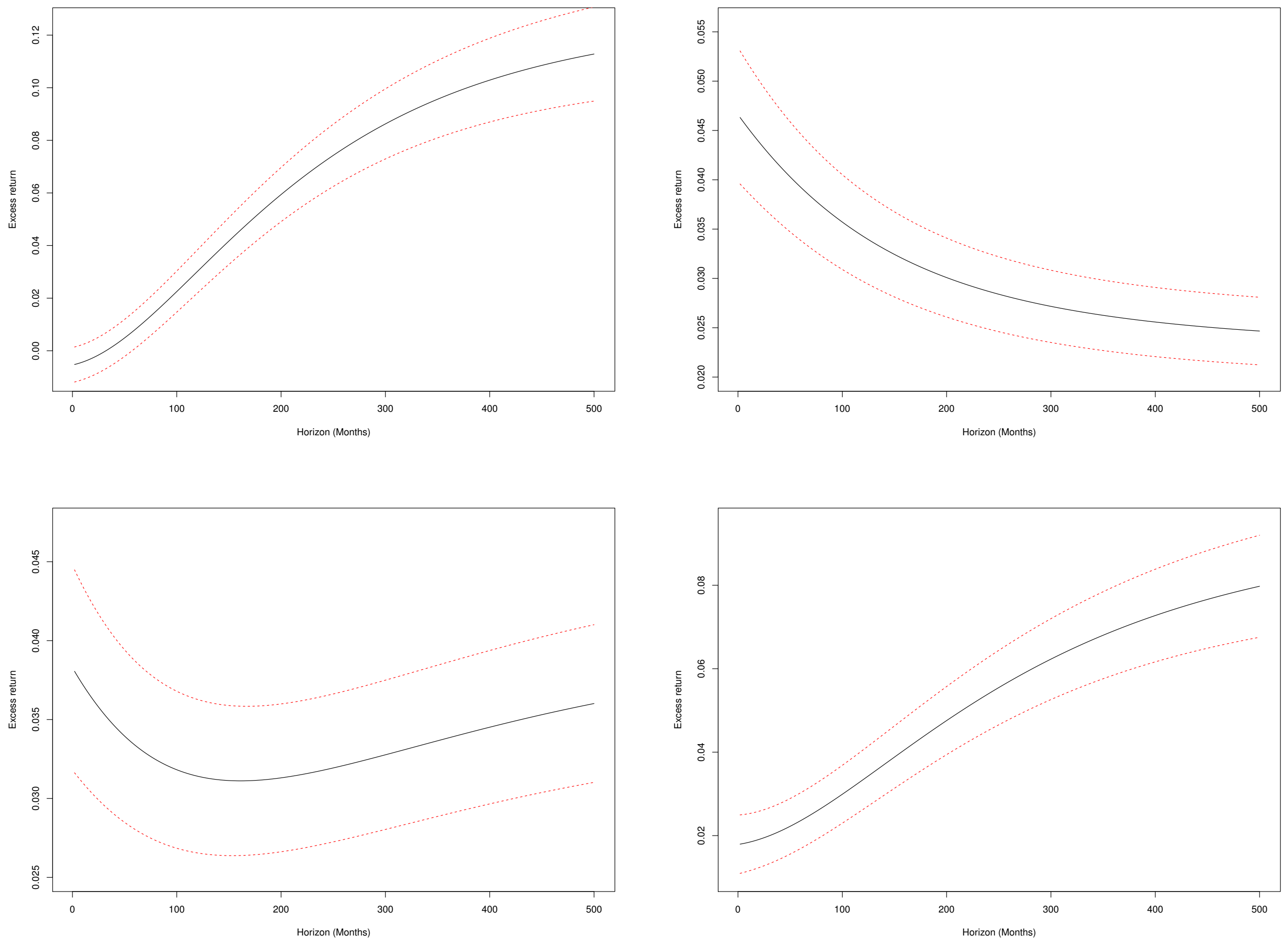
Table 2.2: Basic moments

	One regime model (state1)				One regime model (state2)			Two regimes model		
	Data	Model (50%)	(2.5%)	(97.5%)	Model (50%)	(2.5%)	(97.5%)	Model (50%)	(2.5%)	(97.5%)
EΔc	1.83%	2.03%	0.64%	3.39%	2.06%	0.67%	3.52%	2.07%	0.69%	3.53%
$\sigma(\Delta c)$	2.11%	1.69%	1.21%	2.28%	1.67%	1.19%	2.28%	1.68%	1.20%	2.30%
$\mathrm{AC}_1(\Delta c)$	0.50	0.39	0.12	0.64	0.38	0.12	0.64	0.39	0.12	0.63
$\mathrm{E}(\Delta d)$	4.48%	5.16%	0.62%	9.61%	-0.04%	-2.34%	2.29%	2.60%	-0.63%	5.47%
$\sigma(\Delta d)$	12.17%	8.90%	6.60%	11.75%	8.34%	7.01%	9.78%	8.49%	6.85%	10.49%
$\mathrm{AC}_1(\Delta d)$	0.29	0.29	0.05	0.51	0.23	0.03	0.42	0.24	0.04	0.44
$\mathrm{E}(r_m)$	6.01%	8.65%	4.96%	12.37%	2.79%	0.42%	5.17%	5.97%	3.31%	8.41%
$\sigma(r_m)$	19.10%	20.66%	16.13%	25.76%	6.27%	4.98%	7.66%	13.98%	11.30%	16.96%
$\mathrm{AC}_1(r_m)$	-0.01	-0.05	-0.26	0.17	0.19	-0.05	0.41	-0.04	-0.25	0.17
$\mathrm{E}(r_f)$	0.39%	0.09%	-0.55%	0.58%	0.09%	-0.65%	0.62%	0.10%	-0.58%	0.62%
$\sigma(r_f)$	2.09%	0.84%	0.53%	1.27%	0.84%	0.51%	1.27%	0.84%	0.51%	1.28%
$\mathrm{AC}_1(r_f)$	0.75	0.83	0.67	0.93	0.83	0.66	0.93	0.82	0.65	0.93
$\mathrm{E}(pd)$	3.41	3.37	2.79	3.95	3.56	3.52	3.62	3.41	3.14	3.69
$\sigma(pd)$	0.46	0.40	0.23	0.65	0.07	0.04	0.10	0.19	0.11	0.31
$\mathrm{AC}_1(pd)$	0.89	0.85	0.67	0.95	0.59	0.24	0.83	0.77	0.50	0.92

This table presents the first and second order moments from the data and the model simulations. The simulations are made by calibrating the one regime model with the vector of parameters presented in table 2.1. The transition probability matrix allows to move to an state with a fifty-fifty chance : $P = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$.

down in the short run and then up as the maturity increases. On the bottom right picture, with a fifty-fifty transition probability, the term structure slopes up but with a less steep slope compared to one in the Upward sloping regime. Thus the regime switching model gives more flexibility to the behavior of the term structure while still enabling to capture the key asset pricing moments.

Figure 2.1: Term structures of equity returns



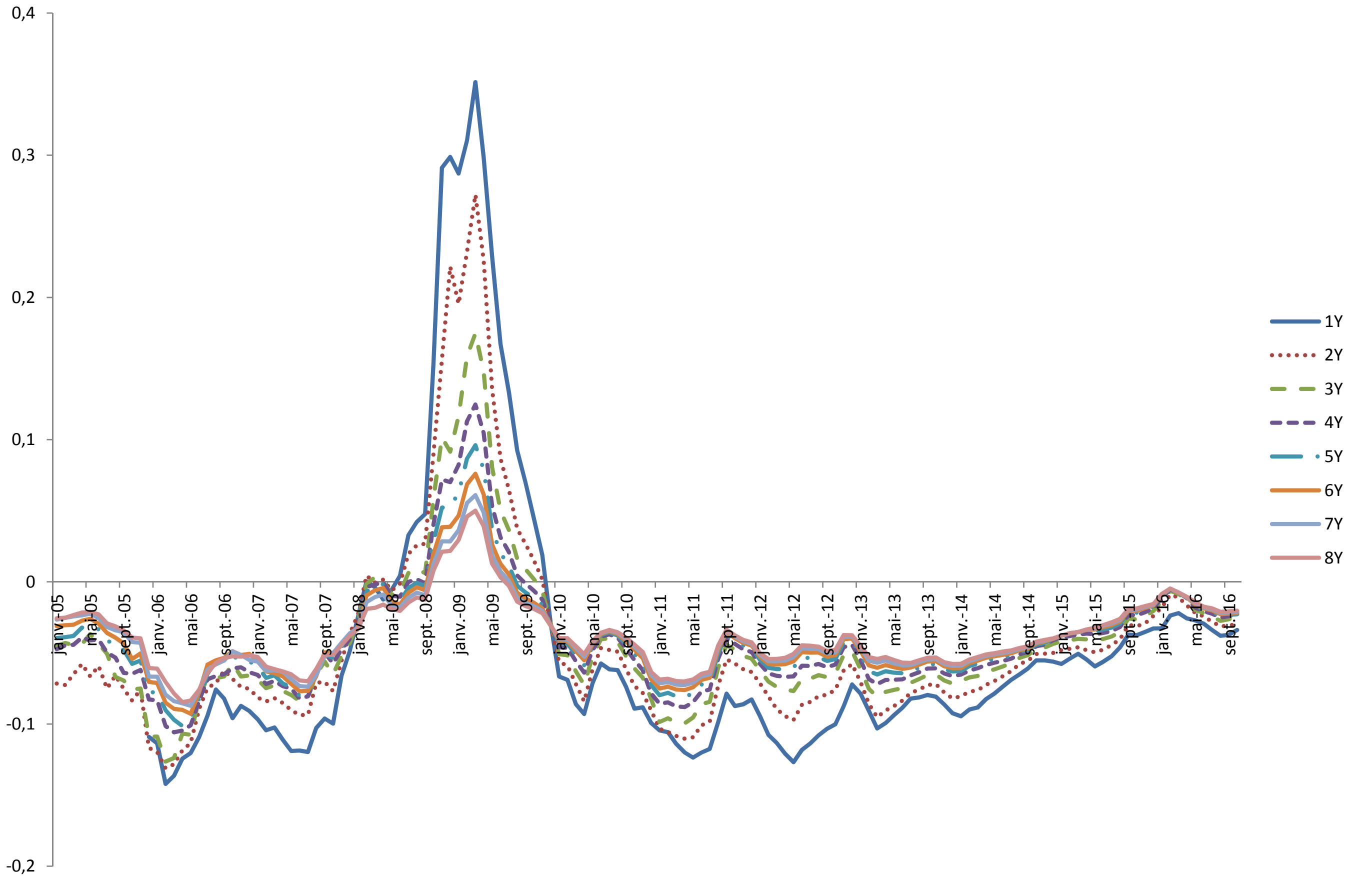
This figure represents the term structures of excess returns in the two regimes model with different transition probability matrices. The upper panel represents the one regime models : the Upward sloping regime on the left and the Downward sloping regime on the right. The bottom panel represents the term structure in the two regimes models : On the left the conditional probability of staying in the upward sloping state is 15 % while the conditional probability of remaining in the downward sloping state is 90 %. On the bottom right picture, these two probabilities are respectively 50 %.

2.3 Conclusion

In this Chapter, we propose a regime switching model in a consumption based asset pricing model with recursive preference and cash flows featuring long run risk and stochastic volatility.

We apply the model in a two regimes context characterized by dividend growth dynamics which put the consumer in an Upward sloping or a Downward sloping term structure regime. This regime switching model offers more flexibility in capturing the changing behavior of the equity returns term structure slope as observed during the crisis periods (During normal period the equity returns term structure slopes upward, while it becomes downward sloping during crisis periods). This work is a first step in that direction. There are many challenges still remaining such as the estimation of the transition matrix and the state specific parameters, but this is left for future work.

Figure 2.2: S&P500 yield curves



This figure displays the evolution of the yield curve of zero coupon equities with maturities ranging from 1 to 8 years. The yield is defined as one over the maturity times the logarithm of the dividend to price ratio ($dy_t^{(n)} = -\frac{1}{n} \log \frac{P_t^{(n)}}{D_t}$). We see that during the crisis period (between May 2008 and December 2009, the yield on the 1 year dividend strip is higher than the yield on dividend strips with higher maturities. The yield spread (difference between higher maturity dividend strip yield and shorter maturity dividend strip yield) is negative during that period while during normal times it is positive on average. So the figure speaks in favor of two regimes characterizing the term structure of equity returns : One with an upward sloping term structure or a positive yield spread and another with a downward sloping term structure or a negative yield spread.

3 Expectation Hypothesis on Bond and Equity markets

The Expectation Hypothesis (EH) states that the yield spread between long-term and short-term interest rates should predict future changes in the interest rate. It is a consequence of the Efficient Market Hypothesis. Indeed, if Market are efficient then returns on stocks or bonds are not predictable and all the variations in the long term spot yield spread comes from the expected variation in the future short-term yield spread. The EH has meanly been tested and rejected on the bond market. In this paper we extend this test both theoretically and empirically to the equity market. To the best of our knowledge, this paper is the first to formulate and test the expectation hypothesis on the equity market. The question we want to answer is whether the future changes in the 1-period dividend yields can be predicted by the current yield spread between the long-term and the short-term dividend strips. This question is interesting for the forecasting of future changes in the short-term dividend yield and to understand the evolution of the slope of the term-structure of the equity returns. Following Campbell and Shiller. [1991], we show that if the EH holds for dividend strips meaning that the expectations of the 1-period future returns on any maturity dividend strip conditional on the information at time t are the same, then the current weighted dividend yields spread over long maturity should predict the future changes in the dividend yields spread for shorter maturities. We test the implication of this hypothesis using the dividend strips forward prices on the S&P 500 and we found a slope coefficient statistically not different from 1 in most cases. Thus the EH is not rejected in the dividends strip data. This "positive" result about the EH has the negative consequence that we can not say something about the predictability of future 1-period return on dividend strips[1]. To further investigate the predictability of 1-period returns on dividend strips with different maturities, we run the Cochrane and Piazzesi [2005] type regression of dividend strips future

[1] Recall that a rejection of the EH would have implied that the 1-period returns on dividends strips are not predictable.

excess returns on a linear combination of forward rate. We find that a linear combination of forward rates up to 5 years (one factor model) predict the 1-period excess return on dividend strips with maturities ranging from 2 years to 8 years.

3.1 Dividend strip regressions

3.1.1 Testing the Expectation Hypothesis

Following Campbell and Shiller. [1991], we define the expectations theory of the term structure of returns as the relationship between a long-term n-period return and a shorter-term m-period return. Let us define by $F_t^{(n)}$ the forward price at t of a $n-$periods zero coupon equity; it is the price settled at time t but paid at time $t+n$ at the maturity of the dividend strip contract. The spot and forward prices of dividend claim maturing in n periods are linked by the following no-arbitrage relation:

$$F_{t,n} = P_{t,n} \exp(ny_{t,n}) \tag{3.1.1}$$

Where $y_{t,n} = -\frac{1}{n} \log P^B_{t,n}$ denotes the yield on the n-period bond at time t and $P^B_{t,n}$ is the spot price a time t of a bond maturing n periods later. The EH test can be formulate using the spot or forward prices; here we present the formulation based on forward prices since we have data on dividend swap prices which are settled today but executed at the maturity of the contract. The maturities are denominated in years and the prices are listed daily. We compute the monthly prices by averaging the daily prices. The maturity increment are made in years and the time increments are made in multiple of 12 months to match the annual frequency of the maturity increments; so $t+n$- period in the time dimension means date $t+12n$ in the monthly observed data.

Proposition 1. *The expectation hypothesis can be tested by regressing the future changes in the forward dividend yield on the current forward dividend yield spread as follows:*

$$dfy_{t+m}^{(n-m)} - dfy_t^{(m)} = a_0^{m,n} + a_1^{m,n} \left(\frac{n}{n-m} \right) \left(dfy_t^{(n)} - dfy_t^{(m)} \right) + \varepsilon_{t+m} \tag{3.1.2}$$

where $dfy_t^{(n)} = \frac{1}{n} \log \left(\frac{D_t}{F_t^{(n)}} \right)$ is the n-year forward dividend yield. Under the expectation hypothesis, the coefficient $a_1^{m,n}$ should be equal to 1.

Proof. (See Appendix.) □

Proposition 1 says that the forward dividend yield spread is a constant risk premium plus an optimal forecast of the changes in future forward dividend yields. It generalizes the Campbell and Shiller (1991) regressions to test the EH on bonds to the equity market. For the bond market, applying the restriction that the dividend is constantly equal to 1 allows to recover the bond regressions formula. A current positive yield spread between the longer n-year maturity and the shorter m-year maturity dividend strips predicts positive future changes in the future yield spread between the shorter maturity and the longer maturity dividend strips. A difficulty in applying the forward yield spread regressions in equation 3.1.2 arises from the quick lost of data as the maturity increases. The maximum year of contracts to mature is fixed at 2024, thus as the time passes the maximum achievable maturity for the contracts reduces and the data length for longer maturity contracts becomes shorter[2]. For example, during the year 2015, we still observe date for 9 years dividend strip which will mature in 2024, but in 2016 this contract does no more exist and the maximum maturity contract become the 8 years dividend strip.

3.1.2 Data and empirical results

Estimates of the forward yield spread regressions in equation 3.1.2 are given in Table 3.1. It shows the results of the regressions to test the EH using the S&P 500 dividend strips data with maturities ranging from 1 year to 8 years [3]. The slope coefficients on the forward yield spread is estimated around 1 and it is statistically not different from at the 95 % confidence test. So the expectation hypothesis in not rejected in the dividend strips sample data under study. We also see that the current forward yield spread can explain a big part of the future development in the future yield spread with R-squared ranging from 38 % to 97 %. Going back to the proof of proposition 3.1, the failure to obtain a slope coefficient equals to 1 in the forward dividend yield regressions implies a rejection of the unpredictability of returns. On the contrary, when returns are unpredictable we should obtain a slope coefficient of 1 as it is the case from the results in Table 3.1.

[2]This explains the reduction in the sample sizes of the forward regressions as shown in table 3.1 as the maturity increases.

[3]See Appendix for details on the data construction.

Table 3.1: Forward yield spread regressions

m=1							
	n=2	n=3	n=4	n=5	n=6	n=7	n=8
Cst. ($a_0^{m,n}$)	-0.011	-0.011	-0.002	0.008	0.011	0.008	-0.014
s.e	*(0.020)*	*(0.016)*	*(0.009)*	*(0.011)*	*(0.010)*	*(0.006)*	*(0.005)*
Slope ($a_1^{m,n}$)	0.688	1.070	1.195	1.015	0.940	0.918	0.630
s.e	*(0.187)*	*(0.113)*	*(0.057)*	*(0.098)*	*(0.091)*	*(0.051)*	*(0.046)*
R-squared	0.377	0.670	0.879	0.858	0.896	0.965	0.886
Sample size	118	106	94	82	70	58	46
m=2							
		n=3	n=4	n=5	n=6	n=7	-
Cst. ($a_0^{m,n}$)	-	-0.020	-0.010	0.006	-0.004	-0.031	-
s.e	-	*(0.014)*	*(0.010)*	*(0.008)*	*(0.008)*	*(0.012)*	-
Slope ($a_1^{m,n}$)	-	1.380	1.291	1.046	0.796	0.552	-
s.e	-	*(0.112)*	*(0.063)*	*(0.050)*	*(0.197)*	*(0.299)*	-
R-squared	-	0.622	0.831	0.911	0.558	0.110	-
Sample size		118	94	70	46	22	
m=3							
	-	-	n=4	n=5	n=6	-	-
Cst. ($a_0^{m,n}$)	-	-	-0.0090	0.0104	-0.0401	-	-
s.e	-	-	*(0.011)*	*(0.009)*	*(0.005)*	-	-
Slope ($a_1^{m,n}$)	-	-	1.149	0.865	-0.075	-	-
s.e	-	-	*(0.145)*	*(0.098)*	*(0.304)*	-	-
R-squared	-	-	0.665	0.789	0.007	-	-
Sample size			106	70	34		

This table shows the results of regressions in equation 3.1.2 for $m = 1, 2, 3$ and $n \in \{2, 3, ..., 8\}$. The standard errors have been obtained using the Newey and West [1987] weighting matrix that account for auto-correlation and heteroskedasticity in the errors.

3.1.3 Other formulations of the EH test

Fama and Bliss [1987] provide a test of the EH by regressing the excess return on the bond against the same maturity forward spread.

Let us denote $P_t^{(n)}$ the price at t of a $n-$periods zero coupon equity. This contract pays a the realized dividends D_{t+n} on the underlying asset at the maturity (in n periods from today's date t) and nothing between its inception and the maturity date. We have an endowment W_t that we would like to invest on financial markets. We can decide to invest on the bond market or on the equity market or both. In any case, we can use the absence of arbitrage to relate the spot and forward prices. Suppose we invest on the equity market, then we can buy a zero-coupon equity and hold it until maturity or we can hold it for just one period and reinvest the outcome the next period and restart the process each period until maturity. With the first strategy, at time t we will buy $\frac{W_t}{P_t^{(n)}}$ units of zero-coupon equity at the price of $P_t^{(n)}$ per unit and at maturity we will get $\frac{W_t}{P_t^{(n)}}D_{t+n}$. By using the second strategy, at time t we will buy $\frac{W_t}{P_t^{(1)}}$ units of zero-coupon equity and at $t+1$ we will get $\frac{W_t}{P_t^{(1)}}D_{t+1}$. This outcome will be reinvested to buy $\frac{W_t}{P_t^{(1)}}\frac{D_{t+1}}{P_{t+1}^{(1)}}$ units of zero-coupon equity maturing the next period, the pay-off will be collected and we will restart the investment procedure. By the no arbitrage condition, the expected outcome at time t of this two investment strategies should be the same; meaning that:

$$\mathrm{E}_t\left(\frac{D_{t+n}}{P_t^{(n)}}\right) = \mathrm{E}_t\left(\prod_{j=1}^{n}\frac{D_{t+j}}{P_{t+j-1}^{(1)}}\right) \tag{3.1.3}$$

Or equivalently, by denoting $dy_t^{(n)} = \frac{1}{n}\log\left(\frac{D_t}{P_t^{(n)}}\right)$ the dividend yield at time t on the n-period zero coupon equity, we have :

$$\mathrm{E}_t\exp\left(ndy_t^{(n)} + \Delta d_{t,t+n}\right) = \mathrm{E}_t\exp\left(\sum_{j=1}^{n} dy_{t+j-1}^{(1)} + \Delta d_{t,t+n}\right) \tag{3.1.4}$$

Furthermore, if we assume that the dividend growth is independent from the dividend yields which are normally distributed conditional on the information at time t, then :

$$dy_t^{(n)} - dy_t^{(1)} = \frac{1}{n}\sum_{j=1}^{n-1}\mathrm{E}_t\left(dy_{t+j}^{(1)} - dy_t^{(1)}\right) + \frac{1}{2n}\mathrm{Var}_t\left(\sum_{j=1}^{n-1} dy_{t+j}^{(1)} - dy_t^{(1)}\right) \tag{3.1.5}$$

Equation 3.1.5 means that the long term dividend yields at time t are the average of future expected short dividend yields plus a risk premium. So, the current long term dividend yield should predict the future short-term dividend yields.

Using a forward iteration, the term in the conditional expectation on the left hand side of equation 3.1.3 can be expressed as follows:

$$\frac{D_{t+n}}{P_t^{(n)}} = \prod_{j=1}^{n} \frac{P_{t+j}^{(n-j)}}{P_{t+j-1}^{(n-j+1)}} \quad \text{with} \quad P_{t+n}^{(0)} = D_{t+n} \tag{3.1.6}$$

So taking the log of equation 3.1.6 and then the expectation conditioning on time t information set, we get that the price at time t of a n-period dividend strip is the present value of the expected pay-off D_{t+n} discounted at the time t expected values of the future 1-year holding period returns on zero coupon equities :

$$P_t^{(n)} = \exp\left(-\sum_{j=1}^{n} \mathrm{E}_t\left[\ln P_{t+j}^{(n-j)} - \ln P_{t+j-1}^{(n-j+1)}\right] + \mathrm{E}_t \ln D_{t+n-1}\right) \tag{3.1.7}$$

so

$$P_t^{(n)}/D_t = \exp\left(-\sum_{j=1}^{n-1} \mathrm{E}_t\left[\ln P_{t+j}^{(n-j)}/D_{t+j} - \ln P_{t+j-1}^{(n-j+1)}/D_{t+j-1}\right] + \mathrm{E}_t \ln P_{t+n-1}^{(1)}/D_{t+n-1}\right) \tag{3.1.8}$$

Equation 3.1.7 is the similar to the equation (4) in Fama and Bliss [1987] for the bond market. Indeed, following the notations therein, we can rewrite equation 3.1.7 as follows:

$$P_t^{(n)}/D_t = \exp\left(-\sum_{j=1}^{n-1} \mathrm{E}_t h(n, n-j : t+j) - \mathrm{E}_t r(1 : t+n-1)\right) \tag{3.1.9}$$

$$= \exp\left(-\mathrm{E}_t h(n, 1 : t+n-1) - \mathrm{E}_t r(1 : t+n-1)\right) \tag{3.1.10}$$

where $h(u, v : t+u-v) = \ln P_{t+u-v}^{(v)}/D_{t+u-v} - \ln P_t^{(u)}/D_t = udy_t^{(u)} - vdy_{t+u-v}^{(v)}$ is the holding period return of a zero coupon equity of maturity u bought at time t and sold at time $t+u-v$ when it has v periods remaining to maturity adjusted for the dividend growth between t and $t+u-v$ and $r(u : t) = \ln D_t - \ln P_t^{(u)} = udy_t^{(u)}$ is the (log) dividend yield at time t of a u-periods zero coupon equity. For the bond market, the dividend is normalized to 1$ and there is no dividend growth: $h(u, v : t+u-v) = \ln P_{t+u-v}^{(v)} - \ln P_t^{(u)}$ and $r(u : t) = -\ln P_t^{(u)}$. The first line of equation 3.1.8 says that the time t price of a n-years zero coupon equity that pays the realized D_{t+n} at maturity is equal to the current dividend D_t discounted by at the time t expected value of future one year returns on the zero coupon equity.

Let us define the time t 1-period forward return from the period $t+n-1$ to $t+n$ as follows:

$$f_t^{(n)} = \ln P_t^{(n-1)} - \ln P_t^{(n)} = r(n:t) - r(n-1:t) \tag{3.1.11}$$

Then substituting equation3.1.9 in equation 3.1.11 and subtracting the 1-period zero coupon equity dividend yield $r(1:t)$, we get that :

$$f_t^{(n)} - r(1:t) = [\mathrm{E}_t r(1:t+n-1) - r(1:t)] + [\mathrm{E}_t h(n,1:t+n-1) - r(n-1:t)] \tag{3.1.12}$$

Equation 3.1.12 says that variation in the forward premia (the difference between the forward return and the spot yield) either comes from change in the 1-period zero coupon equity dividend yield $n-1$ periods ahead (future-spot yield spread) or from the future variations in the term premium. As highlighted by Fama and bliss (1987), it allows to test the split of information in the forward rate $f_t^{(n)}$ between the future spot rate $r(1:t+n-1)$ and the term premium in the $(n-1)$-period return on an n-period zero coupon equity through the following linear regressions:

$$r(1:t+n-1) - r(1:t) = a_1 + b_1 \left[f_t^{(n)} - dy_t^{(1)}\right] + u_1(t+n-1) \tag{3.1.13}$$

$$h(n,1:t+n-1) - r(n-1:t) = -a_1 + (1-b_1)\left[f_t^{(n)} - dy_t^{(1)}\right] - u_1(t+n-1) \tag{3.1.14}$$

Table 3.2: Fama-Bliss regressions

$r(1:t+n) - r(1:t)$						$h(n,1:t+n-1) - r(n-1:t)$					
	a_1	$s\,(a_1)$	b_1	$s\,(b_1)$	R^2		-a_1	$s\,(a_1)$	$1-b_1$	$s\,(b_1)$	R^2
$n=1$	0,01	0,03	-1,30	0,23	0,34	$n=1$	0,03	0,01	1,62	0,08	0,83
$n=2$	0,01	0,03	-1,24	0,15	0,54	$n=2$	0,04	0,01	0,98	0,04	0,93
$n=3$	0,01	0,03	-1,45	0,20	0,62	$n=3$	0,03	0,00	1,00	0,02	0,97
$n=4$	0,09	0,03	-3,27	1,01	0,24	$n=4$	0,03	0,00	1,02	0,01	0,98
$n=5$	-0,04	0,08	-0,33	1,17	0,00	$n=5$	0,02	0,01	1,34	0,15	0,88
n =6	-0,14	0,10	2,10	1,59	0,12	n =6	0,02	0,01	1,44	0,17	0,88
$n=8$	-0,22	0,10	5,06	2,25	0,29	$n=8$	0,02	0,01	1,37	0,19	0,81

This table shows the results of the regressions in equations 3.1.14. The left panel corresponds to the regression of the future dividend yield spread on the current forward yield spread and the right panel corresponds to the regression of the future yields spread on the current forward yield spread.

The first equation estimates the expected value of the $n-1$ periods change in the 1 period return conditional on the forward-spot spread. A value of b_1 different from 0 means that forward-spot yield spread at time t allows to forecast the change in the 1-year dividend yield $n-1$ years ahead. The second equation estimates the expected value of term premium conditional on the forward-spot spread. If b_1 is equal to 0 then all the term premium is explained by the forward-spot spread.

3.2 Predictability of excess return on dividend strips

Cochrane and Piazzesi [2005] found that excess returns on zero coupon bonds with various maturities were predictable by a linear combination of current log foward rates with maturities ranging from 1 year to 5 years. We run similar regressions for the equity market by regressing excess return on dividend strips on the log forward rates with maturities ranging from 1 year to 5 years as follows:

$$rx_{t+1}^{(n)} = \beta_0^{(n)} + \beta_1^{(n)} fdy_t^{(1)} + \beta_2^{(n)} f_t^{(2)} + ... + \beta_5^{(n)} f_t^{(5)} + \varepsilon_{t+1}^{(n)} \tag{3.2.1}$$

where $fdy_t^{(1)} =$ is the yield on the 1-year zero coupon bond, $f_t^{(n)}$is the forward return defined in equation 3.1.11 and the one period excess return on the n-period dividend strip in defined by : $rx_{t+1}^{(n)} \equiv \ln P_{t+1}^{(n-1)} - \ln P_t^{(n)} - y_t^{(1)}$.

The coefficients of the regressions in equation 3.2.1 are summarized on the first panel of figure 3.1. There is a remarkable pattern of the coefficients showing up especially for the regressions of the 4 years, 5 years, 6 years, 7 years and 8 years dividend strips excess returns on the forward yields. The coefficients on forward yields are very close to one another and this closeness speaks in favor of a unique factor model to explain the joint dynamics of excess returns on dividend strips with various maturities. The single factor model is specified as follows :

$$rx_{t+1}^{(n)} = b_n \left(\gamma_0 + \gamma_1 y_t^{(1)} + \gamma_2 f_t^{(2)} + ... + \gamma_5 f_t^{(5)} \right) + \varepsilon_{t+1}^{(n)} \tag{3.2.2}$$

Since b_n and γ_n are can not be separately identified, we apply the normalization : $\frac{1}{7} \sum_{i=2}^{8} b_i = 1$. The single factor model is estimated in two steps. First, we estimate the coefficients of the unique factor as follows:

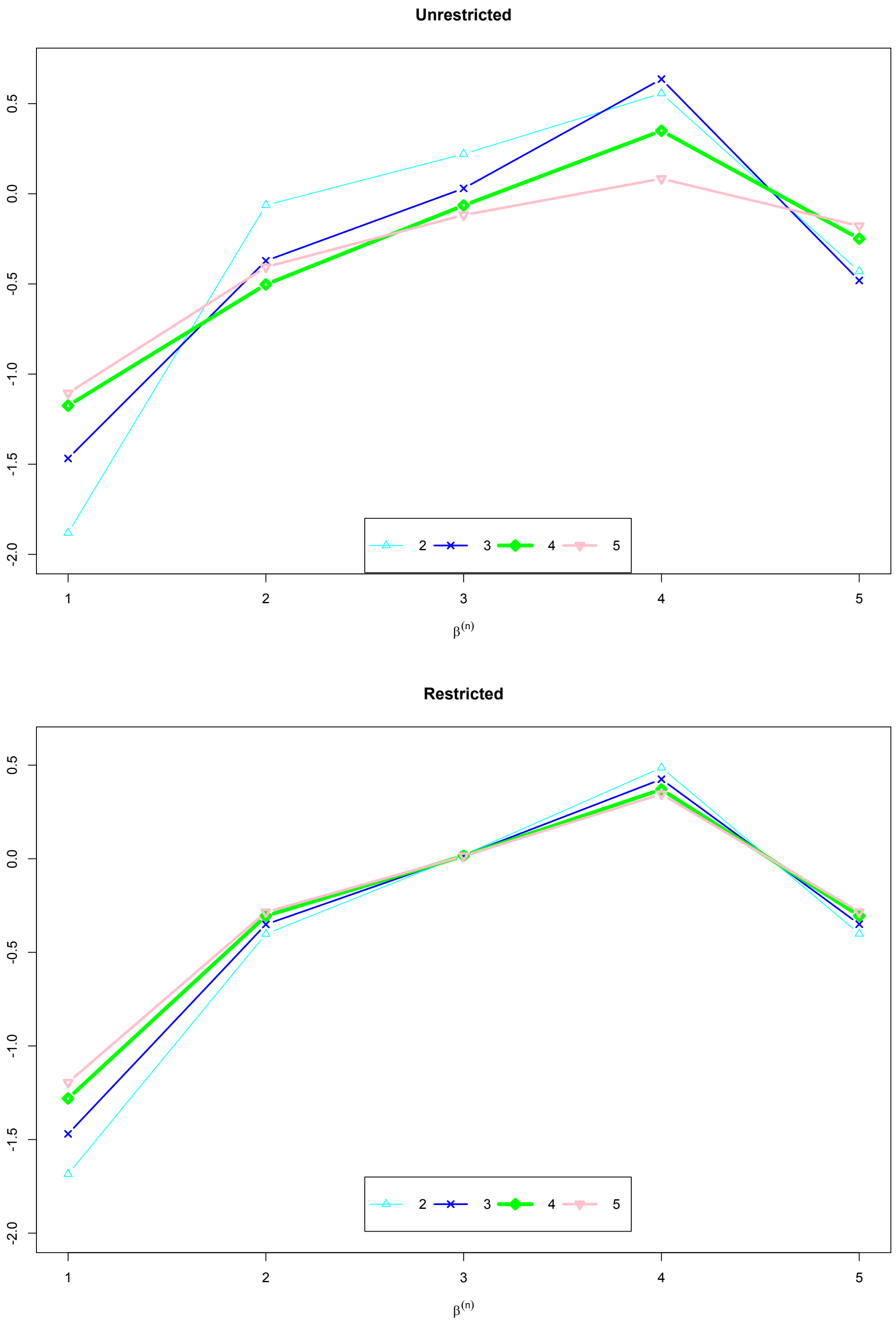
$$\frac{1}{7}\sum_{i=2}^{8} rx^{(i)} = \gamma_0 + \gamma_1 y_t^{(1)} + \gamma_2 f_t^{(2)} + ... + \gamma_5 f_t^{(5)} + \bar{\varepsilon}_{t+1} \quad (3.2.3)$$

Second, the predicted value in equation 3.2.3 is used to run the regressions in equation 3.2.2. We see that the pattern observed for the unrestricted model is also reproduced by the restricted model.

3.3 Conclusion

The expectation hypothesis is a consequence of the efficient market hypothesis (EMH). If markets are efficient then returns on stocks or bonds should not be predictable and thus all the variation in the long-term spot yield spread should come from the expected variation in the future short-term yield spread. So, when the EH holds, there is a one to one relationship between the current long term yield spread and the expected future short-term yield spread; meaning if the current yield curve is upward sloping it is simply because the market expects an increase in the short-term yields. The expectation hypothesis has mainly be tested (and rejected) on the bond market by regressing the future changes of the short-term yields on the current long term yield spread. In this paper, we show that this test can be extended to the equity market by applying the same type of regressions. We formulate an equivalent of the Campbell and Shiller. [1991] regressions that can be applied on equity dividend yields to test the EH. We find that the EH is not rejected when using those equivalent regressions of forward yield spreads. But using the Cochrane and Piazzesi [2005] type regressions, we find that excess return on equity dividend strips with various maturities are predictable in a one factor model, where the factor is obtained as a combination of forward rates up to 5 years of maturity. This unexpected result might sound contradictory but it is not. Indeed, the EH states that if returns are unpredictable then the current yield spreads should predict future yield spread with a slope coefficient of 1. So a failure to obtain a slope coefficient of 1 which means a rejection of the expectation hypothesis implies that the return are predictable. However, obtaining a slope coefficient of 1 (not rejecting the expectation hypothesis) does not say anything about the predictability of return. Thus, it is compatible to have a validation of the EH and a predictability of return as we do here. As future research, we planned to make an international test of the EH on equity markets (EURO STOXX, FTSE, NIKKEI, etc.).

Figure 3.1: Regression coefficients of one year excess returns on forward yields



3.4 Appendix

3.4.1 Proof of the test of EH by regression in 3.1.2

The forward yield can be decomposed as follows (recall that $F^{(0)}_{t+n} = D_{t+n}$):

$$
\begin{aligned}
ndfy^{(n)}_t &= \log\left(\frac{D_t}{F^{(n)}_t}\right) \\
&= \log\left(\frac{D_t}{F^{(n)}_t}\frac{F^{(n-1)}_{t+1}}{D_{t+1}}\frac{D_{t+1}}{F^{(n-1)}_{t+1}}\right) \\
&= \left(r^{n,1}_{t+1} - \Delta d_{t,t+1}\right) + (n-1)dfy^{(n-1)}_{t+1} \\
&= \sum_{i=1}^{n-1}\left(r^{n-i+1,1}_{t+i} - \Delta d_{t+i-1,t+i}\right) + dfy^{(1)}_{t+n-1} \\
&= \sum_{i=1}^{n-1}\left(r^{n-i+1,1}_{t+i} - r^{1,1}_{t+i} - dfy^{(1)}_{t+i-1}\right) + dfy^{(1)}_{t+n-1} \\
&= \sum_{i=1}^{n}\left(r^{n-i+1,1}_{t+i} - r^{1,1}_{t+i} - dfy^{(1)}_{t+i-1}\right)
\end{aligned}
$$

So

$$
dfy^{(n)}_t - dfy^{(m)}_t = \frac{1}{n}\sum_{i=1}^{n}\left(r^{n-i+1,1}_{t+i} - r^{1,1}_{t+i} - dfy^{(1)}_{t+i-1}\right) - \frac{1}{m}\sum_{i=1}^{m}\left(r^{n-i+1,1}_{t+i} - r^{1,1}_{t+i} - dfy^{(1)}_{t+i-1}\right)
$$

if the Expectation Hypothesis holds, then $\forall i, n \mathrm{E}_t\left[r^{n-i+1,1}_{t+i} - r^{1,1}_{t+i}\right] = 0$ such that :

$$
\begin{aligned}
\frac{m}{n-m}\left(dfy^{(n)}_t - dfy^{(m)}_t\right) &= \mathrm{E}_t\left[\frac{1}{n-m}\sum_{i=1}^{m} dfy^{(1)}_{t+i-1} - \frac{m}{n(n-m)}\sum_{i=1}^{n} dfy^{(1)}_{t+i-1}\right] \\
&= \frac{1}{n-m}\mathrm{E}_t\left[\sum_{i=1}^{m}(1-\frac{m}{n})dfy^{(1)}_{t+i-1} - \frac{m}{n}\sum_{i=m+1}^{n} dfy^{(1)}_{t+i-1}\right] \\
&= \frac{m}{n}\left(\frac{1}{m}\mathrm{E}_t\left[\sum_{i=1}^{m} dfy^{(1)}_{t+i-1}\right] - \left[\frac{1}{n-m}\sum_{i=m+1}^{n} dfy^{(1)}_{t+i-1}\right]\right) \\
&= \frac{m}{n}\mathrm{E}_t\left(dfy^{(n-m)}_{t+m} - dfy^{(m)}_t\right)
\end{aligned}
$$

3.4.2 Data construction

We obtain the daily data on the implied dividend swap prices for the S&P 500. These contract allow the buyer to pay at maturity of the contract a price (forward price) settled at the beginning in exchange of the realized dividend on the S&P 500. The data span from 03-Jan.-2005 to 14-Oct.-2016 and the maturities of the contracts range from 1 year (minimum year is 2006) to 16 years (maximum year is 2024). We first compute the monthly prices by averaging the daily prices over the month. Then, we compute the constant maturity prices (CMP). Indeed, all the contracts mature in December of the corresponding year but you can buy it at any date. The CMP are made to correct for the differences in time remaining to maturity. The CMP are computed by averaging weighted consecutive prices. For example, the CMP on March 2006 of a 1 year dividend strip is obtained by summing 9/12 of the price on March 2006 of a dividend strip maturing in 2006 and 3/12 of the price on March 2006 of a dividend strip maturing in 2007. The weighting reflects the fact that 9 months of the 1 year maturity contract are covered in 2006 and the remaining 3 months are covered in 2007. The forward dividend yields are obtained by taking the log of the annualized dividend on the S&P 500 to the CMP ratios. The spot dividend yields are obtained by discounting the forward dividend yield using the corresponding bond yield.

Table 3.3: Cochrane and Piazzesi [2005] regression of excess return on dividend strips

n	2	3	4	5	6	7	8
Cst.	0,002	0,012	0,009	0,010	0,005	0,004	-0,004
s.e	*0,009*	*0,009*	*0,011*	*0,013*	*0,015*	*0,018*	*0,020*
$\beta_0^{(n)}$	-0,002	-0,006	-0,009	-0,010	-0,010	-0,012	-0,012
s.e	*0,002*	*0,002*	*0,002*	*0,003*	*0,003*	*0,004*	*0,004*
$\beta_1^{(n)}$	-0,024	-0,676	1,660	1,858	1,945	2,418	2,771
s.e	*0,192*	*0,423*	*0,681*	*0,874*	*0,967*	*1,117*	*1,242*
$\beta_2^{(n)}$	1,175	-1,740	-8,038	-6,565	-7,833	-9,817	-10,977
s.e	*1,310*	*1,970*	*3,048*	*3,933*	*4,580*	*5,207*	*5,921*
$\beta_3^{(n)}$	-2,402	4,021	8,438	4,310	8,254	11,167	13,744
s.e	*2,874*	*4,056*	*6,369*	*8,138*	*9,505*	*11,092*	*12,361*
$\beta_4^{(n)}$	2,072	-0,742	-1,485	0,807	-2,223	-3,733	-5,818
s.e	*1,903*	*2,792*	*4,207*	*5,293*	*6,125*	*7,349*	*8,048*
R-squared	0,517	0,129	0,190	0,106	0,099	0,115	0,120

This table shows the results of the regressions in equation 3.2.1. The standard errors are computed using the Newey and West [1987] variance-covariance matrix with 8 lags.

4 GMM estimation of the Long Run Risks model

In this paper, we propose a GMM estimation of the structural parameters of the Long Run Risk model that allows for the separation between the consumer optimal decision's frequency and the frequency by which the econometrician observes the data. Our inference procedure is also robust to weak identification. The key finding is that the Long Run Risk model adapts well to the data but could not be so good at forecasting or telling the true story about what drives the evolution of asset prices. Indeed, the model is able to reproduce the qualitative behavior of targeted moments in the long run when the corresponding estimates of the structural parameters are used for simulations, but it also faces a urge tension in keeping in track with all the observed moments considered.

4.1 Introduction

One of the long lasting challenges in asset pricing is to provide a model that could explain the observed level of asset prices and reconcile model's predictions with the observed macro-finance data. Three main competing models, which focus on the decision problem of a representative agent, have received a great attention in this domain : The Habit-formation model (Abel, 1990; Constantinides, 1990; Cochrane and Campbell, 1999), the long run risks model (Bansal and Yaron, 2004) and more recently, the rare disasters model (Rietz, 1988, Barro 2006, 2009; Wachter, 2013). Those alternative models are usually able to account for the observed level of equity premium and risk free rate with a reliable value of the representative agent risk aversion coefficient.

However, each of this model has received some criticisms related to their capacity to yield the same predictions as what have been documented in previous studies from the observed data, especially when we look at higher order moments of the models predictions. Indeed, just to take some few examples, the main intuition behind the habit model is that the drop of consumption toward the habit level rises the risk aversion of the representative agent who asks for a higher equity premium as compensation. So this model implies that past consumption growth should predict future prices but that has not been observed in the data (Bansal et al., 2012a). In the case of the rare disasters model, the eventuality of a rare but large drop in consumption induces a higher equity premium. However, the implied consumption growth process necessary to generate the level of the U.S historical average equity premium is much more volatile and skewed than what we observe in the historical data[1] ([Julliard and Ghosh, 2012]).

More recently, there have also been some criticisms toward the Long run risks (henceforth LRR) model concerning the predictive power that the model confers to asset prices toward cash flows variables. Indeed, as pointed by Beeler and Campbell [2012], the initial calibration of the model implies that the price-dividend ratio has a huge predictive power on the consumption and dividend growth and that is at odd with what is observed in the data. As we know from Cochrane, 2001 Chap. 20, the observed variability in log price-dividend ratio implies that it must predict future excess returns or future dividend growth; but in the data, the log price-dividend ratio seems to predict future excess return and it does not predict future dividend growth. A

[1]This is mainly because the size of the annual drop in consumption growth considered in this model corresponds in the data to a plural-annual summation of consecutive drops. In other words, the probability that such a drop (the one considered in the model) would be observed in the data is more smaller than the rare disaster probability considered to generate the observed equity premium.

subsequent calibration of the model in Bansal et al. [2012a][2] has weakened this predictive power and done a great job in reconciling the model with the data; hence showing that the model implications could be improved just by changing the parameters calibration.

Even though this new calibration of the model has brought the model implications closer to the data by reducing the predictability of future dividend and consumption growth by the asset prices, it has failed to captured the observed predictability of future excess returns by the log price-dividend ratio. Some other mismatches between the model implications and the observed data have also been pointed out and they call for the seek of at least a better calibration of the model in order to reconcile the model predictions, especially in terms of moments matching, with what is observed in the data as advocated by Bansal et al. [2012a], Beeler and Campbell [2012].

The aim of this paper is to estimate the structural parameters that will enable the model to well fit a set of empirical moments derived from the observed data. While this estimation is of particularly importance to test some key assumptions of the model like for example the Elasticity of Inter-temporal Substitution (EIS) being greater than 1, very few studies have attempted to estimate the fundamental parameters of the LRR model. Indeed, the simulation of the model appears to be pretty easy, but the estimation of the fundamental parameters is more challenging due to the unobserved latent variables, the preference structure (Epstein-Zin recursive utility function) and the large number of parameters to be estimated. However, contrary to the calibration exercise which can look like a "fishing expedition", the estimation allows to determined the set of parameters that will bring the model predictions closer to the data in terms of moments we considered, which could also be easily extended to take into account some new stylized facts. Bansal et al. [2012a] recommend the use of a Simulated Method of Moment (SMM) for the estimation of the model parameters instead of a successive regressions approach which produces a downward bias of some coefficients like the EIS and a upward bias of others like the risk aversion coefficient. We try to go further in this direction by deriving the analytical formulas for theoretical moments in order to use the Generalized Method of Moment (GMM) which is more efficient than the SMM.

In this paper, we extend the parametric estimation of the LRR model as done by Constantinides and Ghosh [2011], to take into account some key failures of the previous calibrations

[2] In the Bansal and Yaron, 2004 (henceforth, BY) model the emphasize is put on the persistent of the expected component of the consumption growth process as the main driver of asset prices, while in the Bansal et al. [2012a] (henceforth, BKY) formulation it is the stochastic volatility of the consumption growth process that determines the equity premium level Beeler and Campbell [2012].

of the model pointed by Bansal et al. [2012a], Beeler and Campbell [2012]. Our procedure is closely related to their; the difference being that firstly, we consider an extended version of the LRR model as in BKY where we allow the dividend growth process to be correlated to the short run risk; secondly, we extend the set of moments to be matched, in order to consider some critics formulated against the previous calibrations of the model; and thirdly, we disentangle the agent decision frequency from the one at which the econometrician observes the data through temporal aggregation. Indeed, the LRR model is usually simulated assuming that the agent optimal allocation is made at the monthly frequency while the econometrician observe the data at the annual frequency. We assume that this mapping between the agent decision frequency and the data observation frequency by the econometrician is unknown and could be estimated. Our approach is closely related to Bansal et al. [2012b]who introduced the time aggregation in the estimation of the LRR model, but contrary to them we used a weak identification robust method for the inference. Other approaches have also been used to estimate modified versions of the LRR model. Bansal et al. [2007a] assumed that consumption and dividend growth were co-integrated and used the Efficient Method of Moments (EMM) for the estimation. Schorfheide et al. [2014] disentangled the volatilities of the cash flow processes (consumption and dividend growth), introduced a shock to preferences and used a Bayesian Markov Chain Monte Carlo (MCMC) methods for the estimation[3]. Calvet and Czellar [2015] used an indirect inference method with a restricted version of the LRR model that has a closed form solution as auxiliary model and in their estimation procedure they do not used the market return.

The remaining of the paper is organized as follow: In section 2 we provide the model statement and derive the asset prices formulas in term of underlying parameters. Section 3 provides some comparisons of the model predictions and the data for two common used calibrations of the model. It also highlight the successes and failures of those calibrations and points out the problem we will try to solve by estimating a new set of parameters. Section 4 presents the estimation procedure and some result of the simulations we did in order to verify the reliability of our method before applying it on the observed data. In section 5, we present the results of our estimations and discuss the results. Section 6 concludes. An appendix is provided in section 7 with time aggregation derivations and the analytical formulas of the theoretical moments.

[3]Contrary to them, we prefer to remain parsimonious in terms of model's parameters as our objective is to improve the basic LRR model results through our estimation procedure.

4.2 The model and the solution

The long run risks model initiated by Bansal and Yaron [2004] (BY) provides an explanation to many asset pricing puzzles in the context of the U.S. Economy and the calibrations of the model considered in BY and subsequently in Bansal et al. [2007b] seems to well fit many first and second order moments of the U.S. historical macro-finance data. The success of the model lies in particular in the use of the Epstein and Zin [1989] utility function[4] with an elasticity of inter-temporal substitution (EIS) greater than 1, which features a preference for an early resolution of time uncertainty and the introduction of two predictable latent variables (a persistent component and a time varying volatility) to describe the evolution of the cash flows (consumption and dividend growth) and asset prices processes.

4.2.1 Model statement

We consider a representative agent with Epstein-Zin recursive preferences. For these preferences the Euler equation satisfied by the representative agent's optimal allocation plan is given by:

$$E_t[exp(\theta \log \delta - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{c,t+1} + r_{i,t+1})] = 1 \qquad (4.2.1)$$

with $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$, γ is the coefficient of risk aversion and ψ is the elasticity of inter-temporal substitution (EIS); g_t is the consumption growth process, δ is the discount rate, $r_{i,t}$ is the log-return on a given asset i and $r_{c,t+1}$ is the log-return on wealth portfolio that delivers aggregate consumption as its dividend each period. The evolution of the growth process is driven by x_t the persistent component of the consumption growth process which is unobservable but seems to play a big role in the model andσ_t is the stochastic volatility process.

$$g_{t+1} = \mu + x_t + \sigma_t \eta_{t+1} \qquad (4.2.2)$$

$$x_{t+1} = \rho x_t + \varphi_e \sigma_t e_{t+1}$$

$$\sigma_{t+1}^2 = \bar{\sigma}^2 + \nu(\sigma_t^2 - \bar{\sigma}^2) + \sigma_w w_{t+1}$$

$$g_{d,t+1} = \mu_d + \phi x_t + \pi \sigma_t \eta_{t+1} + \varphi_d \sigma_t u_{t+1}$$

$$w_{t+1}, e_{t+1}, u_{t+1}, \eta_{t+1} \sim N.i.id(0,1)$$

$e_{t+1}, \eta_{t+1}, w_{t+1}$represent the three sources of risk in the aggregate consumption dynamics

[4]The Epstein-Zin recursive utility function allows to disentangle the time preference (characterized by the elasticity of inter-temporal substitution) from the risk aversion. Indeed, the fact that someone is highly risk averse does not necessarily imply that she will also want to smooth consumption over time [Mehra, 2003].

that are priced in the long run risks model and correspond respectively to the short run risk in consumption, the long run risks and the volatility risk. π represents the magnitude of the influence of short run consumption shocks on the dividend growth process. In The Bansal and Yaron [2004] calibration, this coefficient is equal to zero, while it is equal to 2.6 in Bansal et al. [2012a]; furthermore it has not been considered in the Constantinides and Ghosh [2011] estimation of the model parameters and so, this omission could be damaging for the estimation of the other parameters of the model or could implied a rejection of the model. The exposure of the dividend growth process to the long run consumption shocks is captured by ϕ.

4.2.2 Solution

To solve the model, Bansal and Yaron [2004] assume that the price-consumption ratio, the price-dividend ratio and the risk free rate are affine functions of the states variables.

$$\begin{aligned} z_t &= A_0 + A_1 x_t + A_2 \sigma_t^2 \\ z_{m,t} &= A_{0,m} + A_{1,m} x_t + A_{2,m} \sigma_t^2 \\ r_{f,t} &= A_{0,f} + A_{1,f} x_t + A_{2,f} \sigma_t^2 \end{aligned} \tag{4.2.3}$$

They combined this with the Campbell and Shiller [1988] approximation of the returns on the wealth and market portfolios; $r_{c,t+1} = k_0 + k_1 z_{t+1} - z_t + g_{t+1}$ and $r_{m,t+1} = k_{0,m} + k_{1,m} z_{m,t+1} - z_{m,t} + g_{d,t+1}$ with $k_0 = \log(1 + e^{\bar{z}}) - k_1 \bar{z}$, $k_1 = \frac{e^{\bar{z}}}{1+e^{\bar{z}}}$,$k_{0,m} = \log(1 + e^{\bar{z}_m}) - k_1 \bar{z}_m$ and $k_{1,m} = \frac{e^{\bar{z}_m}}{1+e^{\bar{z}_m}}$. Plugging the system of equations (1) and (2) in the Euler equation, we can easily derive the analytical formulas for the coefficients A_0, A_1, A_2, A_{om}, A_{1m}, A_{2m}, A_{0f}, A_{1f} and A_{2f} conditionally on k_0,k_1,$k_{0m}$and k_{1m}which also depend on $\bar{z}$ and $\bar{z}_m$[5]. So, the model is solved by firstly determining the mean values of the log price-consumption and log price-dividend ratios through the resolution of a fixed point problem, secondly the $k's$ are computed and finally the $A's$ coefficients are determined. All this is done using a fixed point algorithm which converges very quickly when the fixed point is feasible. Indeed, the fixed point feasibility has not been emphasized in the previous literature. However, when dealing with an optimization algorithm, this feasibility becomes very relevant because when not satisfied it leads to a crash of the routine. To obtained the fixed point feasibility condition, we substitute k_0, k_1,A_0, A_1, A_2 in the expression of $\bar{z} = A_0 + A_2 \bar{\sigma}^2$ and obtain the following equation:

[5]In many studies of the LRR involving the Campbell and Shiller [1988] approximation, the values of $\bar{z}$ and $\bar{z}_m$are calibrated instead of been determined as the solutions of the fixed point problem. This could lead to a bias estimation of the true parameters of the model. Indeed, we run a simulation experiment showing that calibrating$\bar{z}$ and $\bar{z}_m$to the wrong values leads to a rejection of the true parameters of the model and this happened even when the calibration error is less than 5%.

$$f(\bar{z}) = g(\bar{z}) \tag{4.2.4}$$

where

$$f(\bar{z}) = \bar{z} - \log(\delta) - (1 - \tfrac{1}{\psi})\mu_c - \log(1 + e^{\bar{z}}) \text{ and}$$
$$g(\bar{z}) = \lambda_1 \left(1 + U(\bar{z}, \varphi_e, \rho)^2\right) + \lambda_1 \left(1 + U(\bar{z}, \varphi_e, \rho)^2\right)^2 U(\bar{z}, \lambda_0, \nu)^2$$

$$\text{with } \lambda_0 = 0.5(1-\gamma)\tfrac{\sigma_w}{\bar{\sigma}} \ , \ \lambda_1 = 0.5(1-\gamma)(1-\tfrac{1}{\psi})\bar{\sigma}^2 \text{ and } U(z,a,b) = \tfrac{ae^z}{1+e^z(1-b)}$$

The function f is increasing : $\mathbb{R} \to]-\infty; F[$ and the function

$$g \text{ is } \begin{cases} \textit{decreasing}: \mathbb{R} \to]G_2; G_1[\quad & if \quad \lambda_1 < 0 \\ \\ \textit{increasing}: \quad \mathbb{R} \to]G_1; G_2[\quad & if \quad \lambda_1 > 0 \end{cases}$$

where

$F = \lim_{z \to +\infty} f(z) = -\log(\delta) - (1 - \frac{1}{\psi})\mu_c$, $G_1 = \lim_{z \to -\infty} g(z) = \lambda_1$ and

$$G_2 = \lim_{z \to +\infty} f(z) = \lambda_1 \left(1 + \left(\frac{\varphi_e}{1-\rho}\right)^2\right) \left(1 + \left(1 + \left(\frac{\varphi_e}{1-\rho}\right)^2\right) \left(\frac{\lambda_0}{1-\nu}\right)^2\right)$$

Equation 4.2.4 put a restriction on the preference parameters (δ, γ, ψ) of the representative agent for a given set of consumption growth parameters. For a given set of parameters to be feasible, the pure discount factor should satisfy the following conditions:

$$\begin{cases} \delta \leq e^{-(1-\frac{1}{\psi})\left(\mu_c + \frac{G_2}{1-\frac{1}{\psi}}\right)} & if\ \gamma > 1, \psi > 1\ or\ \gamma < 1, \psi < 1 \\ \delta \leq e^{-(1-\frac{1}{\psi})\left(\mu_c + \frac{G_1}{1-\frac{1}{\psi}}\right)} & if\ \gamma > 1, \psi < 1\ or\ \gamma < 1, \psi > 1 \end{cases}$$

For some vector of parameters, for example if $F < G_2$ while $\lambda_1 < 0$ or $F < G_1$ while $\lambda_1 > 0$, there is no fixed point. A similar condition also derives from the log price dividend ration equilibrium condition (see Appendix). As illustrate in figure 1, the increase of the unconditional mean of consumption growth restricts the feasible combination of agent risk aversion and elasticity of inter-temporal substitution; and the increase of unconditional mean of the dividend growth process further restrict it. Furthermore, once some parameters are kept fixed, the comparison between F and G's determines the fixed point feasible set for the remaining parameters. In the case where $\gamma > 1$and $\psi > 1$ for some values of the unconditional mean of the consumption growth process (μ_c), the upper bound of the pure discount factor will be lower than 1 and in this case, it is an increasing function of the relative risk aversion coefficient (γ). If $\delta = 1$ then there exist $\bar{\gamma}$ such that there is a solution to the fixed point problem only if ($\gamma > \bar{\gamma}$ and $\psi > 1$) or ($\gamma < \bar{\gamma}$ and $\psi < 1$).

This condition is similar to the ones obtained by Burnside [1998] with a CRRA utility function and a Gaussian consumption growth process and extended in de Groot [2015] to a stochastic volatility consumption growth process. They obtained a **sufficient** condition for the price dividend ratio to exist[6], but contrary to them our condition is a **necessary**[7] condition for an equilibrium finite value of the price dividend ratio to exist and it is determined after applying the Campbell and Shiller [1988] approximation. In a model with CRRA utility function ($\gamma = \frac{1}{\psi}$), our condition is necessary and sufficient (analytically it is the same as the one obtained by de Groot [2015]) for the existence of the price-consumption ratio, but with the Epstein-Zin recursive utility function, we show that it is much more less restrictive compare to the one obtained through the series convergence criterion. Creal and Wu [2015] also studied the fixed problem in a consumption based model with recursive preferences, where they introduced some shocks to preferences and affine dynamics of the state variables. There conclusions are similar to the ones we obtained here.

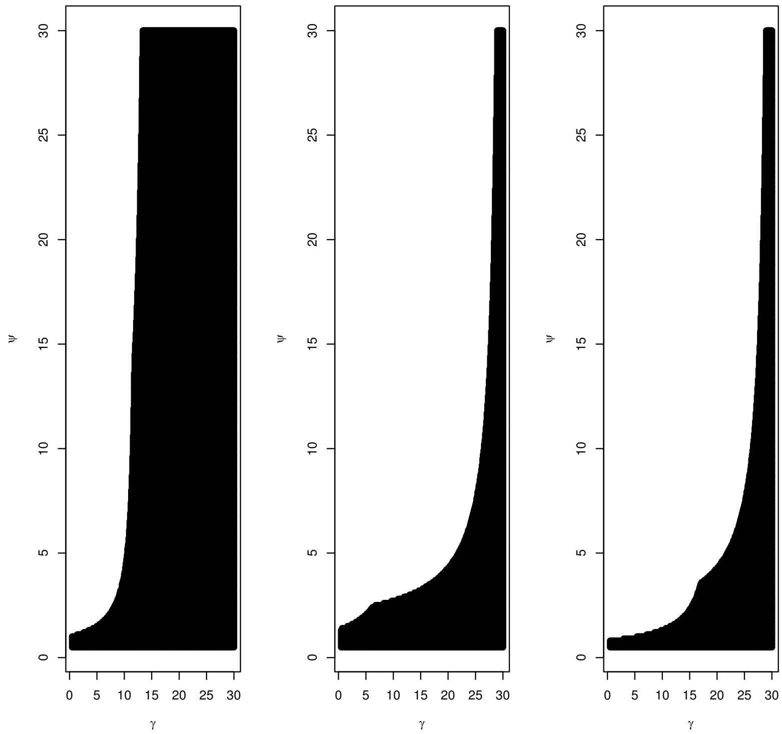
To handle this fixed point existence problem in the estimation, we set a ruling out condition in the optimization algorithm for all the parameter's vector for which the fixed point does not exist; in the case where the algorithm is trying such a vector of parameters, the objective function is automatically set to a very high value.This penalty value imposes the optimizers to seek the parameters in other regions (Holly et al. [2011]). By trying various combinations of parameters, the optimization algorithm allows to expand the set of possible parameter values compared to a calibration exercise where some parameters are given less attention (like for example, the ones driving the consumption and dividend processes) and are preliminary fixed, hence constraining the set of feasible parameters. This fixed point feasibility might also cause difficulties to optimization method based on the gradient if its analytical expression is not provided. Indeed, as the numerical approximation of the gradient value needs two points where

[6]When the stochastic volatility is shut down, Their condition is is given by $\delta \leq e^{-\left((1-\gamma)\mu_c + 0.5\left(\frac{1-\gamma}{1-\rho}\right)^2\sigma^2\right)}$, which is obtained from the convergence criterion of a given series

[7]When the sufficient condition is satisfied, we are sure that the price dividend ratio exist, but it does not mean that it will not exist when the condition is not satisfied, instead the sufficient condition can be too much restrictive and should not be used to construct the set of unfeasible parameters. On the other hand, when the necessary condition is satisfied, the price dividend ratio could possibly exist and we are sure that it does not exist when that condition is not verified, so this condition in more appropriate to construct the set of unfeasible parameters.

the objective function will be evaluated, one of the two might fall in the infeasible region. That is why we preferred an optimization method based uniquely on the objective function (The Nelder and Mead [1965]'s algorithm).

Figure 4.1: Feasible set forγ and ψ with the remaining parameters fixed



This figure shows the feasible values (for which the model can be simulated) of RRA (γ) and EIS (ψ) when all the other parameters are kept fixed. We restrict the values of γ and ψ in $[0.5, 30]$. In the baseline picture, the value of the other parameter are as in the BKY monthly calibration: $\delta = 0.9989, \mu_c = 0.0015, \mu_d = 0.0015, \phi = 2.5, \varphi = 5.96, \rho = 0.975, \varphi_e = 0.038, \sigma = 0.0072, \nu = 0.999, \sigma_w = 0.0000028, \pi = 2.6$. in the middle picture the value of the average consumption growth is put at $\mu_c = 0.0020$ and in the right picture: $\mu_c = 0.0020, \mu_d = 0.0025$.

4.2.3 Temporal aggregation

The structural parameters used to simulate the LRR model are usually assumed to generate the data at the monthly frequency. Then these data are aggregated at the annual frequency and the moment are computed to be compared to those computed in the real annual data. We choose to follow this tradition here. But we could also, as done by Bansal et al. [2012b], assume that the frequency at which the model evolves is unknown but we still observed the data at the annual frequency. Then the aggregation frequency becomes another parameter that could be determined optimally in order to match the observed data. This implies that the annual theoretical moments should be computed while taking the temporal aggregation from an unknown frequency, which would be within a year, to the annual frequency. For example, it could be that the representative agent in the model takes its decisions about the optimal consumption allocation every 2 months; then the annual data will be the results of 6 period's temporal aggregation. Appendix B. presents the analytical formulas for cash flows and asset prices allowing to move from the agent decision frequency to the annual frequency while assuming that there are h-periods within a year. The theoretical moment derived from the aggregate processes will be used to match the corresponding ones observed in the data and this is how our GMM procedure will estimate the structural parameters of the model to match the observed data.

4.3 Moments matching

In this section we simulate the model using the standard calibrations as in Bansal and Yaron [2004], Bansal et al. [2012a] and compare the model predictions to the observed annual data. This is done in order to highlight the successes and failures of the previous calibrations of the model relative to the data. Later after the estimation, we will re-do this simulation exercise using the estimated parameters to calibrate the model and see whether the model predictions have been improved. Simulated data from the model are generated at the monthly frequency, then aggregated at the annual frequency to make the comparison with the observed annual data. Our observed annual data set goes from 1926 to 2009, which corresponds to 1008 months.

4.3.1 Data and Basics moments

We use the U.S annual real data on consumption growth per capita of non durable goods and services from Robert Shiller Website. The financial data (risk free rate, market return, dividend growth and log price dividend ratio) are also from Shiller's Website, the real values are computed using data on inflation from the same source. The log price dividend ratio is obtained by dividing the real prices by the real dividend and taking the log. The real market return is obtained by subtracting the inflation rate from the nominal market return; the same procedure is applied to obtained the real risk free rate. the growth rates are obtained by taking the differences in the log of the corresponding variables.

The simulation procedure is done as described in appendix A.1. The moments computed from the simulated data will give us an idea of how far the model is close to the data by comparing their values with what can be computed using the recorded data. We also report the results obtained by Constantinides and Ghosh [2011]using the same calibration from Bansal and Yaron [2004] as we do. In general the model is very close to the data as far as we consider moments of the consumption growth, dividend growth and market return processes; the moments computed using the observed data are within the 95 % confidence intervals of the model predictions. But for the risk free rate and the log price-dividend ratio, the model predictions are less convincing and in many cases, the observed data moments fall out of the 95 % confidence interval of the model prediction.

Indeed, the average level of the risk free rate predicted by the model is above what has been recorded in the data, meaning that the representative agent in the model is less risk adverse than the real one, or the preference for the future is higher (a lower time preference) in the reality than what the calibrated model assumes. It means that, the risk aversion coefficient or the preference for the future coefficients (the EIS or the pure discount rate) should be increased in the model. As we see in the BKY calibration, the pure discount rate has been increased and the model yields a 95 % confidence interval which now contains the small observed value of the average risk free rate. So increasing the pure discount factor, and by the way lowering the time preference for the representative agent in the model, reduces the demand for saving and allows the model to generate a low average risk free rate that is closer to what is observed in the data. However, some improvement still need to be done. Indeed, the observed volatility of the risk free rate, the volatility and auto-correlation of the log-price dividend ratio are all out of the 95 % confidence intervals predicted by the model for all the calibrations considered.

Table 4.3.1 also reveals that the BY and the BKY calibrations deliver almost the same results in term of basic moments even though they are quite different. So, by using only the basic moments, we could identify many sets of parameters that could fit the model and in order to discriminate among then we need to add other moment conditions and select among the candidate sets of parameters allowing to fit the basic moments, those that will also be able to fit the added moments.

These observations clearly advocate for the use of an estimation procedure in order to select the parameters that will bring the model closer to the observed data. This exercise will be very important also because the mechanism at play in the model depends on the calibration. Indeed, as it has be pointed by Beeler and Campbell [2012], the relative importance of movements in consumption growth and volatility is not the same in the BY and BKY calibrations. In the former calibration, time variation in the volatility almost plays no role while in the later, it is a key determinant of the level of the equity premium and the volatility of the stock prices.

Table 4.1: Basic moments implied by the model

Moments	Data	BY	BKY
$E(\Delta c)$	0.020	0.018 [0.005; 0.031]	0.018 [0.006; 0.030]
$\sigma(\Delta c)$	0.03	0.028 [0.023; 0.035]	0.028 [0.015;0.044]
$AC1(\Delta c)$	0.359	0.473 [0.245; 0.668]	0.402 [0.168;0.608]
$E(\Delta d)$	0.008	0.010 [-0.034;0.054]	0.018 [-0.032;0.066]
$\sigma(\Delta d)$	0.110	0.112 [0.091;0.137]	0.155 [0.084;0.247]
$AC1(\Delta d)$	0.146	0.367 [0.149;0.570]	0.261 [0.034;0.466]
$E(r_m)$	0.075	0.067 [0.027;0.109]	0.065 [0.019;0.121]
$\sigma(r_m)$	0.200	0.167 [0.139;0.198]	0.205 [0.118;0.321]
$AC1(r_m)$	-0.003	0.017 [-0.203;0.236]	-0.002 [-0.227;0.230]
$E(r_f)$	0.005	0.026 [0.018;0.034]	0.001 [-0.004;0.018]
$\sigma(r_f)$	0.039	0.012 [0.009; 0.017]	0.011 [0.005;0.018]
$AC1(r_f)$	0.563	0.807 [0.671;0.900]	0.806 [0.660;0.908]
$E(p-d)$	3.320	3.007 [2.894;3.118]	3.05 [2.593;3.319]
$\sigma(p-d)$	0.441	0.188 [0.140;0.250]	0.204 [0.106;0.331]
$AC1(p-d)$	0.936	0.654 [0.441;0.811]	0.628 [0.314;0.850]

The table displays the results from 10.000 simulations of 948 months each (to mimic the period 1930-2009) of the long run risks model based on the following BY calibration: $\delta = 0.998, \gamma = 10, \psi = 1.5, \mu_c = 0.0015, \mu_d = 0.0015, \phi = 3, \varphi = 4.5, \rho = 0.979, \varphi_e = 0.044, \sigma = 0.0078, \nu = 0.987, \sigma_w = 0.0000023, \pi = 0$. The statistics are for annually aggregated variables (consumption growth rate, dividend growth rate, market return, risk free rate, log price-dividend ratio). Aggregation has been done through a 12 months summation for growth rates and returns. For the log price-dividend ratio, the annual values have been computed using the year's December values of the price divided by the summation of monthly dividends within the year. The last column reports the result obtained using Bansal et al. [2012a] calibration: $\delta = 0.9989, \gamma = 10, \psi = 1.5, \mu_c = 0.0015, \mu_d = 0.0015, \phi = 2.5, \varphi = 5.96, \rho = 0.975, \varphi_e = 0.038, \sigma = 0.0072, \nu = 0.999, \sigma_w = 0.0000028, \pi = 2.6$.

4.3.2 Stylized facts and Model predictions

Beeler and Campbell [2012] have extended the basic set of moments matching to verified that the long run risks model is coherent with some stylized facts that have been documented by earlier studies. Specifically, they tested 2 stylized facts related to the predictability of cash flows by stock prices:

- log price-dividend ratio predicts excess stock returns and not dividend growth or real interest rate (Campbell and Shiller [1988]; Fama and French [1988]; Hodrick [1992]);
- there is a strong relationship between stock prices and consumption volatility but no link between stock prices and return volatility [Beeler and Campbell, 2012];

As shown in Table 4.3.2, in the data the log price-dividend ratio predicts excess stock returns with a negative and significant sign at 5% confidence level; meaning that an increase in stock prices predicts a drop in future returns. This predictability increases with the horizon; at one year horizon, log price-dividend ratio explained only 6 % of future excess return. This explanation increases to 29 % at 5 year horizon. The predictability of future excess return by the log price-dividend ratio is due to the persistence in the later variable (Cochrane [2001], Chap. 20), but it is not captured by both the BK and the BKY calibrations of the long run risks model: at 5 year horizon, the log price-dividend ratio explains less than 5% on the excess return and the regression's slope even sometime has the wrong sign or is not significant.

For the consumption process, the picture is inverted, the model overstates the predictability of future consumption growth by stock prices : in the data, there is no evidence of predictability of future consumption growth by the log price-dividend ratio while in the BK and BYK calibrations, log price dividend ratio predicts future consumption growth with a significant slope coefficient. Even though the power of the predictability of the future consumption growth by the log price dividend ratio has been greatly reduced from BK to BYK calibration, the slope coefficients still remain significant at 5 % confidence level; what is at odd with the data.

As we know from (Cochrane [2001], sect.20.1.2), log price-dividend ratio only moves if it forecasts excess return or dividend growth or if there is a bubble. In the data, log price dividend ratio rather predicts excess return as we can see in Table 2. The BKY calibration does well in this respect compare to the BK calibration where the log price dividend ratio seems to predict future dividend growth.

Table 4.2: Predictability of excess returns, consumption, and dividends.

	$\hat{\beta}$	$\widehat{R^2}$	$\beta(50\%)$	$\beta(50\%)$	$R^2(50\%)$		$\%\left(\widehat{R^2}\right)$	
	data	data	BY	BKY	BY	BKY	BY	BKY
$\sum_{j=1}^{J} r^a_{m,t+j} - r^a_{f,t+j} = const. + \beta(p^a_t - d^a_t) + \varepsilon_{t+J}$								
1Y	-0.107 -3,446	0,059	0.0614 0.712	0.079 0.925	0.006	0.010	0.944	0.908
3Y	-0.304 -4.078	0.222	-0.171 -1.082	-0.219 -1.315	0.024	0.026	0.950	0.884
5Y	-0.440 -4.420	0.293	-0.299 -1.298	-0.447 -1.365	0.021	0.040	0.960	0.862
$\sum_{j=1}^{J} (\triangle c^a_{t+j}) = const. + \beta(p^a_t - d^a_t) + \varepsilon_{t+J}$								
1Y	0.001 0.165	0.0003	0.067 4.845	0.056 5.694	0.225	0.150	0	0.009
3Y	-0.002 -0.116	0.0002	0.175 5.361	0.108 1.713	0.267	0.114	0.001	0.027
5Y	-0.009 -0.562	0.005	0.220 4.648	-0.089 -1.269	0.219	0.089	0.010	0.086
$\sum_{j=1}^{J} (\triangle d^a_{t+j}) = const. + \beta(p^a_t - d^a_t) + \varepsilon_{t+J}$								
1Y	-0,003 -0,133	0.000	0.262 3.217	0.317 4.898	0.118	0.198	0.002	0.001
3Y	0,024 0,436	0.003	0.663 3.955	0.351 2.301	0.165	0.085	0.015	0.068
5Y	0,043 0,612	0.007	0.673 3.622	0.737 1.581	0.145	0.062	0.047	0.157

There is a great improvement in the model implications from BK calibration to BKY own. However, the BKY calibration of the model failed to capture excess return predictability by the log price dividend ratio that is observed in the data. Each column represent the slope and the t-stat of the predictive regression of future excess return, consumption and dividend growth on the log price dividend ratio for different horizons (1 year, 3 years and 5 years). The R-square are also provided. The last column represents the percentage of sample simulated R-squared that are below the R-squared observed in the data; a percentile below 5% or above 95 % correspond to a one side test rejection of the model at 5% significant level.

Some of these stylized facts have been shown to be a failure of the BY and BYK calibrations to match the data. But, as the matching of the data has been greatly improved from the BY calibration to the BYK own, there is a hope that the failure highlighted by Beeler and Campbell [2012] and Garcia et al. [2008] could be just a failure of finding the good set of parameters for the model rather than a rejection of the economic mechanism presented in the model as the explanation of many asset pricing puzzles. So, the addition of the predictive moment constraints implied by these stylized facts in the GMM procedure should improved the model prediction through a better selection of the set of parameters. The derivation of the theoretical annual moment involving the log price-dividend ratio is not an easy task because the annual values of the log price-dividend ratio are obtained by dividing the end of year (December) monthly prices by the sum of the monthly dividends within the year. So,

$$p_t^a - d_t^a = \log P_{ht} - \log \sum_{j=0}^{h-1} D_{ht-j} \tag{4.3.1}$$

Then using the following approximation presented in Bansal et al. [2007b]. [8]:

$$\log \sum_{j=0}^{h-1} D_{h(t+1)-j} \approx \log D_{ht} + \log h + \sum_{j=0}^{h-1} \frac{j+1}{h} \triangle d_{h(t+1)-j} \tag{4.3.2}$$

Equation (4.3.1) can be rewritten as follow[9]:

$$p_t^a - d_t^a = \mu_d^a + x_t^a + A_{2m}\sigma_{ht}^2 - \pi\eta_t^a - \varphi_d u_t^a \tag{4.3.3}$$

Where

$$\mu_d^a = A_{0m} - \log h - \sum_{j=0}^{h-1} \left(\frac{j+1}{h}\right) \mu_d, \quad x_t^a = A_{1m} x_{ht} - \sum_{j=0}^{h-1} \left(\frac{j+1}{h}\right) \phi x_{ht-j-1}$$

$$\eta_t^a = \sum_{j=0}^{h-1} \left(\frac{j+1}{h}\right) \sigma_{ht-j-1}\eta_{ht-j}, \quad u_t^a = \sum_{j=0}^{h-1} \left(\frac{j+1}{h}\right) \sigma_{ht-j-1} u_{ht-j}$$

Equation (4.3.1)) allows to write the annual values of the log price-dividend ratio in terms of monthly values of the log price-dividend ratio and monthly dividend growth rate. This formula will enable to express the annual log price-dividend ratio in term of monthly shocks and later to compute the theoretical moments mentioned in the stylized facts that will be added to the previous usual set of moments. Equation (4.3.3) shows that the annual value at year t of the log price-dividend ratio depend on the monthly values of the persistent component of the consumption growth process, on the December volatility, on the monthly short run consumption shocks and monthly dividend growth shocks. We can see that negative shocks will increased asset prices while past month positive shocks will lower the level log price-dividend ratio.

To compute the theoretical value of the slope in the predictive regressions, we need to express the annual value of the cash flows in term of monthly processes. Future annual excess returns could be expressed in term of present year last month states variables (x_{ht} and σ_{ht}) and weighted sums of within years future monthly consumption and dividend shocks ($\sigma_{h(t+j)-p-1}\eta_{h(t+j)-p}$,

[8] This approximation is exact when the arithmetic mean and the geometric mean are equal; it is also assumed in Hansen et al. [1996]. It could also be obtained through a Taylor approximation around the mean of the one period (log) consumption growth and assuming that mean to be equal to zero, as shown in Ghattassi and Meddahi [2012] .

[9] This expression of the aggregate log price-dividend ratio is mainly used to compute the theoretical moments, but for simulations we rather used the following intermediary formula: $p_t^a - d_t^a = z_{m,ht} + \sum_{j=0}^{h-1} \triangle d_{ht-j} - \log h - \sum_{j=0}^{h-1} \left(\frac{j+1}{h}\right) \triangle d_{ht-j}$

$\sigma_w w_{h(t+j)-p}, \sigma_{h(t+j)-p-1} e_{h(t+j)-p}, \sigma_{h(t+j)-p-1} u_{h(t+j)-p})$ as follow:

$$\sum_{j=1}^{J} \left(r^a_{m,t+j} - r^a_{f,t+j} \right) = \mu_r^{a,J} + \phi_x^{a,J} x_{h(t-1)} + \phi_\sigma^{a,J} \sigma^2_{h(t-1)} + e_t^{a,J} + w_t^{a,J} + \varphi_d u_t^{a,J} + \pi \eta_t^{a,J} \quad (4.3.4)$$

where $\mu_r^{a,J}$, $\phi_x^{a,J}$, $\phi_\sigma^{a,J}$, $e_t^{a,J}$,$w_t^{a,J}$,$u_t^{a,J}$,η_t^a, Λ_0, Λ_1,Λ_2, $\beta_{m,w}$ and $\beta_{m,e}$ are given in Appendix A.3.

We choose here to focus on the predictive regression of the future excess return on the log price-dividend ratio because this is the key mismatch that both the BK and BKY calibrations failed to capture. For the other predictive regressions (with future consumption or dividend growth as the dependent variables), the BKY calibration delivered a pattern that seems to be close to the data.

Concerning the second stylized fact, the derivation of analytical formulas is more tricky because, as in Beeler and Campbell [2012], we need to compute the realized volatility suggested by Bansal et al. [2012a]. For that, the first step is to run an AR(1) regression of each variable of interest y_t (consumption or dividend growth or excess return):

$$y_{t+1} = b_0 + b_1 y_t + u_{t+1} \quad (4.3.5)$$

Then we can compute the K-period realized volatility as the sum of the absolute values of the residuals over the K periods:

$$Vol_{t+1,t+K} = \sum_{k=1}^{K} \mid u_{t+k} \mid \quad (4.3.6)$$

and finally we regress the log of K period realized volatility on the present log price-dividend ratio:

$$log[Vol_{t+1,t+K}] = \alpha_c + \beta_c (p_t^a - d_t^a) + \xi_{t+K} \quad (4.3.7)$$

Keeping track of analytical formulas along this procedure becomes difficult; that is why we will use it (this procedure) as an out of sample test to see whether the new calibration derived from estimation performs better than the previous one as far as this predictability is concerned.

Table 4.3: Predictability of the volatility of excess returns , consumption, and dividends.

	$\hat{\beta}$ data	$\widehat{R^2}$ data	$\beta(50\%)$ BY	$\beta(50\%)$ BKY	$R^2(50\%)$ BY	$R^2(50\%)$ BKY	$\%\left(\widehat{R^2}\right)$ BY	$\%\left(\widehat{R^2}\right)$ BKY
Excess return volatility predictability								
1Y	0.015 / 0.553	0.003	-0.040 / -0.5732	-0.090 / -1.756	0.007	0.027	0.351	0.173
3Y	0.008 / 0.106	0.0003	-0.149 / -1.102	-0.256 / -1.480	0.015	0.057	0.077	0.038
5Y	-0.012 / -0.114	0.0004	-0.225 / -0.812	-0.415 / -2.292	0.022	0.073	0.067	0.040
Consumption growth volatility predictability								
1Y	-0.013 / -3.615	0.120	-0.006 / -1.019	-0.016 / -1.474	0.007	0.033	0.996	0.864
3Y	-0.043 / -4.138	0.306	0.012 / 1.010	-0.034 / -1.996	0.016	0.072	0.9995	0.928
5Y	-0.072 / -4.119	0.397	0.0280 / 1.065	-0.049 / -1.899	0.021	0.091	0.999	0.9403
Dividend growth volatility predictability								
1Y	-0.025 / -1.409	0.020	0.022 / 0.837	-0.118 / -1.541	0.007	0.033	0.763	0.386
3Y	-0.069 / -1.158	0.031	-0.095 / -1.060	-0.316 / -2.192	0.0147	0.069	0.667	0.327
5Y	-0.141 / -1.165	0.054	-0.103 / -1.035	-0.318 / -1.425	0.021	0.088	0.720	0.385

There is a great improvement in the model implications from BK calibration to BKY concerning the consumption growth volatility predictability. However, the BKY calibration of the model overstated the excess return volatility predictability by the log price dividend ratio that is observed in the data. Each column represent the slope and the t-stat of the predictive regression of future excess return, consumption and dividend growth on the log price dividend ratio for different horizons (1 year, 3 years and 5 years). The R-square are also provided. The last column represents the percentage of sample simulated R-squared that are below the R-squared observed in the data; a percentile below 5% or above 95 % correspond to a one side test rejection of the model at 5% significant level.

4.4 Estimation procedure and Simulations

In the data, the latent variables (x_t and σ_t^2) are not observable but we observe the log price-dividend ratio (z_m), the real market return (r_m), the real risk free rate (r_f), the market dividend growth rate (g_d) and the consumption growth rate (g). We use the Euler equation and the affine link between the observed variables and the latent ones to set up some moment restrictions that will be used for the GMM procedure. Our approach is similar to Constantinides and Ghosh [2011] who solved for the 12 parameters of the model by imposing 15 unconditional moment restrictions. But contrary to them we choose to remain closer to how the original LRR model works by expressing the theoretical moments taking temporal aggregation into account to match the annual data ($h = 12$). Indeed it is usually assume during calibrations that the long run risks model operates at the monthly frequency, but the monthly data are reputed to be of poor quality especially because of seasonality and other measurement problems Wilcox [1992]. We also estimate the model assuming that the agent decisions are made at the same frequency as the econometrician observes the data (meaning at the annual frequency, $h = 1$), but we reject the model in that case. This shows that the temporal aggregation also play an important role for the LRR model to match the data.

The long run risks model we consider here has 13 parameters[10] The moment restrictions are made of 16 basic moment restrictions whose empirical counterparts are summarized in Table 1. and are known to be matched pretty well by the previous calibrations of the model; except the variances of the risk free rate and of the log price dividend ratio. In a first step, we will only use those restrictions for the estimation. In a second step, we add 3 other moment restrictions to capture the predictability (at 1, 3 and 5 years horizon) of the cash flows (future excess returns) on asset prices (log price-dividend ratio) summarized in Table 2. and which have been presented as a failure of the previous calibrations of the model to match the data. Finally, in a third step we extended the previous 19 moment restrictions to 22 (resp. 25) moment restrictions by adding 3 other moment restrictions for the predictive regressions of consumption (resp. dividend) growth on log price-dividend ratio.

4.4.1 Weak identification evidence in the LRR model

In this section, we will show that some parameters in the long run risk model are weakly identified, mainly because of the sample size, and thus the standard inference procedure, which relies on the delta method to obtain the confidence intervals and in testing is not valid. The type of weak identification that we observe in the LRR features the special case described by [Andrews and Cheng 2012, 2014].

Following Newey and McFadden [1986] (Sec. 2.2.3), Stock and Wright [2000], we defined a parameter θ as being identified if $\theta = \theta_0$ is the unique solution of the moment condition $E(g(y_t, \theta)) = 0$[11] for $\theta \in \Theta$, where the expectation is taken with respect to the true distribution

[10] In the model considered by Constantinides and Ghosh [2011], the parameter π which captures the exposure of dividend growth to short run consumption shocks is ignored. The time aggregation parameter h could also be estimated, that will extend the number of structural parameters to be estimated to 14.

[11] The vector function $g(y_t, \theta)$ can be seen as the difference between the empirical moment from the data evaluated

(for which $\theta = \theta_0$). Some parameters can be weakly identified while others are strongly identified; we can then split the parameter vector as follow $\theta = (\alpha, \beta)$, where α is the sub-vector of weakly identified parameters and β is the sub-vector of strongly identified parameters. This means that the moment conditions are zero at $\theta_0 = (\alpha_0, \beta_0)$, but are also very nearly zero for $\alpha \neq \alpha_0$: In other words, the population objective function is steep in β around β_0, but nearly flat in α (Stock and Wright [2000], Sect 2.3).

Andrews and Cheng [2012] presented another specific situation where some parameters of the model might be weakly identified when the true value of a given parameter is close to 0[12]. Then, they split the vector of parameters into three components : $\theta = (\beta, \xi, \pi)$ where β and ξ are strongly identified but π becomes weakly identified when $\beta_0 = 0$ and there is a drifting sequence of parameters, depending on the sample size ($\theta_n = (\beta_n, \xi_n, \pi_n)$ that converges to θ_0 such that $\beta_n \neq 0$ and $n^{1/2}\beta_n \rightarrow b \in R^{d_\beta}$(category I(b) in their paper). The situation we encounter in the LRR model really looks like that one but we should be cautious about this categorization. Indeed, if we assume that the true value of volatility of the volatility (σ_w) is equal to 0 then the persistence of the volatility (ν) will not be identified in population; and from the literature on the LRR model we know that the true value of the volatility of the volatility must be different from 0; if not the volatility of the consumption growth will be constant and thus the equity premium will not be time varying. But from the subsequent calibrations of the model[13], it seems that it is close to 0, which implies the weak identification of the persistence of the volatility (ν) in finite sample even if it could be identified in population.

By the same argument, we can deduce that the closeness to 0 of the true value on the vector of parameters (σ, σ_w) induces the weak identification of the vector of parameters $(\phi, \varphi_d, \varphi_e, \rho, \nu, \pi)$ and the remaining parameters $(\delta, \gamma, \psi, \mu_c, \mu_d)$ are strongly identified. Hence we have the decomposition of our vector of parameters which satisfies the Assumption A in Andrews and Cheng [2012], and we can rely on their robust procedure for estimation and inference. Furthermore, we observe that when the EIS become greater (let say more than 10), the objective function becomes almost insensitive with respect to that parameter changes[14] and thus the jacobian matrix is of reduced rank. So, this parameter might also be weakly identified in a part of the parameters

at the true value of the parameter θ_0 and the population moment evaluated at θ: $g(y_t, \theta) = \bar{m}(y_t, \theta_0) - m(Y_t, \theta)$, where y_t denotes the sample observation of the random variable Y_t

[12] The value 0 has been used without lost of generality and can be achieved by using an appropriate re-parametrization

[13] It was 2.3 e-06 in the BY calibration and 2.8 e-06 in the BYK calibration

[14] More precisely, the moment conditions depend on the EIS only through the coefficient of the affine functions and it is those coefficients which are almost insensitive to the changes of ψ as it becomes high

space.

4.4.2 Simulations: finite sample approximation of the asymptotic distribution

As we know that commonly used testing procedures, in particular those robust to weak identification (on which we will come back in later), are derived asymptotically (Dufour, 2009), to evaluate the accuracy of our unconditional moment matching procedure in finite sample, we run a Monte Carlo experiment, where we simulate the model at various sample sizes and plot kernel density of the Gmm objective function. On the same plot (4.7) , we also represent the corresponding chi-squared limiting distribution. When the sample size is small (e.g: T=84) as what we have in the data, the finite sample approximation of the asymptotic distribution of the Gmm objective function is very poor; the right tail is very huge and that will lead to an over-rejection of the model. It means that while using the asymptotic distribution to evaluate the model in this case, we will end up wrongly rejecting the model more often that usual. So the rejection probability of a 95 % confidence test will be much more greater than the size of the test and the true size of the obtained confidence set will be much lower than 95 %. However, for a large sample (e.g: T=4200) , we see that the asymptotic test becomes conservative because the asymptotic distribution of the Gmm objective distribution is dominated by Chi-squared limiting distribution; meaning that rejection rate for a true model is lower than the significant level of the test. this conservativeness is the usual drawback of this against-weak identification robust methods[15].

4.5 Data and results

4.5.1 Estimation results

We use the Continuously Update GMM Estimator (CUE) for the estimation because it is partially robust to weak identification Stock et al. [2002]. We ran the estimations for four sets of moment conditions: (i) The first one is made of 16 moment restrictions mainly targeting the basic moments summarized in Table 1 and does not include the predictive moments; we call it the restricted Gmm (RCUGmm). (ii) The second one, named CUGmm, is made of the moment conditions considered in the restricted Gmm and the excess returns predictive moments (covariance between excess returns at different horizons and the log price dividend ratio). (iii) The

[15]See Stock et al. [2002]

third is made of all the previous moment conditions and the consumption growth predictive moments (covariance between consumption growth at different horizons and the log price dividend ratio); we call it the Extended Gmm (ECUGmm). (iv) The last one is made of all the moment restrictions in (iii) and 3 predictive moments of the dividend growth by the log-price dividend ratio, we call it EECUGmm .

In a first place, we follow the usual procedure in a Gmm estimation, meaning that we minimize the objective function and then for the achieve minimum we test the model by comparing that minimum to a 95 % quantile of a chi-square distribution with the number of moment equations minus the number of estimated parameters as degree of freedom. In a second place, we follow the procedure describe in Stock and Wright [2000] when there is weak identification, which consist at fixing the weak identified parameters and estimating the remaining ones using continuously updated Gmm estimator with the optimal weighting matrix. The objective function value could then be compared to a chi-squared distribution and the model will not be rejected if we are able to find some fixed values of the weakly identified parameters for which the objective function at the minimum will be lower to a threshold given by the chi-square quantile at 95% level and the degree of freedom being equal to the number of moment restrictions minus the number of parameters estimated[16]. This method also allows to construct the confidence interval of the fixed parameter, by collecting all the values for which the model is not rejected. We repeated the procedure by fixing one parameter each time.

Table 4.4 shows that the model is not rejected at 5 % confidence level for two of the four set of moment conditions that we considered. In fact, the model is not rejected when we consider the basic moments made of the first and second order moments of the cask flows and stock prices variables, and the covariance between excess return and the log price dividend ratio. But it is rejected when we add the other predictive restrictions. So the LRR model is able to match the first and second order moment of asset prices and cash flows data, hence to solve the equity premium and risk free rate puzzles. It can also reproduce the predictability of excess return by the price-dividend ratio. But, it fails at matching simultaneously the observed level predictability of cash flows (consumption and dividend growth) by the price dividend ratio. In fact, there is a huge tension for the model to capture the excess return predictability without overstating the cash flows (consumption growth and dividend growth) predictabilities compare to what is observed o real data. We tried man different starting values and still in the last two

[16]See appendix B.

cases (ECUGmm and EECUGmm) we where not able to achieve a set of parameters for which the model could not be rejected at the standard confidence level.

Some parameters (δ, μ_c, μ_d,ρ,ν,σ) are estimated very precisely and their values are stable whatever the moments restriction we considered. The value of the pure discount rate is very closed to 1, the monthly consumption growth is around 0.19 %. For the RCUGmm, where there is not predictability moment in the restrictions, the model is not rejected at the 5 % critical error probability. The estimates of γ and ψ are very large but also not significant. this is also the case for many parameters describing the dividend growth process (ϕ,φ_d,π). When the model is not rejected, the estimated level of risk aversion coefficient (2.7) is quite below the one that is usually used for calibration (10). The estimate of the EIS (1.27) is above 1 as usual assumed in the model, even though the 95 % confidence interval contains values below 1.

We find some difficulties while computing the confidence intervals for the minimands we obtained. Indeed, applying the formula in equation 4.10.6, the matrix we computed at the minimum was not invertible ($G(\hat{\theta})'\Omega(\hat{\theta})^{-1}G(\hat{\theta})$). The singularity of this matrix comes from the fact that the estimate of the jacobian matrix is not of full rank. We think that this lack of full rankness does not come from the non identification of the model, if not this would have happened for every vector of parameters. But instead, it comes from the fixed point feasibility condition we highlighted in section 2.2. Indeed, the numerical computation of the Jacobian requires two points where the objective function will be evaluated and when the minimand is close to the feasibility bound this can be difficult do[17].

Figures 4.8 and 4.9 show respectively the confidence intervals of γ and ψ obtained with the CUGmm restriction setup, following the weak identification robust procedure described in 4.10.2. We see that those confidence intervals are different from the ones obtained in 4.4. The confidence intervals for γ and ψ are included in $[1; 20]$. No value below 1 is admitted in the 95% confidence interval of ψ; this is in adequacy with the hypothesis made in the LRR model that the value of the EIS is above 1. However, this confidence sets can still be extended because for computational reason[18], we took values of γ and ψ ranging between 0.1 and 20, but we can see from figures 4.8 and 4.10, that some values of ψ even above 50 can still be admitted.

[17]We still need to work on this point.

[18]For each value of the fixed parameter, we run several optimizations restarting from the previous obtained minimand until convergence. This can take more than several minutes for each value of the fixed parameter.

Table 4.4: Results of the Gmm estimations

	RCUGmm	CUGmm	ECUGmm	EECUGmm
δ	0.99738 [0.915 ;1.0798]	9.999999e-01 [0.994 ; 1.006]	0.999815 [0.9998146 ; 0.9998164]	9.997412e-01 [0.9994 ;1.0001]
γ	21.659 [-132.336 ;175.654]	2.685 [-7.346e-02 ; 5.443]	5.374 [-10.554 ; 21.302]	5.128 [4.671 ; 5.586]
ψ	6.812 [-632.82 ;646.44]	1.250 [-0.922 ; 3.423]	1.475 [3.938e-01 ; 2.557]	1.4816 [1.480; 1.484]
μ_c	1.93e-03 [1.567e-03 ;2.292e-03]	2.082e-03 [1.788e-03; 2.375e-03]	1.905e-03 [1.903e-03 ; 1.907e-03]	1.900e-03 [1.657e-03 ;2.142e-03]
μ_d	2.066e-03 [4.298e-04 ;3.703e-03]	2.398e-03 [1.129e-03 ;3.667e-03]	2.824e-03 [1.563e-03 ;4.085e-03]	2.437e-03 [1.770e-03 ;3.104e-03]
ϕ	20.631 [-121.87 ;163.132]	2.648 [0.751 ;4.545]	3.880 [-0.880 ;8.640]	3.439 [3.020 ;3.859]
φ_d	2.379 [-1.232 ;5.990]	-4.053 [-6.388 ;-1.718]	17.699 [-72.341 ;107.740]	17.471 [17.011 ;17.930]
ρ	0.9944 [0.988 ; 1.001]	0.9965 [0.988 ; 1.005]	0.957 [0.879 ; 1.034]	9.475e-01 [9.005e-01 ; 9.945e-01]
φ_e	2.342e-03 [-1.141e-02 ; 1.609e-02]	1.664e-02 [-9.714e-03 ; 4.300e-02]	0.355 [-1.642 ; 2.353]	6.368e-01 [2.942e-01 ; 9.794e-01]
σ	7.865e-03 [4.655e-03 ; 1.108e-02]	4.805e-03 [2.504e-03 ; 7.106e-03]	8.136e-04 [0 ;4.277e-03]	5.333e-04 [1.461e-04 ; 9.205e-04]
ν	9.208e-01 [8.867e-01 ; 9.548e-01]	9.900e-01 [0.975 ;1.005]	0.997 [0.995 ;0.998]	9.972e-01 [9.965e-01 ; 9.980e-01]
σ_w	1.071e-04 [0 ;1.426e-03]	2.922e-05 [0 ;1.494 e-04]	5.4872e-07 [2.6126e-07 ; 8.3617e-07]	2.005e-07 [0 ; 2.336e-07]
π	3.956e-01 [-1.711 ;2.502]	0.8172 [-2.872 ;4.506]	3.422 [-209.713 ;216.557]	1.823 [1.363 ;2.283]
$TJ_T(\widehat{\theta_p^u})$	6.715 0.0816	10.862 0.0927	23.252 0.00565	38.626 0.00012

This table presents the parameter's estimates and their 95 % level confidence intervals derived from the minimization of the GMM criterion function. The last row gives the value of the minimum achieved and the corresponding p-value. 13 parameters are estimated. For the RCUGmm we have 16 moment restrictions; for the CUGmm, 19 moment restrictions; for the ECUGmm, 22 moment restrictions; and for the EECUGmm, 25 moment restrictions. So the p-values are computed respectively with respect to a $\chi^2(3)$, $\chi^2(6)$, $\chi^2(9)$ and $\chi^2(12)$.

4.5.2 Predictability

For reminding, one of the goals we had at the beginning was to find through an efficient procedure, a set of parameters that could enable (if possible) to restore the predictability failure highlighted by Beeler and Campbell [2012] in the LRR model. So, we now examine the model implied moments and the predictability of future cash flows by the log price dividend ratio when the model is calibrated at the value we find in the previous estimations.

Table 4.6 shows that concerning the basic moments, all the four restrictions estimates did equivalently good job at generating data whose first and second order moments meet the observed ones. The mean of the consumption growth is well around 2% with a volatility around 3%. The market return is at 6% with a risk free rate below 1 %.

Table 4.6 also highlights the tension that the model faces while attempting to match all the data characteristics. Indeed, we see that the estimate of the first moment restrictions setup (RCUGmm) does almost the same job as the BYK calibration, mainly on one hand it fails at capturing thevolatility of the risk free rate and its auto-correlation. On the other hand, the volatility of the market return is a bit overstate, but the volatility of the price-dividend ratio and its auto-correlation are captured. When we move to the second moment restrictions setup (CUGmm), we observed that the volatility of the risk free rate and the volatility of the market return have been improved, but then the risk free rate auto-correlation is still not captured and the dividend volatility is a bit overstated. For the last two moment restrictions set up (ECUGmm and EECUGmm), we see that the volatility of the risk free rate has been improved again and its auto-correlation is now captured. But the consumption and dividend growth auto-correlations are overstated and the price dividend auto-correlation is understated.

The results concerning the predictability implied by the model are even more interesting. Indeed as Table 4.5 shows, for the first restriction set up (RCUGmm), there is a slight predictability observed for excess return and almost no predictability for the consumption growth; both can hardly but still be rejected as being equal to the ones observed in the data at the 95 % confidence level. For dividend growth, there is some predictability in the simulated data, which is strongly rejected as being equal to the one observed in the data. While in the second restriction setup (CUGmm), where the predictability of the excess return by the log price dividend ratio has been targeted, there is indeed a predictability pattern emerging, which mimics the one observed in the data. However, we see also that the consumption growth and the dividend growth both became highly predictable, which is not the case in the data. For the predictability of the excess return by the log price-dividend ratio, the equality of the model's R-squared to the sample counterpart is not empirically rejected at the 5 % confidence level. When the model is calibrated at the CUGmm estimates, at 1 year horizon, the log price-dividend ratio explains 2 % of the future excess return variability and this explanation increases to 7 % at the 5 years horizon. Furthermore, as in the observed data, the slope in the predictive regression has a negative sign and its significance increases with the horizon of prediction. Similarly, in the third restriction setup (ECUGmm) where the predictability of the excess returns and the consumption growth has been targeted, we see that the predictability of the excess return is almost similar to the one obtained in the previous restriction setup and the predictability of the consumption growth is more closer two the one observed in the data (it is declining with the horizon). There is also an improvement for the predictability of the dividend growth. In the last restriction setting (EECUGmm), The predictability of excess return by the log price dividend ratio has continue to improve, moving from 3 % part in the explanation of the future excess return by the log price dividend ratio at one year horizon to 11 % of that explanation at 5 years horizon. the predictabilities of the consumption growth and the dividend growth by the log price dividend ratio are both declining with the horizon.

We also look at the volatility predictability to see whether the model could reproduce the predictability of the volatility of the consumption growth by the log price dividend ratio observed in the data. As we can see from Table 4.10.2, there is an improvement when we move from the basic moment set up to the extended Gmm estimation set up. The median r-squared has increased compare to the BYK calibration. However, there is an over-state of the predictability of excess return volatility by the log price-dividend ratio.

These results say that the long run risk model is very good at reproducing many features of the observed data when they are targeted. But as we saw this could come at the cost of introducing more bias in mimicking the non targeted ones, and thus increasing the targeted moment does not seem to be the solution. This cast some doubt about the reliability of this model to be used for forecasting and policy recommendations.

To resume, the results obtained in the empirical part still need to be examined and could be improved by the choice of a better starting value in the optimization program. Indeed, because of the complexity of the objective GMM criterion, the presence of non-feasible regions of parameters and also due to the identification problem highlighted, the finally selected set of parameters is obviously a local minimum and depends on the starting point. The choice of another starting point could yield to a better parameter vector, for which the moments restriction will be validated and the moments predictability improved. Another possible source of bias, could be the approximations on which we relied. Indeed, there are two main approximations that we made in order to solve the model and later to make a bridge between the monthly simulated data in the model and the annualized observed data. The first approximation is the Campbell-Shiller approximation of log return that allows to expressed asset log returns in terms of log price dividend ratio and therefore as linear function of states variables. By comparing the model implications using this approximation and the computational power, Bansal et al. [2007b] argued that this approximation has little impact on the results. The second approximation of the arithmetic mean by the geometric mean in order to derived analytical expressions of annualized variables in terms of monthly counterparts. Since the one-month growth rate of cash flows or asset prices are usually small in magnitude, this approximation based on a Taylor expansion could be quite reliable to compare first order moments but could break down when the higher order terms in the Taylor expansion are no more negligible (Campbell et al., 1996-sect.1.4).

4.6 Conclusion

The Long Run Risks model explains the long run behavior of asset prices by the presence of a persistence movement in consumption growth, a time varying volatility, and a consumer preference structure described by the Epstein-Zin recursive utility function with an EIS greater than 1. The calibration of this model has remarkably well performed as far as the matching of the macro-finance first and second order moments is concerned. However, this model has recently received some criticism concerning the ability of the model to match the predictability of cash flows by asset prices observed in the data. More specifically it has failed to match the predictability that price-dividend has on the future returns. The correction of this failure and the rehabilitation of the model require to go beyond the calibration and to provide an estimation procedure that will find the parameter's set which brings the model closer to the data and restore the predictability. The difficulties being that on one hand, the return on the aggregate consumption asset embedded in the Euler equation of a representative agent with an Epstein-Zin recursive utility function is not observable; on the other hand the states variables driving the risk of consumption growth are not observable and even difficult to estimated in a reduced form model because of bias induced by the time aggregation. However, by assuming an affine link between stock prices and the states variables, and using the Euler equation, we can express the cash flows and the stock prices moments in terms of the model underlying parameters. We apply this procedure to derive the analytical expression of a set of moment conditions and apply the GMM method to estimate the structural model's parameters.

Empirically, we find that using the standard asymptotic theory, the model succeeded to pass the test of validation of the moment conditions when we target only the basic moments or extended the basic moments to capture predictability of future excess returns by the price dividend ratio, but it became rejected when the other predictability moments are added. We also saw that the LRR model faces a huge tension while trying to match all the stylized facts highlighted in the literature. Furthermore, even considering the best fitting of the model achieved while taking into account the predictability concerns, it is still very difficult for the model to match the quantitative levels observed in the data, even though the qualitative behavior of the model is very similar in the long run to what appears in the data. However, the parameters estimates achieved using extended moment conditions allow to improve the model predictions and to bring them closer to the observed data.

4.7 Appendix

4.7.1 Model solution : Case of $EIS \neq 1$

- For the wealth portfolio,

$$A_1 = \frac{1-\frac{1}{\psi}}{1-k_1\rho},\ A_2 = \frac{0.5[(-\frac{\theta}{\psi}+\theta)^2 + (\theta k_1 A_1 \varphi_e)^2]}{\theta(1-k_1\nu)}$$

$$A_0 = \frac{\log\delta + (1-\frac{1}{\psi})\mu_c + k_0 + k_1 A_2 \sigma^2(1-\nu) + 0.5\theta k_1^2 A_2^2 \sigma_w^2}{1-k_1}$$

For the price-consumption ratio to produce a positive response to an increase in the long run consumption growth, A_1needs to be positive; that means the EIS (ψ) should be greater than one. By the same way, for the price-consumption ratio to decrease with an increase in the consumption volatility, we need that θshould be negative; which correspond to an EIS greater than 1 given that the risk aversion is greater than 1.

- For the market portfolio,

$$A_{1,m} = \frac{\phi - \gamma + (\theta-1)(k_1\rho - 1)A_1}{1-k_{1,m}\rho}$$

$$A_{2,m} = \frac{(\theta-1)(k_1\nu-1)A_2 + 0.5[(\pi-\gamma)^2 + \varphi_d^2 + ((\theta-1)k_1A_1 + k_{1,m}A_{1,m})^2\varphi_e^2]}{1-k_{1,m}\nu}$$

$$A_{0,m} = \frac{\theta\log\delta - \gamma\mu_c + (\theta-1)[k_0 + (k_1-1)A_0 + k_1A_2\sigma^2(1-\nu)]}{1-k_{1,m}}$$

$$\frac{+k_{0,m} + \mu_d + k_{1,m}A_{2,m}\sigma^2(1-\nu) + 0.5[(\theta-1)k_1A_2 + k_{1,m}A_{2,m}]^2\sigma_w^2}{1-k_{1,m}}$$

- For the risk free asset,

$$A_{1,f} = \gamma + (\theta-1)(1-k_1\rho)A_1$$

$$A_{2,f} = -[(\theta-1)(k_1\nu-1)A_2 + 0.5\left(\gamma^2 + (\theta-1)^2k_1^2A_1^2\varphi_e^2\right)]$$

$$A_{0,f} = -\theta\log\delta + \gamma\mu_c - (\theta-1)[k_0 + (k_1-1)A_0 + k_1A_2(1-\nu)\sigma^2] - 0.5(\theta-1)^2k_1^2A_2^2\sigma_w^2$$

Because the price consumption ratio is not observable, the values of the affine function coefficients are obtained numerical by solving a fixed point problem to get the value of $\bar{z}$ and then compute the k's values and later solve for the coefficient values.

4.7.2 Model solution : Case of $EIS = 1$

The value function is given by:

$$V_t = C_t^{1-\delta}\left(E_t\left(V_{t+1}^{1-\gamma}\right)\right)^{\frac{\delta}{1-\gamma}}$$

which implies the following log price-consumption ratio:

$$vc_t = log(V_t/C_t) = \frac{\delta}{1-\gamma} log\left(E_t\left[exp\left((1-\gamma)\left(vc_{t+1}+\Delta c_{t+1}\right)\right]\right)\right.$$

We conjecture that the log price-consumption ratio can be expressed as an affine function of the state variable, meaning that : $vc_t = \tilde{A}_0 + \tilde{A}_1 x_t + \tilde{A}_2\sigma_t^2$. Given the conditional normality of all the process involved, we have that:

$$\begin{aligned}
vc_t &= \frac{\delta}{1-\gamma}\left\{E_t\left[\left((1-\gamma)\left(vc_{t+1}+\Delta c_{t+1}\right)\right]+0.5(1-\gamma)^2 Var_t\left(vc_{t+1}+\Delta c_{t+1}\right)\right\}\right. \\
&= \delta\left(\tilde{A}_0+\tilde{A}_2(1-\nu)\bar{\sigma}^2+\mu_c+0.5(1-\gamma)(\tilde{A}_2\sigma_w)^2\right)+\delta\left(1+\tilde{A}_1\rho\right)x_t \\
&+ \delta\left(\nu\tilde{A}_2+0.5(1-\gamma)(\varphi_e\tilde{A}_1)^2+0.5(1-\gamma)\right)\sigma_t^2
\end{aligned}$$

By identification, we get that:

$$\begin{aligned}
\tilde{A}_0 &= \delta\left(\tilde{A}_0+\tilde{A}_2(1-\nu)\bar{\sigma}^2+\mu_c+0.5(1-\gamma)(\tilde{A}_2\sigma_w)^2\right) \\
\tilde{A}_1 &= \delta\left(1+\tilde{A}_1\rho\right) \\
\tilde{A}_2 &= \delta\left(\nu\tilde{A}_2+0.5(1-\gamma)(\varphi_e\tilde{A}_1)^2+0.5(1-\gamma)\right)
\end{aligned}$$

Then,

- For the wealth portfolio,

$$\tilde{A}_1 = \frac{\delta}{1-\rho\delta},\ \tilde{A}_2 = \frac{0.5\delta\left(1-\gamma\right)\left(1+(\tilde{A}_1\varphi_e)^2\right)}{1-\nu\delta}$$

$$\tilde{A}_0 = \frac{\delta}{1-\delta}\left[\mu_c+\tilde{A}_2(1-\nu)\bar{\sigma}^2+0.5(1-\gamma)\tilde{A}_2^2\sigma_w^2\right]$$

- For the market portfolio (we still use the Campbell and Shiller [1988] approximation for the log return on the market portfolio)

$$\tilde{A}_{1,m} = \frac{\tilde{\Gamma}_1 + \phi}{1 - k_{1,m}\rho}$$

$$\tilde{A}_{2,m} = \frac{\tilde{\Gamma}_2 + 0.5[(\pi-\gamma)^2 + \varphi_d^2 + [k_{1,m}\tilde{A}_{1,m} + (1-\gamma)\tilde{A}_1]^2\varphi_e^2]}{1 - k_{1,m}\nu}$$

$$\tilde{A}_{0,m} = \frac{\tilde{\Gamma}_0 + k_{0m} + \mu_d + k_{1m}\tilde{A}_{2m}(1-\nu)\bar{\sigma}^2 + 0.5[k_{1m}\tilde{A}_{2m} + (1-\gamma)\tilde{A}_2]^2\sigma_w^2}{1 - k_{1,m}}$$

$\tilde{\Gamma}_1 = -\gamma + (1-\gamma)\tilde{A}_1\frac{(\delta\rho-1)}{\delta}$, $\tilde{\Gamma}_2 = (1-\gamma)\tilde{A}_2\frac{(\delta\nu-1)}{\delta}$,$\tilde{\Gamma}_0 = \log\delta + (1-\gamma)\tilde{A}_0\frac{(\delta-1)}{\delta} - \gamma\mu_c + (1-\gamma)\tilde{A}_2(1-\nu)\bar{\sigma}^2$,

- For the risk free asset,

$$\tilde{A}_{1,f} = \gamma + (1-\gamma)\tilde{A}_1\frac{(1-\delta\rho)}{\delta}$$

$$\tilde{A}_{2,f} = (1-\gamma)\tilde{A}_2\frac{(1-\delta\nu)}{\delta} - 0.5\left(\gamma^2 + [(1-\gamma)\tilde{A}_1\varphi_e]^2\right)]$$

$$\tilde{A}_{0,f} = -\log\delta - (1-\gamma)\tilde{A}_0\frac{(\delta-1)}{\delta} + \gamma\mu_c - (1-\gamma)\tilde{A}_2(1-\nu)\bar{\sigma}^2 - 0.5(\gamma-1)^2\tilde{A}_2^2\sigma_w^2$$

4.8 Fixed point feasibility condition for the log price-dividend ratio

The fixed point condition is determined by the equilibrium value of the log price dividend ratio, which should satisfy: $\bar{z}_m = A_{0m} + A_{2m}\bar{\sigma}^2$

Substituting the affine coefficient A_{0m} and A_{2m} by their following expressions:

$$A_{0m} = \frac{R_1 + k_{0,m} + k_{1,m}A_{2,m}\sigma^2(1-\nu) + 0.5[R_2 + k_{1,m}A_{2,m}]^2\sigma_w^2}{1 - k_{1,m}}$$

Where $R_1 = \theta\log\delta - \gamma\mu_c + (\theta-1)[k_0 + (k_1-1)A_0 + k_1A_2\sigma^2(1-\nu)] + \mu_d$ and $R_2 = (\theta-1)k_1A_2$

$$A_{2m} = \left[\Delta_{123} + \frac{e^{2\bar{z}_m}}{(1+e^{\bar{z}_m}(1-\rho))^2}\left(\phi - \frac{1}{\psi}\right)^2\varphi_e^2 + \frac{2e^{\bar{z}_m}}{(1+e^{\bar{z}_m}(1-\rho))}\Delta_3\left(\phi - \frac{1}{\psi}\right)\varphi_e^2\right]\frac{(1+e^{\bar{z}_m})}{(1+e^{\bar{z}_m}(1-\nu))}$$

Where $\Delta_1 = \frac{-(\theta-1)\frac{1}{2}\left[(-\frac{\theta}{\psi}+\theta)^2 + (\theta k_1A_1\varphi_e)^2\right]}{\theta}$, $\Delta_2 = (\pi-\gamma)^2 + \varphi_d^2$, $\Delta_3 = (\theta-1)k_1A_1$ and $\Delta_{123} = \Delta_1 + 0.5\Delta_2 + \varphi_e^2\Delta_3^2$

We get that the fixed point condition is equivalent to :

$$f(\bar{z}_m) = g(\bar{z}_m)$$

where

$$f(\bar{z}_m) = \bar{z}_m - log(1+e^{\bar{z}_m}) - \left[\Delta_{123} + \frac{e^{2\bar{z}_m}}{(1+e^{\bar{z}_m}(1-\rho))^2}\left(\phi - \frac{1}{\psi}\right)^2\varphi_e^2 + \frac{2e^{\bar{z}_m}}{(1+e^{\bar{z}_m}(1-\rho))}\Delta_3\left(\phi - \frac{1}{\psi}\right)\varphi_e^2\right]\bar{\sigma}^2$$

$$g(\bar{z}_m) = R_1 + 0.5\sigma_w^2 \left[R_0 + \frac{e^{\bar{z}_m}}{(1+e^{\bar{z}_m}(1-\nu))}\left(\Delta_{123} + \frac{e^{2\bar{z}_m}}{(1+e^{\bar{z}_m}(1-\rho))^2}\left(\phi - \frac{1}{\psi}\right)^2 \varphi_e^2 + \frac{2e^{\bar{z}_m}}{(1+e^{\bar{z}_m}(1-\rho))}\Delta_3\left(\phi-\frac{1}{\psi}\right)\varphi_e^2\right)\right]^2$$

$$\lim_{\bar{z}_m \to +\infty} f(\bar{z}_m) = -\left[\Delta_{123} + \frac{\left(\phi - \frac{1}{\psi}\right)^2}{(1-\rho)^2}\varphi_e^2 + \frac{2\Delta_3\left(\phi - \frac{1}{\psi}\right)\varphi_e^2}{(1-\rho)}\right]\bar{\sigma}^2 = B^I$$

$$f(0) = -\left[\Delta_{123} + \frac{\left(\phi - \frac{1}{\psi}\right)^2}{(2-\rho)^2}\varphi_e^2 + \frac{2\Delta_3\left(\phi - \frac{1}{\psi}\right)\varphi_e^2}{(2-\rho)}\right]\bar{\sigma}^2 = B^S$$

$$\lim_{\bar{z}_m \to +\infty} g(\bar{z}_m) = R_1 + 0.5\sigma_w^2\left(R_2 - \frac{B^S}{(1-\nu)\bar{\sigma}^2}\right)^2 = G^S$$

$$g(0) = R_1 + 0.5\sigma_w^2\left(R_2 - \frac{B^S}{(2-\nu)\bar{\sigma}^2}\right)^2 = G^I$$

When ($\gamma > 1$ and $\psi < 1$) or ($\gamma < 1$ and $\psi > 1$) (meaning that $\theta < 0$), f is decreasing from $[0, +\infty[\to]B^I, B^S]$ and g is increasing from $[0, +\infty[\to [G^I, G^S[$

4.9 Temporal aggregation

- **Annual consumption growth**

As shown in Bansal et al. [2007b], the h-period aggregated consumption growth rate can be expressed as a weighted average of monthly consumption growth, with the weight taking a $\Lambda - shape$:

$$\Delta c_{t+1}^a = log\frac{\sum_{j=0}^{h-1} C_{h(t+1)-j}}{\sum_{j=0}^{h-1} C_{ht-j}} \approx \sum_{j=0}^{2h-2} \tau_j \Delta c_{h(t+1)-j}$$

where the index t is used to count the aggregated time and $h(t-1)+1$ to ht are the corresponding month within the aggregate period t

$$\tau_j = \frac{j+1}{h} \quad if \quad j < h \qquad and \qquad \tau_j = \frac{2h-j-1}{h} \quad if \quad j \geqslant h$$

It follows that :

$$\Delta c_{t+1}^a = \mu_c^a + \phi^a x_{h(t+1)-2h-1} + \eta_{t+1}^a$$

with

$$\mu_c^a = \sum_{j=0}^{2h-2} \tau_j \mu_c, \quad , \phi^a = \sum_{j=0}^{2h-2} \tau_j \rho^j$$

and

$$\eta_{t+1}^a = \sum_{j=0}^{2h-2} \tau_j \sigma_{h(t+1)-j-1} \eta_{h(t+1)-j} + \sum_{j=0}^{2h-3} \tau_j \left[\sum_{k=0}^{2(h-1)-j-1} \rho^k \varphi_e \sigma_{h(t+1)-j-k-2} e_{h(t+1)-j-k-1} \right]$$

As,

$$\sum_{j=0}^{n} \tau_j \rho^{n-j} = \begin{cases} \frac{\rho^{(n+2)}-(n+2)\rho+n+1}{h(1-\rho)^2} & n < h \\ \\ \frac{\rho^{(n+2)}-2\rho^{(n-h+2)}+(2(h-1)-n)(1-\rho)+1}{h(1-\rho)^2} & n \geq h \end{cases}$$

Δc_{t+1}^acan also be written the following way which enables a compact writing of the analytical formulas:

$$\Delta c_{t+1}^a = h\mu_c + \rho b_{h-2} x_{h(t-1)} + \sum_{j=0}^{h-1} a_j \varphi_e v_{h(t+1)-j-1} + \sum_{j=0}^{h-2} b_j \varphi_e v_{ht-j-1}$$

$$+ \sum_{j=0}^{2(h-1)} \tau_j \sigma_{h(t+1)-j-1} \eta_{h(t+1)-j}$$

with

$$\begin{cases} a_j = \sum_{k=0}^{j} \left(\frac{k+1}{h}\right) \rho^{j-k} = \frac{\rho^{j+2}-(j+2)\rho+j+1}{h(1-\rho)^2} & if \quad j < h \\ \\ b_j = \sum_{k=0}^{j+h} \tau_k \rho^{j+h-k} = \frac{\rho^{j+h+2}-2\rho^{j+2}+(h-2-j)(1-\rho)+1}{h(1-\rho)^2} & if \quad j \leq h-2 \end{cases}$$

and $v_{h(t+1)-j-1} = \sigma_{h(t+1)-j-2} e_{h(t+1)-j-1}$.

Notice that $\forall h \geq 1$, $b_{h-2} = \frac{\left(1-\rho^h\right)^2}{h(1-\rho)^2}$, but in the summations from 0 to $h-2$, $b_j = 0$ if $j < 0$ (for example when $h = 1$). So out of the summation, $b_{h-2} = \frac{\left(1-\rho^h\right)^2}{h(1-\rho)^2}$ to avoid writing this ugly fraction every time but in the summations from 0 to $h-2$ $b_j = 0$, if $h = 1$

Therefore,

$$E(\Delta c_{t+1}^a) = E(\sum_{j=0}^{2(h-1)} \tau_j \mu_c) = h\mu_c$$

$$var(\Delta c_{t+1}^a) = var(\sum_{j=0}^{2(h-1)} \tau_j \Delta c_{h(t+1)-j})$$

$$= var\left(\sum_{j=0}^{2(h-1)} \tau_j x_{h(t+1)-j-1} + \sum_{j=0}^{2(h-1)} \tau_j \sigma_{h(t+1)-j-1} \eta_{h(t+1)-j}\right)$$

$$= \sum_{j=0}^{2(h-1)} \tau_j^2 var(x_{ht}) + 2 \sum_{j=0}^{2(h-1)-1} \sum_{k=j+1}^{2(h-1)} \tau_j \tau_k cov(x_{h(t+1)-J-1}, x_{h(t+1)-k-1}) + \sum_{j=0}^{2(h-1)} \tau_j^2 \bar{\sigma}^2$$

$$var(\Delta c_{t+1}^a) = \left[\sum_{j=0}^{2(h-1)} \tau_j^2 + 2 \sum_{j=0}^{2h-3} \sum_{k=j+1}^{2(h-1)} \tau_j \tau_k \rho^{k-j} \right] \left(\frac{\varphi_e^2 \bar{\sigma}^2}{1-\rho^2} \right) + \left(\sum_{j=0}^{2h-2} \tau_j^2 \right) \bar{\sigma}^2$$

$$var(\Delta c_{t+1}^a) = (\rho b_{h-2})^2 \left(\frac{\varphi_e^2 \bar{\sigma}^2}{1-\rho^2} \right) + \left(\sum_{j=0}^{h-1} a_j^2 + \sum_{j=0}^{h-2} b_j^2 \right) \varphi_e^2 \bar{\sigma}^2 + \left(\sum_{j=0}^{2(h-1)} \tau_j^2 \right) \bar{\sigma}^2$$

$$cov(\Delta c_t^a, \Delta c_{t+1}^a) = cov(\sum_{j=0}^{2(h-1)} \tau_j \Delta c_{ht-j}, \sum_{j=0}^{2(h-1)} \tau_j \Delta c_{h(t+1)-j})$$

$$= cov \left(\sum_{j=0}^{2(h-1)} \tau_j x_{ht-j-1} + \sum_{j=0}^{2(h-1)} \tau_j \sigma_{ht-j-1} \eta_{ht-j}, \sum_{m=0}^{2(h-1)} \tau_m x_{h(t+1)-m-1} + \sum_{m=0}^{2(h-1)} \tau_m \sigma_{h(t+1)-m-1} \eta_{h(t+1)-m} \right)$$

$$= \left(\sum_{j=0}^{h-2} \left[\sum_{m=0}^{j+h-1} \tau_j \tau_m \rho^{h-m+j} + \sum_{m=j+h}^{2(h-1)} \tau_j \tau_m \rho^{m-j-h} \right] + \sum_{j=h-1}^{2(h-1)} \sum_{m=0}^{2(h-1)} \tau_j \tau_m \rho^{h-m+j} \right) \left(\frac{\varphi_e^2 \bar{\sigma}^2}{1-\rho^2} \right)$$

$$+ \left(\sum_{m=0}^{h-2} \tau_{m+h} \tau_m \right) \bar{\sigma}^2$$

$$cov(\Delta c_t^a, \Delta c_{t+1}^a) = \rho^h (\rho b_{h-2})^2 \left(\frac{\varphi_e^2 \bar{\sigma}^2}{1-\rho^2} \right) + \sum_{j=0}^{h-2} \left[\rho^{j+2} b_j b_{h-2} + a_j b_j \right] \varphi_e^2 \bar{\sigma}^2 + \rho b_{h-2} a_{h-1} \varphi_e^2 \bar{\sigma}^2$$

$$+ \left(\sum_{m=0}^{h-2} \tau_{m+h} \tau_m \right) \bar{\sigma}^2$$

More generally, we can show that for $s \geq 1$,

$$cov(\Delta c_t^a, \Delta c_{t+s}^a) = \rho^{hs} (\rho b_{h-2})^2 \left(\frac{\varphi_e^2 \bar{\sigma}^2}{1-\rho^2} \right) + \sum_{j=0}^{h-2} \left[\rho^{j+h(s-1)+2} b_j b_{h-2} \right] \varphi_e^2 \bar{\sigma}^2$$

$$+\mathbb{1}\{s=1\}\left\{\rho b_{h-2}a_{h-1}+\sum_{j=0}^{h-2}a_jb_j\right\}\varphi_e^2\bar{\sigma}^2+\mathbb{1}\{s=1\}\left(\sum_{m=0}^{h-2}\tau_{m+h}\tau_m\right)\bar{\sigma}^2$$

$$+\mathbb{1}\{s>1\}\left\{\sum_{j=0}^{h-1}\left[\rho^{j+h(s-2)+2}a_jb_{h-2}\right]\varphi_e^2\bar{\sigma}^2\right\}$$

- **Annual dividend growth**

Using the same formulation as for the consumption growth rate, the annual dividend growth rate can be expressed in term of monthly dividend growth rates as follow:

$$\Delta d_{t+1}^a = log\frac{\sum_{j=0}^{h-1}D_{h(t+1)-j}}{\sum_{j=0}^{h-1}D_{ht-j}} \approx \sum_{j=0}^{2(h-1)}\tau_j\Delta d_{h(t+1)-j}$$

$$\Delta d_{t+1}^a = h\mu_d+\phi\rho b_{h-2}x_{h(t-1)}+\phi\varphi_e\sum_{j=0}^{h-1}a_jv_{h(t+1)-j-1}+\phi\varphi_e\sum_{j=0}^{h-2}b_jv_{ht-j-1}$$

$$+\pi\sum_{j=0}^{2(h-1)}\tau_j\sigma_{h(t+1)-j-1}\eta_{h(t+1)-j}+\varphi_d\sum_{j=0}^{2(h-1)}\tau_j\sigma_{h(t+1)-j-1}u_{h(t+1)-j}$$

Then

$$E(\Delta d_t^a)=E(\sum_{j=0}^{2(h-1)}\tau_j\Delta d_{ht-j})=h\mu_d$$

$$var(\Delta d_t^a)=var\left(\phi\sum_{j=0}^{2(h-1)}\tau_jx_{ht-j-1}\right)+var\left(\varphi_d\sum_{j=0}^{2(h-1)}\tau_j\sigma_{ht-j-1}u_{ht-j}\right)+var(\pi\sum_{j=0}^{2(h-1)}\tau_j\sigma_{ht-j-1}\eta_{ht-j})$$

$$=\phi^2\left[\sum_{j=0}^{2(h-1)}\tau_j^2+2\sum_{j=0}^{2h-3}\sum_{k=j+1}^{2(h-1)}\tau_j\tau_k\rho^{k-j}\right]\left(\frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}\right)+\left(\varphi_d^2+\pi^2\right)\sum_{j=0}^{2(h-1)}\tau_j^2\bar{\sigma}^2$$

$$var(\Delta d_t^a)=(\phi\rho b_{h-2})^2\left(\frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}\right)+\left(\sum_{j=0}^{h-1}a_j^2+\sum_{j=0}^{h-2}b_j^2\right)\phi^2\varphi_e^2\bar{\sigma}^2+\left(\pi^2+\varphi_d^2\right)\left(\sum_{j=0}^{2(h-1)}\tau_j^2\right)\bar{\sigma}^2$$

$$cov(\Delta d_t^a,\Delta d_{t+1}^a)=cov\left(\sum_{j=0}^{2(h-1)}\phi\tau_jx_{ht-j-1}+\sum_{j=0}^{2(h-1)}\tau_j\varphi_d\sigma_{ht-j-1}u_{ht-j}+\sum_{j=0}^{2(h-1)}\tau_j\pi\sigma_{ht-j-1}\eta_{ht-j}\right.,$$

$$\sum_{m=0}^{2(h-1)} \phi\tau_m x_{h(t+1)-m-1} + \sum_{m=0}^{2(h-1)} \tau_m\varphi_d\sigma_{h(t+1)-m-1}u_{h(t+1)-m} + \sum_{m=0}^{2(h-1)} \tau_m\pi\sigma_{h(t+1)-m-1}\eta_{h(t+1)-m}\Bigg)$$

$$= \phi^2 \left(\sum_{j=0}^{h-2}\left[\sum_{m=0}^{j+h-1}\tau_j\tau_m\rho^{h-m+j} + \sum_{m=j+h}^{2(h-1)}\tau_j\tau_m\rho^{m-j-h}\right] + \sum_{j=h-1}^{2(h-1)}\sum_{m=0}^{2(h-1)}\tau_j\tau_m\rho^{h-m+j}\right)\left(\frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}\right)$$

$$+ \left(\varphi_d^2+\pi^2\right)\left(\sum_{m=0}^{h-2}\tau_{m+h}\tau_m\right)\bar{\sigma}^2$$

$$cov(\Delta d_t^a, \Delta d_{t+1}^a) = \rho^h\left(\phi\rho b_{h-2}\right)^2\left(\frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}\right) + \sum_{j=0}^{h-2}\left[\rho^{j+2}b_jb_{h-2}+a_jb_j\right]\phi^2\varphi_e^2\bar{\sigma}^2$$

$$+ \left(\rho b_{h-2}a_{h-1}\right)\phi^2\varphi_e^2\bar{\sigma}^2 + \left(\varphi_d^2+\pi^2\right)\left(\sum_{m=0}^{h-2}\tau_{m+h}\tau_m\right)\bar{\sigma}^2$$

$$cov(\Delta c_t^a, \Delta d_t^a) = cov\left(\sum_{j=0}^{2(h-1)}\tau_jx_{ht-j-1} + \sum_{j=0}^{2(h-1)}\tau_j\sigma_{ht-j-1}\eta_{ht-j}, \sum_{j=0}^{2(h-1)}\phi\tau_jx_{ht-j-1}\right.$$

$$\left. + \sum_{j=0}^{2(h-1)}\tau_j\varphi_d\sigma_{ht-j-1}u_{ht-j} + \sum_{j=0}^{2(h-1)}\tau_j\pi\sigma_{ht-j-1}\eta_{ht-j}\right)$$

$$= \phi var(\sum_{j=0}^{2(h-1)}\tau_jx_{ht-j-1}) + \pi var(\sum_{j=0}^{2(h-1)}\tau_j\sigma_{ht-j-1}\eta_{ht-j})$$

$$= \phi\left[\sum_{j=0}^{2(h-1)}\tau_j^2 + 2\sum_{j=0}^{2h-3}\sum_{k=j+1}^{2(h-1)}\tau_j\tau_k\rho^{k-j}\right]\left(\frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}\right) + \pi\left(\sum_{j=0}^{2(h-1)}\tau_j^2\right)\bar{\sigma}^2$$

$$cov(\Delta c_t^a, \Delta d_{t+1}^a) = \phi\left(\rho b_{h-2}\right)^2\left(\frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}\right) + \left[\sum_{j=0}^{h-1}a_j^2 + \sum_{j=0}^{h-2}b_j^2\right]\phi\varphi_e^2\bar{\sigma}^2$$

$$+\pi\left(\sum_{j=0}^{2(h-1)}\tau_j^2\right)\bar{\sigma}^2$$

- **Annual market return**

Let's denote

$$\Gamma_0 = k_{0m} + (k_{1m}-1)A_{0m} + k_{1m}A_{2m}(1-\nu)\bar{\sigma}^2 + \mu_d$$

$$\Gamma_{1m} = A_{1m}(k_{1m}\rho - 1) + \phi \quad \Gamma_{2m} = (k_{1m}\nu - 1)A_{2m}$$

$$\beta_{m,e} = k_{1m}A_{1m}\varphi_e;\ \beta_{m,w} = k_{1m}A_{2m}\sigma_w$$

Then

$$r^a_{m,t+j} = \sum_{k=0}^{h-1}\Big(\Gamma_0 + \Gamma_{1m}x_{h(t+j)-k-1} + \Gamma_{2m}\sigma^2_{h(t+j)-k-1} + \beta_{m,e}v_{h(t+j)-k}$$

$$+\beta_{m,w}w_{h(t+j)-k} + \varphi_d\sigma_{h(t+j)-k-1}u_{h(t+j)-k} + \pi\sigma_{h(t+j)-k-1}\eta_{h(t+j)-k}\Big)$$

$$= h\Gamma_0 + \Gamma_{1m}\left(\sum_{k=0}^{h-1}\rho^{h(j+1)-1-k}x_{h(t-1)} + \sum_{k=0}^{h-1}\sum_{r=0}^{h(j+1)-2-k}\varphi_e\rho^r v_{h(t+j)-k-r-1}\right)$$

$$+\Gamma_{2m}\left(\sum_{k=0}^{h-1}(1-\nu)\left(\sum_{r=0}^{h(j+1)-2-k}\nu^r\right)\bar{\sigma}^2 + \sum_{k=0}^{h-1}\nu^{h(j+1)-1-k}\sigma^2_{h(t-1)} + \sigma_w\sum_{k=0}^{h-1}\sum_{r=0}^{h(j+1)-2-k}\nu^r w_{h(t+j)-k-r-1}\right)$$

$$+\sum_{k=0}^{h-1}\beta_{m,e}v_{h(t+j)-k} + \sum_{k=0}^{h-1}\beta_{m,w}w_{h(t+j)-k}$$

$$+\varphi_d\sum_{k=0}^{h-1}\sigma_{h(t+j)-k-1}u_{h(t+j)-k} + \pi\sum_{k=0}^{h-1}\sigma_{h(t+j)-k-1}\eta_{h(t+j)-k}$$

$$= h\Gamma_0 + \Gamma_{2m}\left(h - \nu^{hj}\left(\frac{1-\nu^h}{1-\nu}\right)\right)\bar{\sigma}^2 + \Gamma_{1m}\rho^{hj}\left(\frac{1-\rho^h}{1-\rho}\right)x_{h(t-1)} + \Gamma_{2m}\nu^{hj}\left(\frac{1-\rho^h}{1-\rho}\right)\sigma^2_{h(t-1)}$$

$$+\sum_{k=0}^{h-1}\beta_{m,e}v_{h(t+j)-k} + \varphi_e\Gamma_{1m}\left(\sum_{r=0}^{h-1}\left(\frac{1-\rho^{r+1}}{1-\rho}\right)v_{h(t+j)-r-1} + \left(\frac{1-\rho^h}{1-\rho}\right)\sum_{r=h}^{h(j+1)-2}\rho^{(r-h+1)}v_{h(t+j)-r-1}\right)$$

$$+\sum_{k=0}^{h-1}\beta_{m,w}w_{h(t+j)-k} + \sigma_w\Gamma_{2m}\left(\sum_{r=0}^{h-1}\left(\frac{1-\nu^h}{1-\nu}\right)w_{h(t+j)-r-1} + \left(\frac{1-\nu^h}{1-\nu}\right)\sum_{r=h}^{h(j+1)-2}\nu^{(r-h+1)}w_{h(t+j)-r-1}\right)$$

$$+\varphi_d\sum_{k=0}^{h-1}\sigma_{h(t+j)-k-1}u_{h(t+j)-k} + \pi\sum_{k=0}^{h-1}\sigma_{h(t+j)-k-1}\eta_{h(t+j)-k}$$

and

$$E(r^a_{m,t}) = h\Gamma_0 + h\Gamma_{2m}\sigma^2$$

$$var(r^a_{m,t}) = var(\sum_{j=0}^{h-1}r_{m,ht-j}) = var(\sum_{j=0}^{h-1}\Gamma_0 + \Gamma_{1m}x_{ht-j-1} + \Gamma_{2m}\sigma^2_{ht-j-1} + \beta_{m,e}\sigma_{ht-j-1}e_{ht-j}$$

$$+\beta_{m,w}w_{ht-j} + \varphi_d\sigma_{ht-j-1}u_{ht-j} + \pi\sigma_{ht-j-1}\eta_{ht-j})$$

$$= \sum_{j=0}^{h-1}(\Gamma^2_{1m}var(x_{ht-j-1}) + \Gamma^2_{2m}var(\sigma^2_{ht-m-1}) + \beta^2_{m,e}\bar{\sigma}^2 + \beta^2_{m,w} + \varphi^2_d\bar{\sigma}^2) + 2\sum_{j=0}^{h-2}\sum_{k=j+1}^{h-1}cov(\Gamma_{1m}x_{ht-j-1}, \Gamma_{1m}x_{ht-k-1})$$

$$+\Gamma^2_{2m}cov(\sigma^2_{ht-j-1}, \sigma^2_{ht-k-1}) + \Gamma_{1m}\beta_{m,e}cov(x_{ht-j-1}, \sigma_{ht-k-1}e_{ht-k}) + \Gamma_{2m}\beta_{m,w}cov(\sigma^2_{ht-m-1}, w_{ht-k})$$

$$= \left[h + \frac{2h\rho}{1-\rho} - \frac{2\rho(1-\rho^h)}{(1-\rho)^2}\right] \frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}\Gamma_{1m}^2 + \left[h + \frac{2h\nu}{1-\nu} - \frac{2\nu(1-\nu^h)}{(1-\nu)^2}\right] \frac{\sigma_w^2}{1-\nu^2}\Gamma_{2m}^2 + h[(\beta_{m,e}^2 + \varphi_e^2 + \pi^2)\bar{\sigma}^2 + \beta_{m,w}^2]$$

$$+2\Gamma_{1m}\beta_{m,e}\varphi_e\bar{\sigma}^2 \left(\frac{h}{1-\rho} - \frac{1-\rho^h}{(1-\rho)^2}\right) + 2\Gamma_{2m}\beta_{m,w}\sigma_w \left(\frac{h}{1-\nu} - \frac{1-\nu^h}{(1-\nu)^2}\right)$$

$$cov(r_{m,t}^a, r_{m,t+1}^a) = \rho\left(\frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}\right)\left(\frac{1-\rho^h}{1-\rho}\right)^2 \Gamma_{1m}^2 + \nu\left(\frac{\sigma_w^2}{1-\nu^2}\right)\left(\frac{1-\nu^h}{1-\nu}\right)^2 \Gamma_{2m}^2 + \beta_{m,e}\Gamma_{1m}\varphi_e\bar{\sigma}^2\left(\frac{1-\rho^h}{1-\rho}\right)^2$$

$$+\beta_{m,w}\Gamma_{2m}\sigma_w\left(\frac{1-\nu^h}{1-\nu}\right)^2$$

- **Annual risk free rate**

$$r_{f,t+j}^a = \sum_{k=0}^{h-1} A_{0,f} + A_{1,f}x_{12(t+j)-k} + A_{2,,f}\sigma_{h(t+j)-k}^2$$

$$= hA_{0f} + \sum_{k=0}^{h-1}\left[A_{1f}\left(\rho^{h(j+1)-k}x_{h(t-1)} + \sum_{r=0}^{h(j+1)-k-1}\rho^r\varphi_e v_{h(t+j)-k-r}\right)\right.$$

$$\left.+A_{2f}\left(\nu^{h(j+1)-k}\sigma_{h(t+1)}^2 + \sum_{r=0}^{12j+11-k}\nu^r\sigma_w w_{h(t+j)-k-r} + (1-\nu^{h(j+1)-k})\bar{\sigma}^2\right)\right]$$

$$= hA_{0f} + A_{1f}\rho^{hj+1}\left(\frac{1-\rho^h}{1-\rho}\right)x_{h(t-1)} + \sum_{k=0}^{h-1}A_{1f}\varphi_e\left(\sum_{r=0}^{h(j+1)-k-1}\rho^r v_{h(t+j)-k-r}\right)$$

$$+A_{2f}\nu^{hj+1}\left(\frac{1-\nu^h}{1-\nu}\right)\sigma_{h(t-1)}^2 + \sum_{k=0}^{h-1}(1-\nu^{h(j+1)-k})A_{2,f}\bar{\sigma}^2 + \sum_{k=0}^{h-1}A_{2,f}\sigma_w\left(\sum_{r=0}^{12j+11-k}\nu^r w_{h(t+j)-k-r}\right)$$

$$= hA_{0f} + A_{1f}\rho^{hj+1}\left(\frac{1-\rho^h}{1-\rho}\right)x_{h(t-1)} + A_{2f}\nu^{hj+1}\left(\frac{1-\nu^h}{1-\nu}\right)\sigma_{h(t-1)}^2 + A_{2,f}\left(h - \nu^{hj+1}\left(\frac{1-\nu^h}{1-\nu}\right)\right)\bar{\sigma}^2$$

$$+A_{1f}\varphi_e\left[\sum_{r=0}^{h-1}\left(\frac{1-\rho^{r+1}}{1-\rho}\right)v_{h(t+j)-r} + \sum_{r=12}^{h(j+1)-1}\rho^{r-h+1}\left(\frac{1-\rho^h}{1-\rho}\right)v_{h(t+j)-r}\right]$$

$$+A_{2f}\sigma_w\left[\sum_{r=0}^{h-1}\left(\frac{1-\nu^{r+1}}{1-\nu}\right)w_{h(t+j)-r} + \sum_{r=12}^{h(j+1)-1}\nu^{r-h+1}\left(\frac{1-\nu^h}{1-\nu}\right)w_{h(t+j)-r}\right]$$

In particular,

$$r_{f,t}^a = \sum_{j=0}^{h-1} r_{f,ht-j} = \sum_{j=0}^{h-1} A_{0f} + A_{1f}x_{ht-j} + A_{2f}\sigma_{ht-j}^2$$

$$= hA_{0f} + A_{1f}\rho\left(\frac{1-\rho^h}{1-\rho}\right)x_{h(t-1)} + A_{2f}\nu\left(\frac{1-\nu^h}{1-\nu}\right)\sigma_{h(t-1)}^2 + \left[h - \nu\left(\frac{1-\nu^h}{1-\nu}\right)\right]A_{2f}\bar{\sigma}^2$$

$$+A_{1f}\varphi_e\sum_{j=0}^{h-1}\left(\frac{1-\rho^{(j+1)}}{1-\rho}\right)v_{ht-j} + A_{2f}\sigma_w\sum_{j=0}^{h-1}\left(\frac{1-\nu^{(j+1)}}{1-\nu}\right)w_{ht-j}$$

$$E(r^a_{f,t}) = hA_{0f} + hA_{2f}\bar{\sigma}^2$$

$$var(r^a_{f,t}) = \left[h + \frac{2h\rho}{1-\rho} - 2\rho\frac{\left(1-\rho^h\right)}{\left(1-\rho\right)^2}\right]\frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}A_{1f}^2 + \left[h + \frac{2h\nu}{1-\nu} - 2\nu\frac{\left(1-\nu^h\right)}{\left(1-\nu\right)^2}\right]\frac{\sigma_w^2}{1-\nu^2}A_{2f}^2$$

$$cov(r^a_{f,t}, r^a_{f,t+1}) = \left(\frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}\right)\rho\left(\frac{1-\rho^h}{1-\rho}\right)^2 A_{1f}^2 + \left(\frac{\sigma_w^2}{1-\nu^2}\right)\nu\left(\frac{1-\nu^h}{1-\nu}\right)^2 A_{2f}^2$$

4.9.1 Expression of asset prices and cash flows in terms of the latent processes

- **Annual price-dividend ratio**

Form (3) and (4), we get that

$$p^a_t - d^a_t = \log P_{ht} - \log\sum_{j=0}^{h-1} D_{ht-j}$$

$$= \log P_{ht} - \log D_{ht} + \sum_{j=0}^{h-1}\left(\log D_{ht-j} - \log D_{ht-j-1}\right) - \log h - \sum_{j=0}^{h-1}\frac{j+1}{h}\triangle d_{ht-j}$$

$$= z_{m,ht} + \sum_{j=0}^{h-1}\triangle d_{ht-j} - \log h - \sum_{j=0}^{h-1}\frac{j+1}{h}\triangle d_{ht-j}$$

$$p^a_t - d^a_t = z_{m,ht} - \log h - \sum_{j=0}^{h-1}\left(\frac{j+1}{h} - 1\right)\triangle d_{ht-j}$$

$$= A_{0m} + A_{1m}x_{ht} + A_{2m}\sigma^2_{ht} - \log(h) - \sum_{j=0}^{h-1}\left(\frac{j+1}{h} - 1\right)\left(\mu_d + \phi x_{ht-j-1} + \pi\sigma_{ht-j-1}\eta_{ht-j} + \varphi_d\sigma_{ht-j-1}u_{ht-j}\right)$$

$$= [A_{0m} - \log(h) - \sum_{j=0}^{h-1}\left(\frac{j+1}{h} - 1\right)\mu_d] + A_{1m}x_{ht} - \sum_{j=0}^{h-1}\left(\frac{j+1}{h} - 1\right)\phi x_{ht-j-1}$$

$$- \sum_{j=0}^{h-1}\left(\frac{j+1}{h} - 1\right)\pi\sigma_{ht-j-1}\eta_{ht-j} + A_{2m}\sigma^2_{ht} - \sum_{j=0}^{h-1}\left(\frac{j+1}{h} - 1\right)\varphi_d\sigma_{ht-j-1}u_{ht-j}$$

$$= \left[A_{0m} + 0.5(h-1)\mu_d - \log(h) + A_{2m}\left(1-\nu^h\right)\bar{\sigma}^2\right] + \sum_{j=1}^{h-1}\left[A_{1m}\rho^j - \phi a'_{j-1}\right]\varphi_e v_{ht-j}$$

$$+ \left[A_{1m}\rho^h - \phi a'_{h-1}\right]x_{h(t-1)} + A_{2m}\nu^h\sigma^2_{h(t-1)} - \varphi_d\sum_{j=0}^{h-1}\left(\frac{j+1}{h} - 1\right)\sigma_{ht-j-1}u_{ht-j}$$

$$+ A_{1m}\varphi_e v_{ht} - \pi\sum_{j=0}^{h-1}\left(\frac{j+1}{h} - 1\right)\sigma_{ht-j-1}\eta_{ht-j} + A_{2m}\sigma_w\sum_{j=0}^{h-1}\nu^j w_{ht-j}$$

where

$$a'_j = a_j - \left(\frac{1-\rho^{(j+1)}}{1-\rho}\right) \qquad for \quad j \in \{0, ..., h-1\}$$

$$E(p^a_t - d^a_t) = [A_{0m} - \log(h) + 0.5(h-1)\mu_d] + A_{2m}\bar{\sigma}^2$$

$$var(p_t^a - d_t^a) = \left[A_{1m}\rho^h - \phi a'_{h-1}\right]^2 \left(\frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}\right) + \sum_{j=1}^{h-1}\left[A_{1m}\rho^j - \phi a'_{j-1}\right]^2 \varphi_e^2\bar{\sigma}^2$$

$$+(A_{1m}\varphi_e)^2\bar{\sigma}^2 + \sum_{j=0}^{h-1}\left(\frac{j+1}{h} - 1\right)^2 \left(\pi^2 + \varphi_d^2\right)\bar{\sigma}^2 + \frac{A_{2m}^2\sigma_w^2}{1-\nu^2}$$

$$cov(p_t^a - d_t^a, p_{t+1}^a - d_{t+1}^a) =$$

$$cov\Bigg(A_{2m}\sigma_{ht}^2 + A_{1m}\varphi_e v_{12t} + \left(A_{1m}\rho^h - \phi a'_{h-1}\right)x_{h(t-1)} + \sum_{j=1}^{h-1}\left(A_{1m}\rho^j - \phi a'_{j-1}\right)\varphi_e v_{ht-j}$$

$$-\sum_{j=0}^{h-1}\left(\frac{j+1}{h} - 1\right)\pi\sigma_{ht-j-1}\eta_{ht-j} - \sum_{j=0}^{h-1}\left(\frac{j+1}{h} - 1\right)\varphi_d\sigma_{ht-j-1}u_{ht-j}, A_{2m}\sigma_{h(t+1)}^2 + A_{1m}\varphi_e v_{h(t+1)}$$

$$+\left(A_{1m}\rho^h - \phi a'_{h-1}\right)x_{ht} + \sum_{j=1}^{h-1}\left(A_{1m}\rho^j - \phi a'_{j-1}\right)\varphi_e v_{h(t+1)-j} - \sum_{j=0}^{h-1}\left(\frac{j+1}{h} - 1\right)\pi\sigma_{h(t+1)-j-1}\eta_{h(t+1)-j}$$

$$-\sum_{j=0}^{h-1}\left(\frac{j+1}{h} - 1\right)\varphi_d\sigma_{h(t+1)-j-1}u_{h(t+1)-j}\Bigg)$$

$$cov(p_t^a - d_t^a, p_{t+1}^a - d_{t+1}^a) = \left[A_{1m}\rho^h - \phi a'_{h-1}\right]^2 \rho^h \left(\frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}\right) + A_{1m}\left[A_{1m}\rho^h - \phi a'_{h-1}\right]\varphi_e^2\bar{\sigma}^2$$

$$+\sum_{j=1}^{h-1}\left[A_{1m}\rho^j - \phi a'_{j-1}\right]\left[A_{1m}\rho^h - \phi a'_{h-1}\right]\rho^j\varphi_e^2\bar{\sigma}^2 + A_{2m}^2\nu^h\left(\frac{\sigma_w^2}{1-\nu^2}\right)$$

- **Forward annual excess return**

Combining the expressions obtained for the annual market return and the annual risk free rate, we obtain that for $j \geq 1$,

$$r_{m,t+j}^a - r_{f,t+j}^a = h(\Gamma_0 - A_{0,f}) + h(\Gamma_{2m} - A_{2,f})\bar{\sigma}^2 - \nu^{hj}\left(\frac{1-\nu^h}{1-\nu}\right)(\Gamma_{2m} - \nu A_{2,f})\bar{\sigma}^2$$

$$+\rho^{hj}\left(\frac{1-\rho^h}{1-\rho}\right)(\Gamma_{1m} - A_{1,f}\rho)x_{h(t-1)} + \nu^{hj}\left(\frac{1-\nu^h}{1-\nu}\right)(\Gamma_{2m} - A_{2,f}\nu)\sigma_{h(t-1)}^2$$

$$+\sum_{k=0}^{h-1}\left[\beta_{m,e} + \Gamma_{1m}\varphi_e\left(\frac{1-\rho^k}{1-\rho}\right) - A_{1,f}\varphi_e\left(\frac{1-\rho^{k+1}}{1-\rho}\right)\right]v_{h(t+j)-k}$$

$$+\sum_{k=0}^{h-1}\left[\beta_{m,w} + \Gamma_{2m}\sigma_w\left(\frac{1-\nu^k}{1-\nu}\right) - A_{2,f}\sigma_w\left(\frac{1-\nu^{k+1}}{1-\nu}\right)\right]w_{h(t+j)-k}$$

$$+\sum_{k=h}^{hj+h-1}\left[\Gamma_{1m} - \rho A_{1,f}\right]\varphi_e\left(\frac{1-\rho^h}{1-\rho}\right)\rho^{k-h}v_{h(t+j)-k}$$

$$+\sum_{k=h}^{hj+h-1}\left[\Gamma_{2m} - A_{2,f}\nu\right]\sigma_w\nu^{k-h}\left(\frac{1-\nu^h}{1-\nu}\right)w_{h(t+j)-k}$$

$$+\pi\sum_{k=0}^{h-1}\sigma_{h(t+j)-k-1}\eta_{h(t+1)-k}+\varphi_d\sum_{k=0}^{h-1}\sigma_{h(t+j)-k-1}u_{h(t+j)-k}$$

So,

$$\sum_{j=1}^{J}\left(r^a_{m,t+j}-r^a_{f,t+j}\right)=\mu_r^{a,J}+\phi_x^{a,J}x_{h(t-1)}+\phi_\sigma^{a,J}\sigma^2_{h(t-1)}+e_t^{a,J}+w_t^{a,J}+\varphi_d u_t^{a,J}+\pi\eta_t^{a,J}$$

where $\mu_r^{a,J}=hJ(\Gamma_0-A_{0,f})+hJ(\Gamma_{2m}-A_{2,f})\bar{\sigma}^2-\nu^h\left(\frac{1-\nu^{hJ}}{1-\nu}\right)(\Gamma_{2m}-\nu A_{2,f})\bar{\sigma}^2$, $\phi_x^{a,J}=\rho^h\left(\frac{1-\rho^{hJ}}{1-\rho}\right)(\Gamma_{1m}-A_{1,f}\rho)$, $\phi_\sigma^{a,J}=\nu^h\left(\frac{1-\nu^{hJ}}{1-\nu}\right)(\Gamma_{2m}-A_{2,f}\nu)$,

$$e_t^{a,J}=\sum_{j=1}^{J}\left(\sum_{k=0}^{h-1}\left[\beta_{m,e}+\Gamma_{1m}\varphi_e\left(\frac{1-\rho^k}{1-\rho}\right)-A_{1,f}\varphi_e\left(\frac{1-\rho^{k+1}}{1-\rho}\right)\right]v_{h(t+j)-k}\right.$$

$$\left.+\sum_{k=h}^{hj+h-1}\left[\Gamma_{1m}-\rho A_{1,f}\right]\varphi_e\left(\frac{1-\rho^h}{1-\rho}\right)\rho^{k-h}v_{h(t+j)-k}\right)$$

remind that $v_{h(t+1)-j-1}=\sigma_{h(t+1)-j-2}e_{h(t+1)-j-1}$.

$$w_t^{a,J}=\sum_{j=1}^{J}\left(\sum_{k=0}^{h-1}\left[\beta_{m,w}+\Gamma_{2m}\sigma_w\left(\frac{1-\nu^k}{1-\nu}\right)-A_{2,f}\sigma_w\left(\frac{1-\nu^{k+1}}{1-\nu}\right)\right]w_{h(t+j)-k}\right.$$

$$\left.+\sum_{k=h}^{hj+h-1}\left[\Gamma_{2m}-A_{2,f}\nu\right]\sigma_w\nu^{k-h}\left(\frac{1-\nu^h}{1-\nu}\right)w_{h(t+j)-k}\right)$$

$u_t^{a,J}=\sum_{j=1}^{J}\sum_{k=0}^{h-1}\sigma_{h(t+j)-k-1}u_{h(t+j)-k}$ and $\eta_t^{a,J}=\sum_{j=1}^{J}\sum_{k=0}^{h-1}\sigma_{h(t+j)-k-1}\eta_{h(t+1)-k}$

4.9.2 Theoretical moments for the predictive regression

- **Prediction of future excess return by the log price-dividend ratio**

$$cov(r^a_{m,t+j}-r^a_{f,t+j},p^a_t-d^a_t)=\left[A_{1m}\rho^h-\phi a'_{h-1}\right]\rho^{hj}\left(\frac{1-\rho^h}{1-\rho}\right)(\Gamma_{1m}-\rho A_{1,f})\left(\frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}\right)$$

$$+A_{2m}\nu^{h(j+1)}\left(\frac{1-\nu^h}{1-\nu}\right)(\Gamma_{2m}-A_{2,f}\nu)\left(\frac{\sigma_w^2}{1-\nu^2}\right)+A_{1m}\rho^{12(j-1)}\left(\frac{1-\rho^h}{1-\rho}\right)(\Gamma_{1m}-\rho A_{1,f})\varphi_e^2\bar{\sigma}^2$$

$$+\sum_{k=1}^{h-1}\rho^{h(j-1)+k}\left[A_{1m}\rho^k-\phi a'_{k-1}\right]\left(\frac{1-\rho^h}{1-\rho}\right)(\Gamma_{1m}-\rho A_{1,f})\varphi_e^2\bar{\sigma}^2$$

$$+\sum_{k=0}^{h-1}A_{2m}\nu^{h(j-1)+2k}\left(\frac{1-\nu^h}{1-\nu}\right)(\Gamma_{2m}-\nu A_{2,f})\sigma_w^2$$

$$cov\left(\sum_{j=1}^{J}(r^a_{m,t+j}-r^a_{f,t+j}),p^a_t-d^a_t\right)=\sum_{j=1}^{J}cov(r^a_{m,t+j}-r^a_{f,t+j},p^a_t-d^a_t)$$

$$=A_{1m}\left(\frac{1-\rho^{hJ}}{1-\rho}\right)(\Gamma_{1m}-\rho A_{1,f})\left(\frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}\right)+A_{2m}\left(\frac{1-\nu^{hJ}}{1-\nu}\right)(\Gamma_{2m}-\nu A_{2,f})\left(\frac{\sigma_w^2}{1-\nu^2}\right)$$

$$-\phi a_{h-1}^{'}\rho^{h}\left(\frac{1-\rho^{hJ}}{1-\rho}\right)\left(\Gamma_{1m}-\rho A_{1,f}\right)\left(\frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}\right)-\sum_{k=1}^{h-1}\phi a_{k-1}^{'}\rho^{k}\left(\frac{1-\rho^{hJ}}{1-\rho}\right)\left(\Gamma_{1m}-\rho A_{1,f}\right)\varphi_e^2\bar{\sigma}^2$$

- **Prediction of future consumption growth by the log price-dividend ratio**

$$cov\left(\sum_{j=1}^{J}\Delta c_{t+j}^{a},p_t^a-d_t^a\right)=\left(\rho b_{h-2}A_{1m}\right)\left[\frac{1-\rho^{h(J-1)}}{1-\rho^{h}}\right]\left(\varphi_e^2\bar{\sigma}^2\right)\mathbb{1}\left\{J\geq 2\right\}$$

$$+a_{h-1}A_{1m}\left(\varphi_e^2\bar{\sigma}^2\right)+\left(\rho b_{h-2}\right)\left[\frac{1-\rho^{h(J-1)}}{1-\rho^{h}}\right]\left(\sum_{k=1}^{h-1}\rho^{k}\left(A_{1m}\rho^{k}-\phi a_{k-1}^{'}\right)\right)\left(\varphi_e^2\bar{\sigma}^2\right)\mathbb{1}\left\{J\geq 2\right\}$$

$$+\left(\rho b_{h-2}\right)\left(A_{1m}\rho^{h}-\phi a_{h-1}^{'}\right)\left[\frac{1-\rho^{hJ}}{1-\rho^{h}}\right]\left(\frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}\right)+\sum_{k=1}^{h-1}b_{k-1}\left(A_{1m}\rho^{k}-\phi a_{k-1}^{'}\right)\left(\varphi_e^2\bar{\sigma}^2\right)$$

$$-\pi\sum_{k=0}^{h-2}\left(\frac{k+1}{h}-1\right)\tau_{k+h}\bar{\sigma}^2$$

- **Prediction of future dividend growth by the log price-dividend ratio**

$$cov\left(\sum_{j=1}^{J}\Delta d_{t+j}^{a},p_t^a-d_t^a\right)=\phi\left(\rho b_{h-2}A_{1m}\right)\left[\frac{1-\rho^{h(J-1)}}{1-\rho^{h}}\right]\left(\varphi_e^2\bar{\sigma}^2\right)\mathbb{1}\left\{J\geq 2\right\}$$

$$+\phi a_{h-1}A_{1m}\left(\varphi_e^2\bar{\sigma}^2\right)+\phi\left(\rho b_{h-2}\right)\left[\frac{1-\rho^{h(J-1)}}{1-\rho^{h}}\right]\left(\sum_{k=1}^{h-1}\rho^{k}\left(A_{1m}\rho^{k}-\phi a_{k-1}^{'}\right)\right)\left(\varphi_e^2\bar{\sigma}^2\right)\mathbb{1}\left\{J\geq 2\right\}$$

$$+\phi\left(\rho b_{h-2}\right)\left(A_{1m}\rho^{h}-\phi a_{h-1}^{'}\right)\left[\frac{1-\rho^{hJ}}{1-\rho^{h}}\right]\left(\frac{\varphi_e^2\bar{\sigma}^2}{1-\rho^2}\right)+\phi\sum_{k=1}^{h-1}b_{k-1}\left(A_{1m}\rho^{k}-\phi a_{k-1}^{'}\right)\left(\varphi_e^2\bar{\sigma}^2\right)$$

$$-\left(\pi^2+\varphi_d^2\right)\sum_{k=0}^{h-2}\left(\frac{k+1}{h}-1\right)\tau_{k+h}\bar{\sigma}^2$$

4.10 Short note on Gmm estimation and inference

4.10.1 Strong identification case

The Gmm estimation procedure has been very popular since the seminal paper by Hansen [1982]. Let's call $\theta \in R^p$ the vector of parameters of the model and θ_0 its true value. The set of unconditional moment restrictions can be summarized by :

$$E(g(y_t,\theta_0))=0 \tag{4.10.1}$$

θ is identified at θ_0 if $\theta = \theta_0$ is the unique solution of (4.10.1) for $\theta \epsilon \Theta$ (Stock and Wright, 2000). Let's denote $g_T(y_t,\theta) = \frac{1}{T}\sum_{t=1}^{T} g(y_t,\theta)$ and $\bar{g}(\theta) = E(g(y_t,\theta)$

The Gmm estimator of θ minimizes the criterion function :

$$S_T(\theta;\bar{\theta}_T(\theta)) = T.\left(g_T(y_t,\theta)\right)' W_T(\bar{\theta}_T(\theta))\left(g_T(y_t,\theta)\right)$$

Where $W_T(\bar{\theta}_T(\theta))$ is a positive semi-definite weighting matrix and $\bar{\theta}_T(\theta)$ is the value of the parameters used to compute W_T. For the 2-step Gmm estimator, $W_T = I$ in the first step and for the second step,$\bar{\theta}_T(\theta)$ will be equal to the value estimated in the first step. For the CUE Gmm estimator, $\bar{\theta}_T(\theta) = \theta$. The optimal weighting matrix is the inverse of the asymptotic variance-covariance matrix of $\sqrt{T}g_T(y_t,\theta)$[19] which can be consistently estimated using a kernel based estimator (Newey and West [1987]; Andrews [1991]) by:

$$\widehat{\Omega}(\theta) = \widehat{\Omega}_0 + \sum_{j=1}^{m} \bar{k}_j\left(\widehat{\Omega}_j + \widehat{\Omega}_j^{'}\right)$$

where $\widehat{\Omega}_j = T^{-1}\sum_{t=j+1}^{T}\left(g(y_t,\theta) - g_T(y_t,\theta)\right)\left(g(y_{t-j},\theta) - g_T(y_t,\theta)\right)'$ and $\bar{k}_j$is a kernel function (e.g: Bartlett, Quadratic Spectral, Parzen, Truncated, etc.). For the Bartlett kernel with bandwidth m,$\bar{k}_j = \begin{cases} (1 - \frac{j}{m+1})\ if\ 0 \le j \le m \\ 0\ \ if\ \ Not \end{cases}$ where $g(y_t,\theta) \in R^q \times R^T$ [20]. A part from being consistent, this estimator is also positive semi-definite. Newey and West [1994] also showed how to automatically select the number of lag ($m = parameter \times T^{1/3}$)[21]

From the Central Limit Theorem, we know that:

$$\sqrt{T}\bar{g}_T(y_t,\theta_0) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} g(y_t,\theta_0) \to N(0, Avar\left(\sqrt{T}\bar{g}_T(y_t,\theta_0)\right)) \qquad (4.10.2)$$

So

$$TJ_T(\theta_0) = T.\bar{g}(y_t,\theta_0)' W_T(\theta_0)\bar{g}(y_t,\theta_0) \to \chi(q) \qquad (4.10.3)$$

Where $W_T(\theta_0) = \widehat{\Omega}(\theta_0)^{-1}$, $\widehat{\Omega}(\theta_0)$ is a consistent estimate of the asymptotic variance-covariance matrix $\Omega(\theta_0)$.

[19] $\Omega(\theta) = \underset{T\to+\infty}{lim} T.E\left(\left([g_T(y_t,\theta) - \bar{g}(\theta)]\right)\left([g_T(y_t,\theta) - \bar{g}(\theta)]\right)'\right) = \sum_{j=-\infty}^{\infty} E\left(g(y_t,\theta)g(y_{t-j},\theta)'\right)$

[20] p is the number of parameters in the model, q is the number of non redundant moment restrictions; it is equal here to 16 when we consider only the basic moments and to 19 if we add the moment restrictions from the predictive regressions of the future excess return on the log price dividend ratio and 22 (resp. 25) if we add the moment restrictions from the predictive regressions of the future consumption (resp. future dividend) growth on the log price dividend ratio . T is the length of the data series)

[21] The choice of m is made by minimizing the asymptotic mean squared error $(\underset{T\to\infty}{lim} E\{normalised\, w'(\hat{S} - S)w\}^2)$. For the Bartlett kernel that we used here, $m(T) = \hat{\gamma}T^{1/3}$ with $\hat{\gamma} = c_\gamma\{\hat{s}^{(1)}/\hat{s}^{(0)}\}^{2/3}$,$c_\gamma \equiv \left(1/\int_{-1}^{1}(1-\mid x \mid)^2 dx\right)^{1/3} = \sqrt[3]{3/2} \backsimeq 1.1447$,$\hat{s}^{(1)} = 2\sum_{j=1}^{n} j\hat{\sigma}_j$, $\hat{s}^{(0)} = \hat{\sigma}_0 + 2\sum_{j=1}^{n}\hat{\sigma}_j$,$\hat{\sigma}_j = (T-1)^{-1}\sum_{t=j+2}^{T}\{(w'g(y_t,\theta_0))(w'g(y_{t-j},\theta_0))\}$, $j = 0,...,n$ with $n = [4(T/100)^{2/9}]$ and $w = (1,...,1)'$

If k parameters θ_k^u in the vector θ_0 are estimated while the others are known then under regularity conditions (see Newey and McFadden [1986],section 9.5 for proof):

$$TJ_T(\widehat{\theta_k^u},\theta^0_{-k}) = T.\bar{g}(y_t,\widehat{\theta_k^u},\theta_{0,-k})\widehat{W_T}\bar{g}(y_t,\widehat{\theta_k^u},\theta_{0,-k}) \to \chi\,(q-k) \tag{4.10.4}$$

where θ_k^u is the sub-vector of strongly identified parameters, $\theta_{0,-k}$ is the vector of weakly identified parameters and the weighting matrix $\widehat{W_T}$ is evaluated at the $(\widehat{\theta_k^u},\theta^0_{-k})$. This statistic is the same that we also use in the case of weak identification in the next section; the only difference being that in the weak identification case, because of the inconsistency of the weakly identified parameters, the weighting matrix needs to be evaluated at the same value of the parameters as the moment conditions (in other words we must use the continuously update estimator).

It follows that if we choose to fix $p-k$ parameters in the p-vector at their true values and to estimate k parameters, then :

$$TJ_T(\widehat{\theta_k^u},\theta^0_{-k}) - TJ_T(\widehat{\theta_p^u}) \to \chi\,(p-k) \tag{4.10.5}$$

The three statistics $TJ_T(\widehat{\theta_p^u})$, $TJ_T(\theta_0)$,$TJ_T(\widehat{\theta_k^u},\theta^0_{-k}) - TJ_T(\widehat{\theta_p^u})$ are used in Hansen et al. [1996] to study the finite sample properties of alternative GMM estimators and respectively denominated "Minimized", "True" and "Constrained-Minimized". This last statistic correspond to the quasi-likelihood (LR) statistic which is usually used to test $\theta = \theta_0$ and could be used to construct confidence region for the set of constrained parameters (θ^0_{-k}) based on the increment of the objective function from its unconstrained value. The second statistic corresponds to the Gmm criterion function computed using a consistent estimator of the weighting matrix evaluated at the parameter and it also correspond to the S-statistic of Stock and Wright [2000].

From a Taylor expansion of equation 4.10.2, the confidence intervals for the estimates are obtained using the fact that for the efficient Gmm (as our weighting matrix is a consistent estimate of the asymptotic variance-covariance matrix $\Omega(\theta_0)$), we have (see Hayashi (2000), Prop. 7.10):

$$\sqrt{T}(\hat{\theta}-\theta_0) \to N(0,\left[G'\Omega(\theta_0)^{-1}G\right]^{-1}) \tag{4.10.6}$$

Where $G = E\left[\frac{\partial g(y_t,\theta_0)}{\partial \theta'}\right]$

4.10.2 Weak identification case

We further assumed that $\theta = (\alpha, \beta)$, where α is a p_α sub-vector of weakly identified parameters and β is the p_β sub-vector of strongly identified parameters ($p_\alpha + p_\beta = p$). This means that the moment conditions are zero at $\theta_0 = (\alpha_0, \beta_0)$, but are also very nearly zero for $\alpha \neq \alpha_0$: In other words, the population objective function is steep in β around β_0, but nearly flat in α (Stock and Wright, 2000, Sect 2.3). In a structural economic model, α could be also seen as a vector of "less influential" or "nuisance" parameters, because their changes have no or only a slight influence on the equilibrium characterized by (7). Detecting the weakly identified parameters is not an easy task in a model with non linear equations. If $\bar{g}(y_t, \theta)$ is continuously differentiable and $\nabla_\theta E(\bar{g}(y_t, \theta)) = E(\nabla_\theta \bar{g}(y_t, \theta))$, a necessary and sufficient condition for local identification (which is also necessary for global identification), is that $E(\nabla_\theta \bar{g}(y_t, \theta))$ has a full column rank (Newey and McFadden, 1986), which is no more the case under weak identification.

When some parameters are weakly identified, the limiting distribution of the Gmm estimator is no more a standard normal distribution (Stock and Wright [2000]; Section 2.4). Indeed, the weakly identified parameter α is no more consistently estimated; it converges to the solution α^* of a non-quadratic rather than a local quadratic minimization problem, and thus this solution might be different to the true value of the parameter α_0 which put the objective function (when $\beta = \beta_0$) to 0 (the minimum could be negative and achieved at a parameter different from the true one). This inconsistency in the estimation of the weakly identified parameter affects the limiting distribution of the strongly identified parameters. In the end, *because the limiting distributions are nonstandard, confidence intervals for β constructed by inverting the quasi-likelihood ratio (LR) statistic $S_T(\theta_0; \bar{\theta}_T(\theta_0)) - S_T(\hat{\theta}; \bar{\theta}_T(\hat{\theta}))$ or the conventional Wald statistic will not be valid* (Stock and Wright, 2000).

However, if we assume that $\Psi_T(\theta_0) = \sqrt{T}\left([g_T(y_t, \theta_0) - \bar{g}(\theta_0)]\right) \xrightarrow{d} N(0, \Omega(\theta_0, \theta_0))$ and $W_T(\theta_0) \xrightarrow{p} W(\theta_0) = \Omega(\theta_0, \theta_0)^{-1}$ then $S_T(\theta_0; \theta_0) \xrightarrow{d} \chi^2_m$ (Stock and Wright [2000]; Theorem 2) and this could be used to construct the $(1-\lambda)\%$ confidence region of θ_0 by collecting all the parameters θ such that $S_T(\theta; \theta) \leq \chi^2_{(1-\lambda)}(m)$. Under stronger assumptions (Assumptions B,C,D in Stock and Wright, 2000), $S_T(\alpha_0, \hat{\beta}(\alpha_0); \alpha_0, \hat{\beta}(\alpha_0)) \xrightarrow{d} \chi^2_{m-p_2}$ and this could be used to construct the $(1-\lambda)\%$ confidence region of α_0 by collecting all the parameters α such that $S_T(\alpha, \hat{\beta}(\alpha); \alpha, \hat{\beta}(\alpha)) \leq \chi^2_{(1-\lambda)}(m-p_2)$. Notice that for all those statistic the criterion function used is for the CUE Gmm and the weighting matrix must be a consistent estimator of the inverse of the variance covariance matrix of $\sqrt{T}\bar{g}(y_t, \theta)$. Stock and Wright [2000] recommend to be cautious in the interpretation of the confidence set derived by inverting the S-statistic because the confidence region is the result of a joint test of two hypothesis (that $\theta = \theta_0$ and that the over-identifying conditions are valid). *So the S-set could be small either because the model is correctly specified and precisely estimated or because the model is misspecified but the evidence is two weak to reject it entirely.*

Kleibergen [2005] proposed a decomposition of this joint test into its two part using two independent statistics. The S-statistic can be decomposed as:

$$S_T(\theta_0, \theta_0) = J(\theta_0, \theta_0) + K(\theta_0, \theta_0)$$

Where

$$J(\theta_0, \theta_0) = T.\bar{g}(y_t, \theta)' W_T(\theta_0)^{-1/2} M_{W_T(\theta_0)^{-1/2} D_T(\theta_0, Y)} W_T(\theta_0)^{-1/2} \bar{g}(y_t, \theta)$$

and

$$K(\theta_0, \theta_0) = \frac{1}{4}\left(\frac{\partial S_T(\theta, \theta)}{\partial \theta'}|_{\theta_0}\right)\left[D_T(\theta_0, Y)' W_T(\theta_0)^{-1} D_T(\theta_0, Y)\right]^{-1}\left(\frac{\partial S_T(\theta, \theta)}{\partial \theta'}|_{\theta_0}\right)'$$

$$D_T(\theta_0, Y) = T.[q_{1,T}(y_t, \theta_0) - \hat{V}_{\theta g,1}(\theta_0) W_T(\theta_0)^{-1} \bar{g}_T(y_t, \theta_0) ... q_{p,T}(y_t, \theta_0) - \hat{V}_{\theta g,p}(\theta_0) W_T(\theta_0)^{-1} \bar{g}_T(y_t, \theta_0)]$$

is $m \times p$ dimensional matrix,

$q_{i,T}(y_t, \theta_0) = \frac{\partial \bar{g}_T(y_t, \theta)}{\partial \theta_i}|_{\theta_0}$ $\quad i = 1, .., p$ is $m \times 1$ dimensional vector and $\hat{V}_{\theta g}(\theta_0) = (\hat{V}_{\theta g,1}(\theta_0) ... \hat{V}_{\theta g,p}(\theta))$ is a $mp \times m$ dimensional matrix that consistently estimate

$$\lim_{T \to \infty} \frac{1}{T} E\left[\left(\underbrace{(q_1(y_t, \theta_0) - E(q_1(y_t, \theta_0)))' ... (q_p(y_t, \theta_0) - E(q_p(y_t, \theta_0)))'}_{m \times T}\right)\left(\underbrace{g(y_t, \theta_0) - E(g(y_t, \theta_0))}_{T \times m}\right)\right]$$

It is extracted from the variance-covariance matrix estimate of

$$V(\theta) = \lim_{T \to +\infty} var\left[\sqrt{T}\begin{pmatrix} g_T(\theta) \\ q_T(\theta) \end{pmatrix}\right] = \begin{pmatrix} V_{gg}(\theta) & V_{g\theta}(\theta) \\ V_{\theta g}(\theta) & V_{\theta\theta}(\theta) \end{pmatrix}$$

In the case of the moment conditions, we consider here, since there is a separation between the theoretical moment evaluated at θ and the empirical moment computed from the data (the difference between the two forms the moment conditions), $D_T(\theta_0, Y)$ is equal to the jacobian of the moment conditions which only depends on θ, such that $V_{\theta g}$ and $V_{\theta\theta}$ are both null matrices, and

$$V(\theta) = \begin{pmatrix} V_{gg}(\theta) & 0 \\ 0 & 0 \end{pmatrix}$$

Under the null hypothesis, which is $H_0 : \theta = \theta_0$, for the K-statistic, it converges to $\chi^2(p)$ and which is $H_m : E(g(y_t, \theta_0)) = 0$ for the J-statistic, it converges to $\chi^2(m-p)$

More specifically,

$$J(\theta_0,\theta_0) = T.\bar{g}(y_t,\theta)' \Big[W_T(\theta_0)^{-1} - W_T(\theta_0)^{-1}D_T(\theta_0,Y)\left(D_T(\theta_0,Y)'W_T(\theta_0)^{-1}D_T(\theta_0,Y)\right)^{-1}$$

$$\times D_T(\theta_0,Y)'W_T(\theta_0)^{-1}\Big]\bar{g}(y_t,\theta)$$

$$K(\theta_0,\theta_0) = T.\bar{g}(y_t,\theta)'W_T(\theta_0)^{-1}D_T(\theta_0,Y)\left[D_T(\theta_0,Y)'W_T(\theta_0)^{-1}D_T(\theta_0,Y)\right]^{-1}$$

$$\times D_T(\theta_0,Y)'W_T(\theta_0)^{-1}\bar{g}(y_t,\theta)$$

Testing the set of strongly identified parameters (β) is more complicated because α^*could be different from α_0 and then the limiting distribution under the null will be unconventional. The Dufour [2003] approach is to construct a valid $(1-\lambda)$ % confidence set for the all vector of parameters or for the weakly identified ones and to project out the other elements. This projection method becomes difficult to apply when the number of parameters is greater than 3[22] . Kleibergen and Mavroeidis [2009] also propose some subset weak instrument robust statistics (S-statistic, KLM, JKLM, MQLR) to test hypothesis and build confidence intervals for all subset of parameters (even if they are not weakly identified). In there setting, if you want to test $H_0 : \beta = \beta_0$ (here there is no more distinction between, strongly and weakly identified parameters), then you can compute the following statistics:

$$S(\beta_0) = S_T(\tilde{\alpha}(\beta_0),\beta_0;\tilde{\alpha}(\beta_0),\beta_0) \underset{a}{\preccurlyeq} \varphi_{p_\beta} + \varphi_{m-p}$$

$$KLM(\beta_0) = K(\tilde{\alpha}(\beta_0),\beta_0;\tilde{\alpha}(\beta_0),\beta_0) \underset{a}{\preccurlyeq} \varphi_{p_\beta}$$

$$JKLM(\beta_0) = S(\beta_0) - KLM(\beta_0) \underset{a}{\preccurlyeq} \varphi_{m-p}$$

$$MQLR(\beta_0) = \frac{1}{2}[KLM(\beta_0) + JKLM(\beta_0) - rk(\beta_0)$$

$$+(\{KLM(\beta_0) + JKLM(\beta_0) + rk(\beta_0)\}^2 - 4JKLM](\beta_0)rk(\beta_0))^{1/2}]$$

$$\underset{a}{\preccurlyeq} \frac{1}{2}[\varphi_{p_\beta} + \varphi_{m-p} - rk(\beta_0) + (\{\varphi_{p_\beta} + \varphi_{m-p} + rk(\beta_0)\}^2 - 4\varphi_{m-p}rk(\beta_0))^{1/2}]$$

where $rk(\beta_0)$ is a statistic that test the lower rank of the jacobian matrix of the objective function at $(\tilde{\alpha}(\beta_0),\beta_0)$ and which is given by :

[22] For example with a vector of $p = 3$ parameters, if we assume each parameter to belong initially to an interval that we discretize into only 10 values, then our grid will be made of 10^ppoints at which the objective function will be evaluated. So as p increases, the number of points in our grid growth exponentially and it becomes very difficult to settle the S-set.

$$rk(\beta_0) = \underset{v \in R^{p-1}}{min} T \begin{pmatrix} 1 \\ v \end{pmatrix}' D_T(\tilde{\alpha}(\beta_0), \beta_0)' \left[\left(\begin{pmatrix} 1 \\ v \end{pmatrix} \otimes I_m \right)' V_{\theta\theta}(\tilde{\alpha}(\beta_0), \beta_0) \left(\begin{pmatrix} 1 \\ v \end{pmatrix} \otimes I_m \right) \right]^{-1}$$

$$\times D_T(\tilde{\alpha}(\beta_0), \beta_0) \begin{pmatrix} 1 \\ v \end{pmatrix}$$

They show that under the null hypothesis $H_0 : \beta = \beta_0$, the limiting distributions of those statistics is asymptotically bounded by the usual chi-squared distributions with the corresponding degrees of freedom: φ_{p_β} and φ_{m-p} are independent chi-squared distributed random variables with p_β and $m - p$ degrees of freedom, respectively. The drawback being that using the corresponding degree of freedom $(1 - \lambda)$ chi-squared quantile to build the confidence intervals will result in conservative confidence intervals when the unrestricted parameters are weakly identified, meaning that the maximum rejection probability of the tests over all values of the nuisance parameters is lower than the size of the test or equivalently the level of the confidence interval will be greater than $(1 - \lambda)$ %.

Table 4.6: Basic moments implied by the model

Moments	Data	RCUGmm	CUGmm	ECUGmm	EECUGmm
$E(\Delta c)$	0.020	0.023 [0.010; 0.036]	0.025 [-0.016; 0.066]	0.0229 [0.008; 0.037]	0.0229 [0.008 ; 0.038]
$\sigma(\Delta c)$	0.03	0.044 [0.035 ; 0.053]	0.043 [0.031 ; 0.059]	0.030 [0.019 ;0.045]	0.032 [0.020 ; 0.049]
$AC1(\Delta c)$	0.359	0.233 [0.018; 0.429]	0.414 [0.133; 0.678]	0.713 [0.555 ; 0.834]	0.694 [0.535 ; 0.819]
$E(\Delta d)$	0.008	0.024 [-0.073 ;0.121]	0.029 [-0.085 ;0.142]	0.0343 [-0.029 ; 0.097]	0.030 [-0.023 ; 0.082]
$\sigma(\Delta d)$	0.110	0.126 [0.098; 0.163]	0.1640 [0.120; 0.219]	0.164 [0.107 ; 0.241]	0.134 [0.086 ; 0.202]
$AC1(\Delta d)$	0.146	0.427 [0.172 ; 0.666]	0.3190 [0.062 ; 0.561]	0.471 [0.246; 0.664]	0.509 [0.298 ; 0.692]
$E(r_m)$	0.075	0.059 [-0.015; 0.134]	0.066 [-0.025 ; 0.157]	0.073 [0.012 ; 0.140]	0.065 [0.013 ; 0.121]
$\sigma(r_m)$	0.200	0.333 [0.272 ;0.400]	0.326 [0.246 ;0.421]	0.290 [0.205 ; 0.397]	0.243 [0.174 ; 0.334]
$AC1(r_m)$	-0.003	-0.015 [-0.245 ;0.214]	-0.024 [-0.269 ;0.221]	0.009 [-0.227 ;0.250]	0.020 [-0.218 ;0.262]
$E(r_f)$	0.005	-0.0255 [-0.032; -0.02]	0.002 [-0.032; 0.033]	0.009 [-0.0019 ; 0.0183]	0.0115 [0.0004 ; 0.0215]
$\sigma(r_f)$	0.039	0.021 [0.0160 ;0.026]	0.0197 [0.011 ;0.034]	0.021 [0.013 ; 0.031]	0.0232 [0.014 ; 0.036]
$AC1(r_f)$	0.563	0.322 [0.093; 0.530]	0.906 [0.780; 0.977]	0.683 [0.511 ; 0.818]	0.650 [0.474 ; 0.791]
$E(p-d)$	3.320	3.364 [2.403;4.339]	3.298 [2.194;4.353]	3.204 [2.808 ; 3.449]	3.274 [2.832 ; 3.524]
$\sigma(p-d)$	0.441	0.715 [0.423 ; 1.155]	0.641 [0.350 ; 1.114]	0.371 [0.218 ; 0.593]	0.336 [0.192 ; 0.554]
$AC1(p-d)$	0.936	0.873 [0.713; 0.958]	0.857 [0.659; 0.960]	0.691 [0.466 ; 0.855]	0.727 [0.501 ; 0.883]

The table displays the results from 10.000 simulations of 948 months each (to mimic the period 1930-2009) of the long run risks model based on the CUE estimates. The statistics are for annually aggregated variables (consumption growth rate, dividend growth rate, market return, risk free rate, log price-dividend ratio). Aggregation has been done through a 12 months summation for growth rates and returns. For the log price-dividend ratio, the annual values have been computed using the year's December values of the price divided by the summation of monthly dividends within the year.

This figure shows the elasticity of the continuously updated Gmm function with respect to the persistence of the volatility (ν) as the volatility of the volatility (σ_w) goes to 0. As σ_w goes to zero, the objective function becomes almost insensitive to changes in ν. All the parameters (except ν and σ_w) have been calibrated as in BYK for the simulation and the sample length is 840 years.

This figure shows the slopes (on the left) and the R-squared (on the right) of the predictive regression of future excess returns (Ex.Ret), consumption growth (g) and dividend growth (gd) by log price-dividend ratio. The horizontal axis

Table 4.5: Predictability of excess returns, consumption, and dividends

The predictive regressions in the data show that log price-dividend predicts excess return but does not predict consumption growth and dividend growth. When using the estimated parameters to calibrate the LRR model, we see that the values obtained under basic moment restrictions (RCUGmm) do not allow to reproduce the level of the observed predictability of excess return of excess returns. Furthermore the implied predictability of the dividend growth is highly rejected. If we use the values obtained in the extended restrictions (ECUGmm and EECUGmm) to calibrate the model, the implied predictability becomes closer to the observed one. Furthermore, implied consumption growth predictability is closer to the data for EECUGmm restrictions (where this moment conditions have been included) compare to CUGmm restrictions

Horizon	$\beta(50\%)$ RCU	$R^2(50\%)$ RCU	$\%\left(\widehat{R^2}\right)$ RCU	$\beta(50\%)$ CU	$R^2(50\%)$ CU	$\%\left(\widehat{R^2}\right)$ CU	$\beta(50\%)$ ECU	$R^2(50\%)$ ECU	$\%\left(\widehat{R^2}\right)$ ECU	$\beta(50\%)$ EECU	$R^2(50\%)$ EECU	$\%\left(\widehat{R^2}\right)$ EECU
	Excess return predictability : $\sum_{j=1}^{J} r^a_{m,t+j} - r^a_{f,t+j} = const. + \beta(p^a_t - d^a_t) + \varepsilon_{t+J}$											
1Y	-0.049	0.010	0.924	-0.068	0.018	0.850	-0.123	0.022	0.797	-0.115	0.028	0.7971
	-0.996			-1.331			-1.198			-1.12		
3Y	0.099	0.029	0.976	-0.156	0.047	0.945	-0.309	0.056	0.931	-0.247	0.074	0.900
	1.160			-1.385			-1.915			-1.995		
5Y	-0.268	0.048	0.961	-0.264	0.072	0.928	-0.649	0.076	0.926	-0.618	0.108	0.888
	-1.289			-1.675			-1.335			-1.903		
	Consumption growth predictability : $\sum_{j=1}^{J}(\triangle c^a_{t+j}) = const. + \beta(p^a_t - d^a_t) + \varepsilon_{t+J}$											
1Y	0.008	0.013	0.077	0.053	0.245	0.0016	0.059	0.410	0.0001	0.044	0.325	0.002
	0.882			4.182			4.749			4.164		
3Y	-0.025	0.025	0.042	0.072	0.363	0.0019	0.1248	0.247	0.003	0.134	0.182	0.008
	-1.186			5.856			4.522			4.513		
5Y	0.031	0.037	0.207	0.134	0.400	0.013	0.0684	0.1436	0.063	0.073	0.104	0.104
	1.428			4.179			2.422			2.468		
	Dividend growth predictability : $\sum_{j=1}^{J}(\triangle d^a_{t+j}) = const. + \beta(p^a_t - d^a_t) + \varepsilon_{t+J}$											
1Y	0.084	0.338	0.000	0.127	0.194	0.003	0.231	0.302	0.0005	0.206	0.277	0.002
	5.032			5.132			5.133			4.792		
3Y	0.352	0.432	0.002	0.353	0.241	0.034	0.417	0.190	0.045	0.242	0.159	0.077
	4.967			3.792			3.506			3.019		
5Y	0.384	0.449	0.004	0.310	0.270	0.034	0.433	0.119	0.091	0.348	0.095	0.120
	6.270			4.377			2.164			2.641		

The column named $\hat{\beta}$ represents the slope and (below) the t-stat of the predictive regression of future excess return, consumption and dividend growth on the log price dividend ratio for different horizons (1 year, 3 years and 5 years). The R-square are also provided in columns named $\widehat{R^2}$ and $R^2(50\%)$. The first one is the R-squared obtained for the observed data, while the second one is the median of the R-squared obtained for regressions on 10.000 simulated data set using the corresponding estimates of parameters (under RCUGmm, CUGmm , ECUGmm and EECUGmm restrictions) to calibrate the model. The column named $\%\left(\widehat{R^2}\right)$ represents the percentage of sample simulated R-squared that are below the R-squared observed in the data; a percentile below 5% or above 95 % correspond to a one side test rejection of the model at 5% significant level.

Table 4.7: Predictability of the volatility of excess returns, consumption, and dividends

Horizon	$\beta(50\%)$ RCU	$R^2(50\%)$ RCU	$\%\left(\widehat{R^2}\right)$ RCU	$\beta(50\%)$ CU	$R^2(50\%)$ CU	$\%\left(\widehat{R^2}\right)$ CU	$\beta(50\%)$ ECU	$R^2(50\%)$ ECU	$\%\left(\widehat{R^2}\right)$ ECU	$\beta(50\%)$ EECU	$R^2(50\%)$ EECU	$\%\left(\widehat{R^2}\right)$ EECU
	Excess return volatility predictability											
1Y	-0.0248	0.006	0.362	-0.033	0.012	0.2712	-0.114	0.032	0.134	-0.079	0.035	0.123
	-0.952			-1.153			-1.709			-1.420		
3Y	-0.046	0.018	0.069	0.090	0.027	0.0585	-0.266	0.061	0.034	-0.125	0.069	0.029
	-0.769			0.822			-1.633			-1.760		
5Y	-0.128	0.029	0.060	0.107	0.037	0.0531	-0.494	0.068	0.036	-0.288	0.082	0.034
	-0.938			1.119			-1.580			-1.825		
	Consumption growth volatility predictability											
1Y	-0.002	0.007	0.9960	-0.003	0.013	0.9656	-0.009	0.061	0.773	-0.013	0.076	0.701
	-0.843			-0.853			-2.344			-2.648		
3Y	-0.010	0.018	0.998	0.010	0.028	0.9852	-0.0296	0.106	0.908	-0.025	0.140	0.861
	-0.873			1.321			-1.800			-3.240		
5Y	0.015	0.029	0.996	-0.019	0.038	0.9859	-0.0384	0.113	0.943	-0.052	0.152	0.906
	1.080			-1.107			-1.692			-3.178		
	Dividend growth volatility predictability											
1Y	-0.012	0.007	0.753	-0.018	0.013	0.5960	-0.074	0.057	0.224	-0.048	0.074	0.162
	-0.968			-0.842			-1.840			-3.279		
3Y	0.019	0.018	0.623	0.0281	0.028	0.5202	-0.197	0.101	0.215	-0.144	0.137	0.151
	1.045			1.056			-2.303			-2.694		
5Y	-0.032	0.029	0.651	-0.064	0.039	0.5752	-0.252	0.108	0.325	-0.164	0.149	0.239
	-1.030			-1.032			-1.815			-3.939		

The table displays the results of model predictive regressions of future excess return volatility, consumption growth volatility and dividend growth volatility on the log price-dividend ratio as in eq. 4.3.7. We run 10.000 simulations of 948 months each (to mimic the period 1930-2009) of the long run risks model when calibrated at the parameters obtained from the estimations. Each column represents the slope and below the t-stat of the predictive regression for different horizons (1 year, 3 years and 5 years). The median R-square are also provided. The last column represents the percentage of sample simulated R-squared that are below the R-squared observed in the data; a percentile below 5% or above 95 % correspond to a one side test rejection of the model at 5% significant level.

Figure 4.2: Illustration of weak identification of ν driven by σ_w

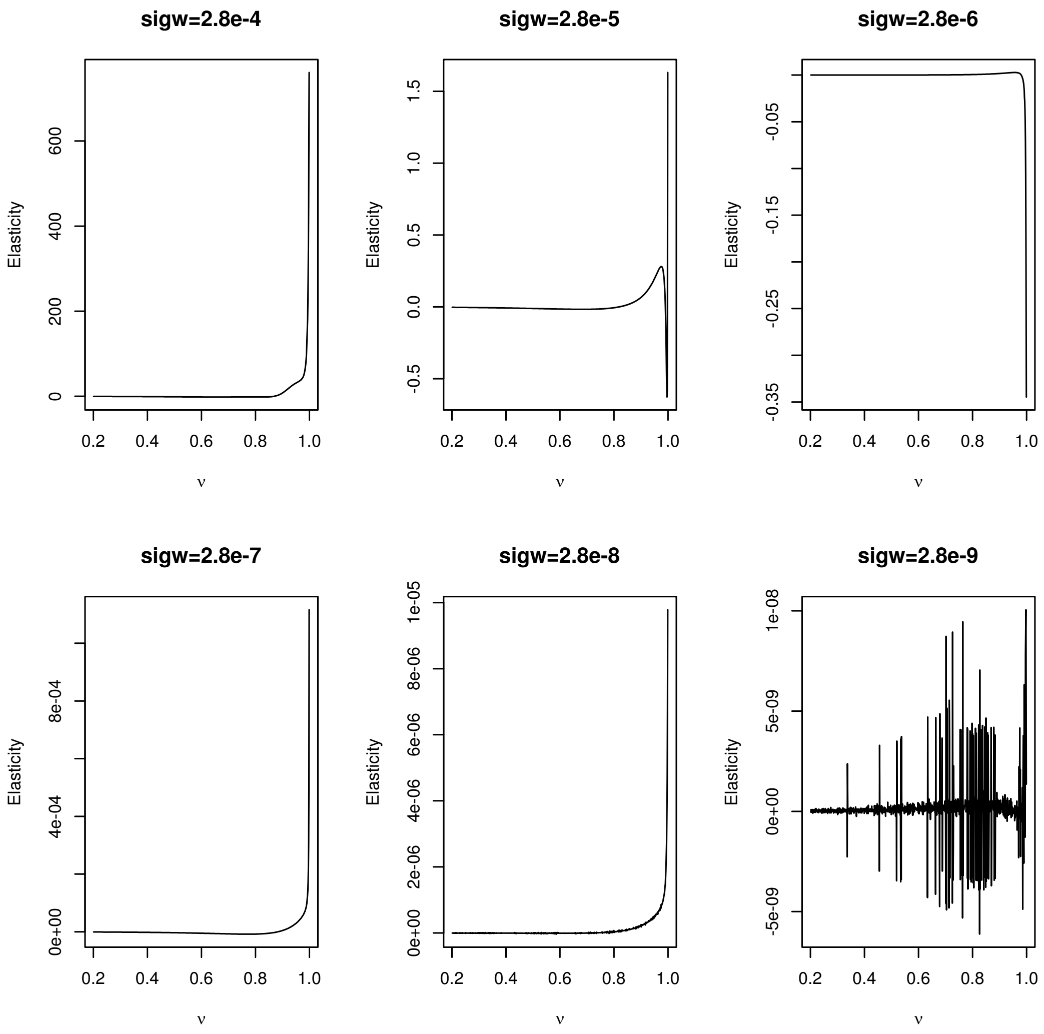
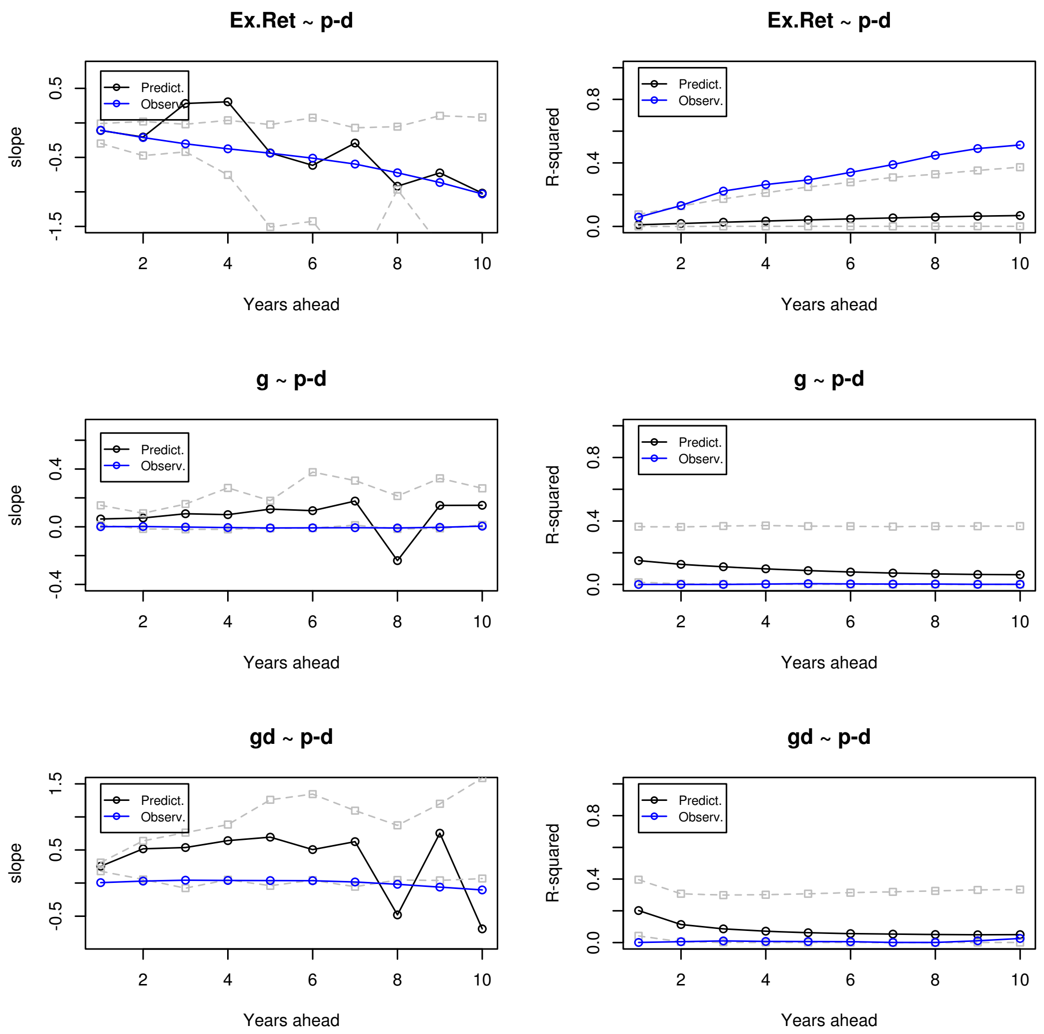
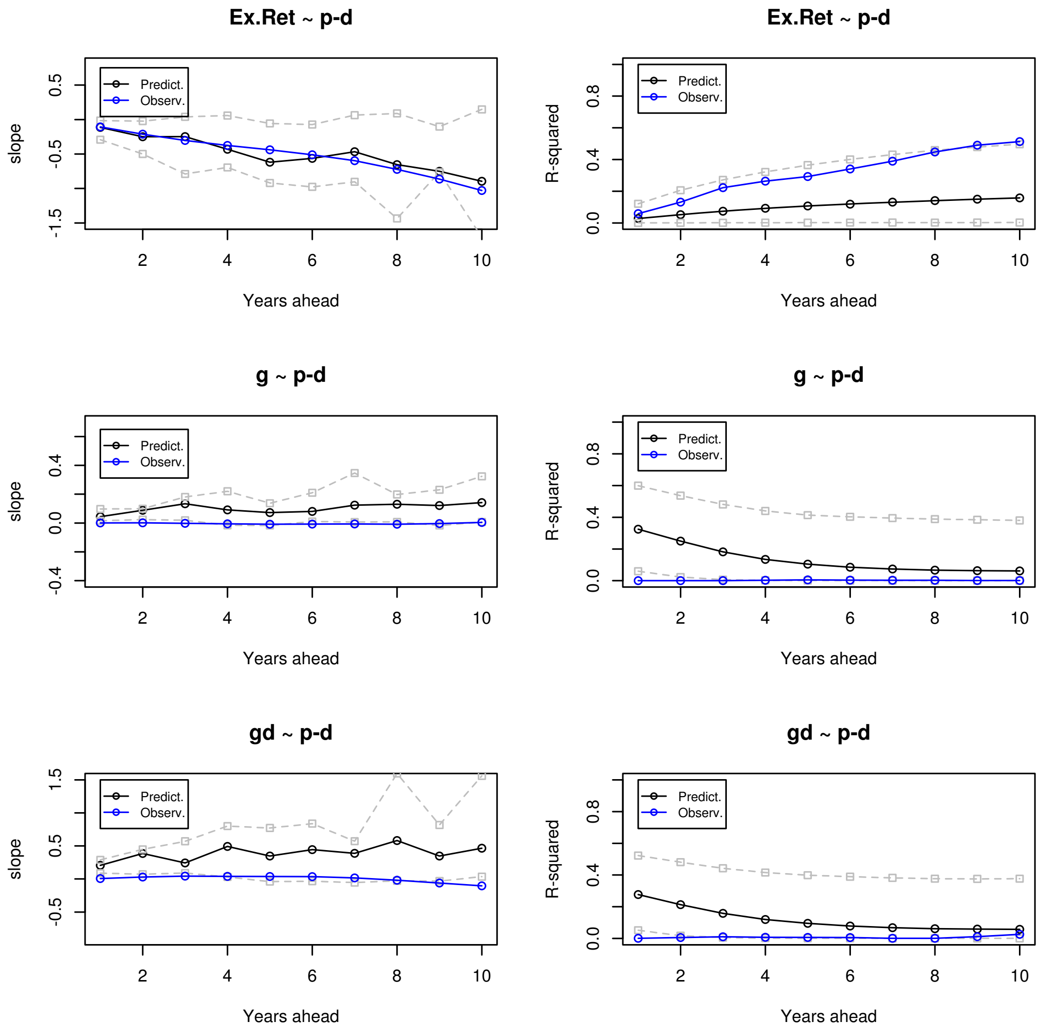


Figure 4.3: Predictability in the BYK calibration of the LRR model



represents the number of ear ahead. The blue line correspond to the observed data (the sample is from 1930 to 2009). The black lines correspond to simulated data using the BYK calibration. The grey lines correspond to the confidence intervals of the black line. We run 10.000 simulations of sample with the same length as the observed data.

Figure 4.4: Predictability in the EECUGmm calibration of the LRR model



This figure shows the slopes (on the left) and the R-squared (on the right) of the predictive regression of future excess returns (Ex.Ret), consumption growth (g) and dividend growth (gd) by log price-dividend ratio. The horizontal axis represents the number of ear ahead. The blue line correspond to the observed data (the sample is from 1930 to 2009). The black lines correspond to simulated data using the estimates obtained from EEUGmm setup for calibration. The grey lines correspond to the confidence intervals of the black line. We run 10.000 simulations of sample with the same length as the observed data.

This figure shows the slopes (on the left) and the R-squared (on the right) of predictive regression of future excess returns volatility (Ex.Ret.Vol), consumption growth volatility (g.Vol) and dividend growth volatility (gd.Vol) by log price-dividend ratio.The horizontal axis represents the number of ear ahead. The blue lines correspond to the observed data (the sample is from 1930 to 2009). The black lines correspond to simulated data using the BYK calibration. The grey lines correspond to the confidence intervals of the black line. We run 10.000 simulations of sample with the same length as the observed data.

This figure shows the slopes (on the left) and the R-squared (on the right) of the predictive regression of future excess returns volatility (Ex.Ret.Vol), consumption growth volatility (g.Vol) and dividend growth volatility (gd.Vol) by log price-

Figure 4.5: Volatility predictability in the BYK calibration of the LRR model

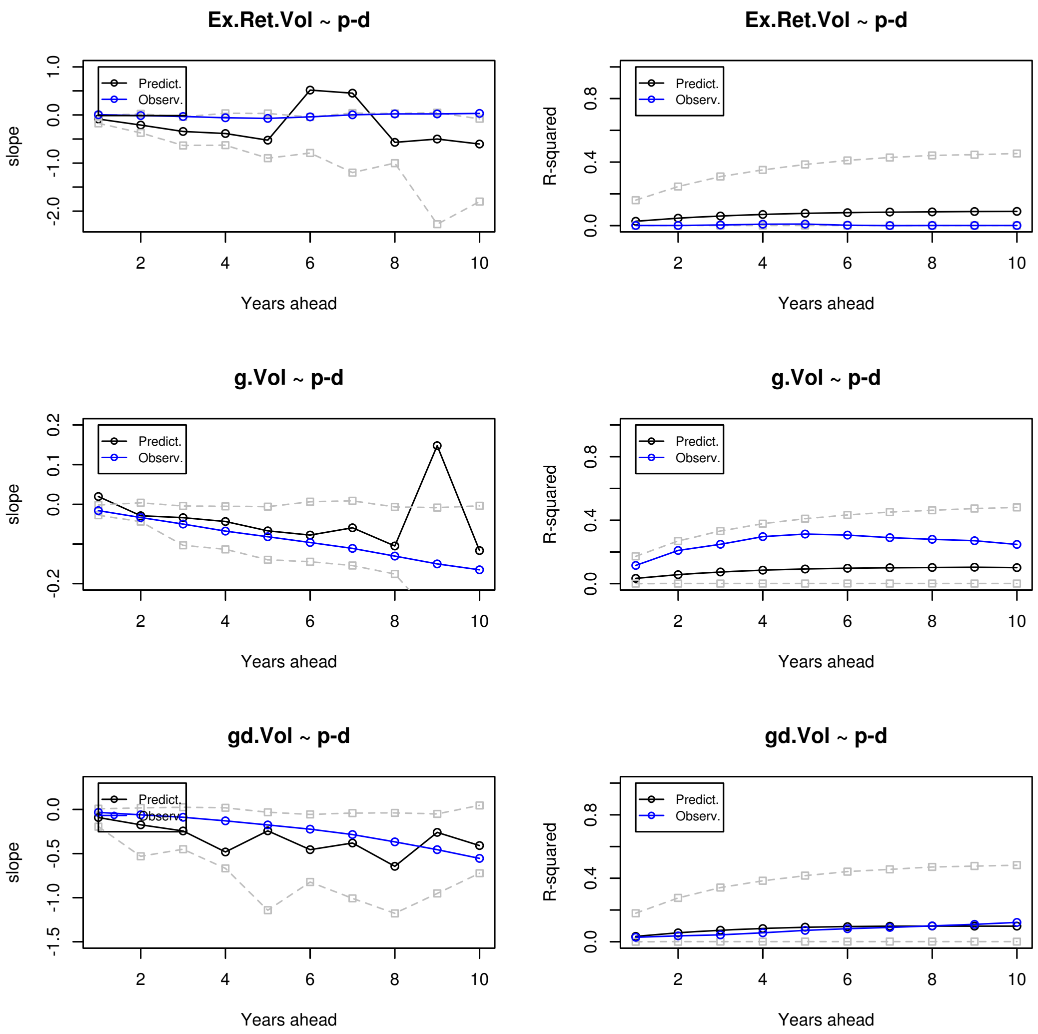
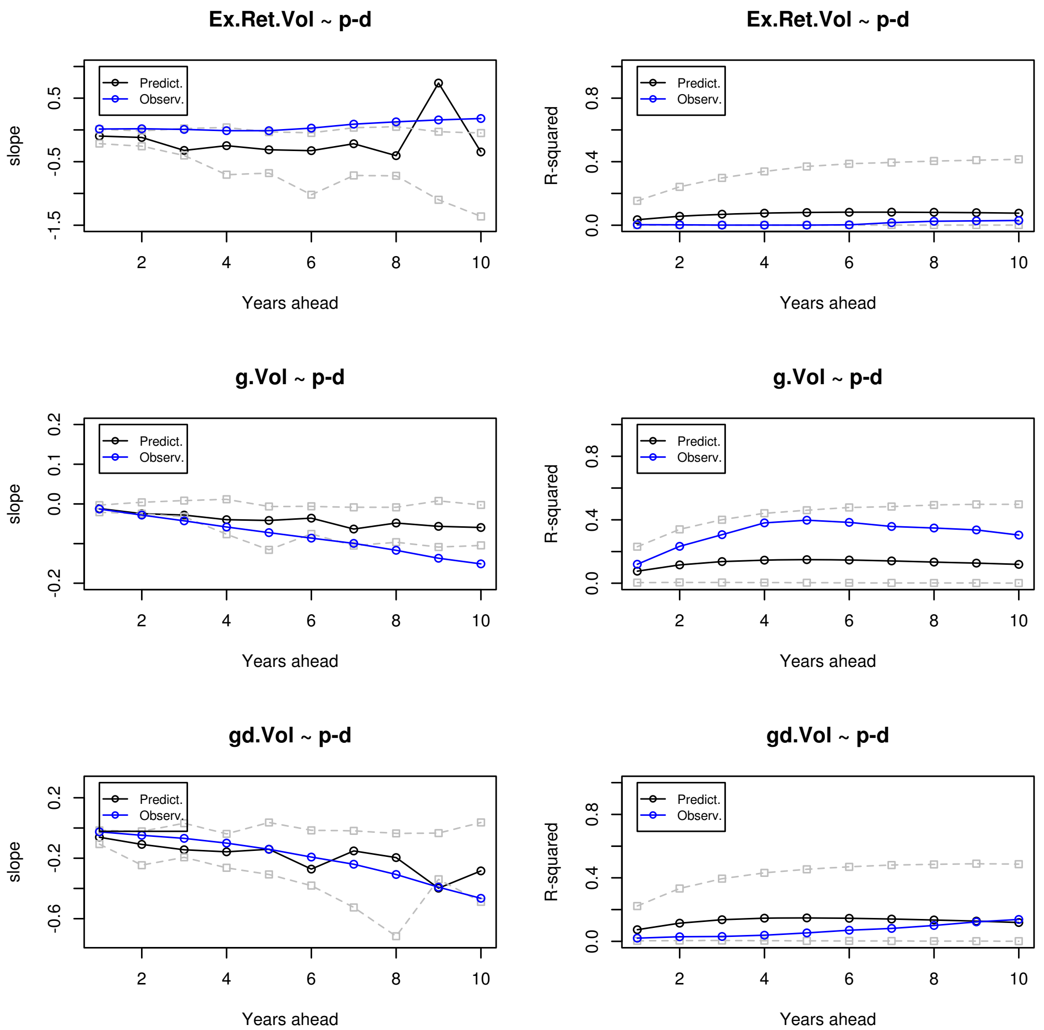
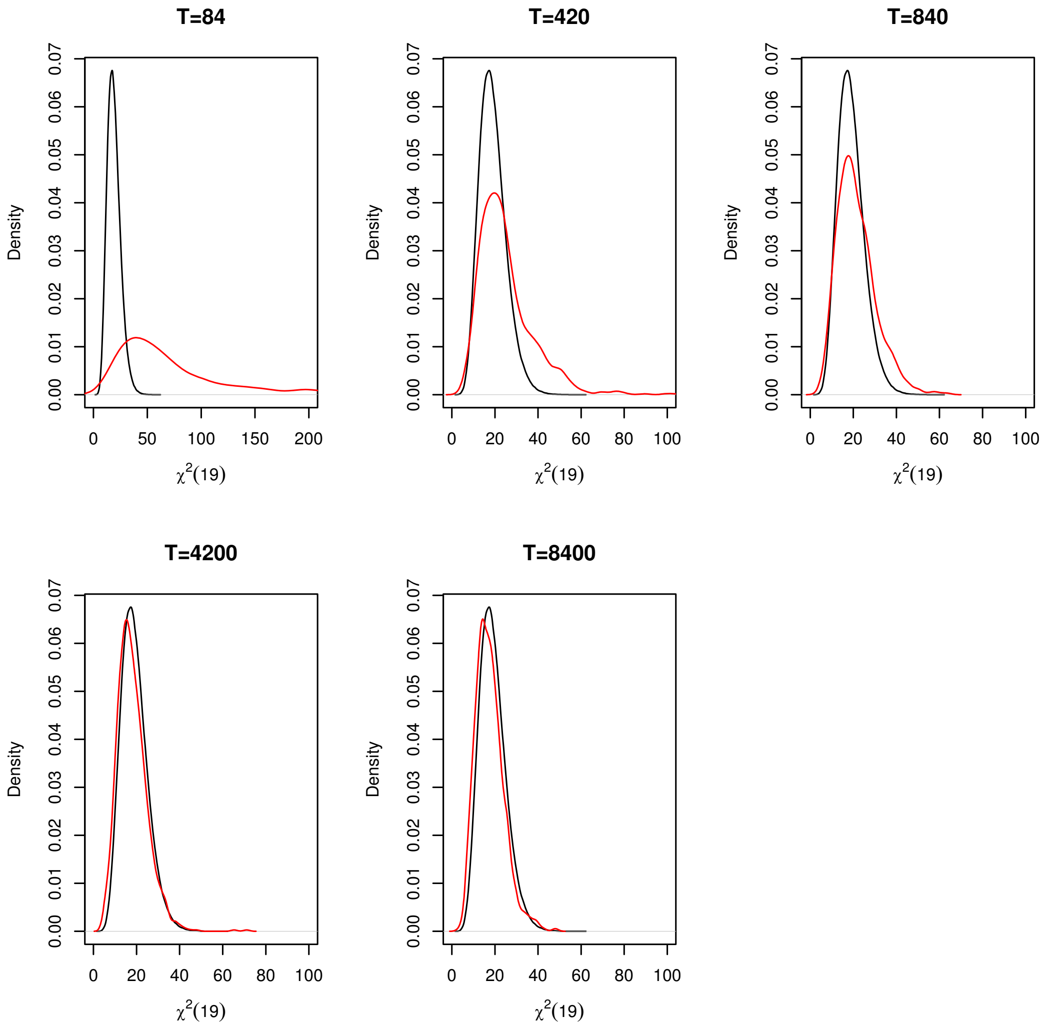


Figure 4.6: Volatility predictability in the EECUGmm calibration of the LRR model



dividend ratio. The horizontal axis represents the number of ear ahead. The blue lines correspond to the observed data (the sample is from 1930 to 2009). The black lines correspond to simulated data using the estimates obtained from EEUGmm setup for calibration. The grey lines correspond to the confidence intervals of the black line. We run 10.000 simulations of sample with the same length as the observed data.

Figure 4.7: Finite sample distribution of the CUGmm objective function in the Model



This figure shows the finite sample distribution of the objective function evaluated at the true value of the parameter for the model simulated with various sample sizes. The number of simulation used to compute the kernel densities (red plots) is the same for all the plots (N=1000), but the sample size of the draws varies: T (in years) =84, 420, 840, 4200, 8400. The black lines represent the density of a chi-squared random variable with 19 degrees of freedom (which correspond to the number of moment conditions). When the sample size is small (e.g: T=84) as what we have in the data, the finite sample approximation of the asymptotic distribution of the CUGmm objective is very poor and that will lead to an over-rejection of the model.

This plot shows the joint 95 % confidence region for γ and others parameters. It has been constructed by selecting all the vectors of parameters for which the CUGmm objective function value is below the 95 % quantile of $\chi^2(7)$. In the CUGmm objective function, we have 19 moment conditions, γ is fixed here and the remaining parameters are estimated.

This plot shows the joint 95 % confidence region for ψ and others parameters. It has been constructed by selecting all

Figure 4.8: 95% Confidence region for γ with respect to the other parameters

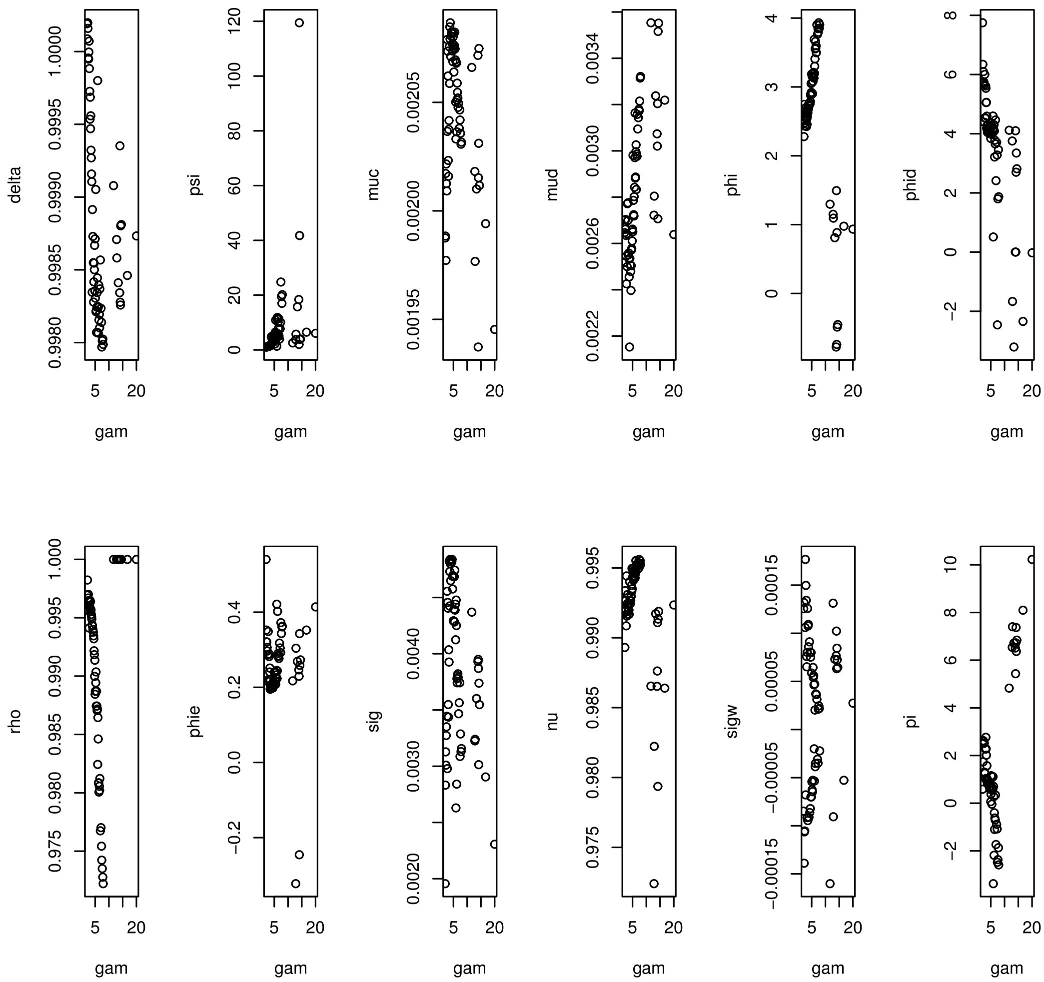
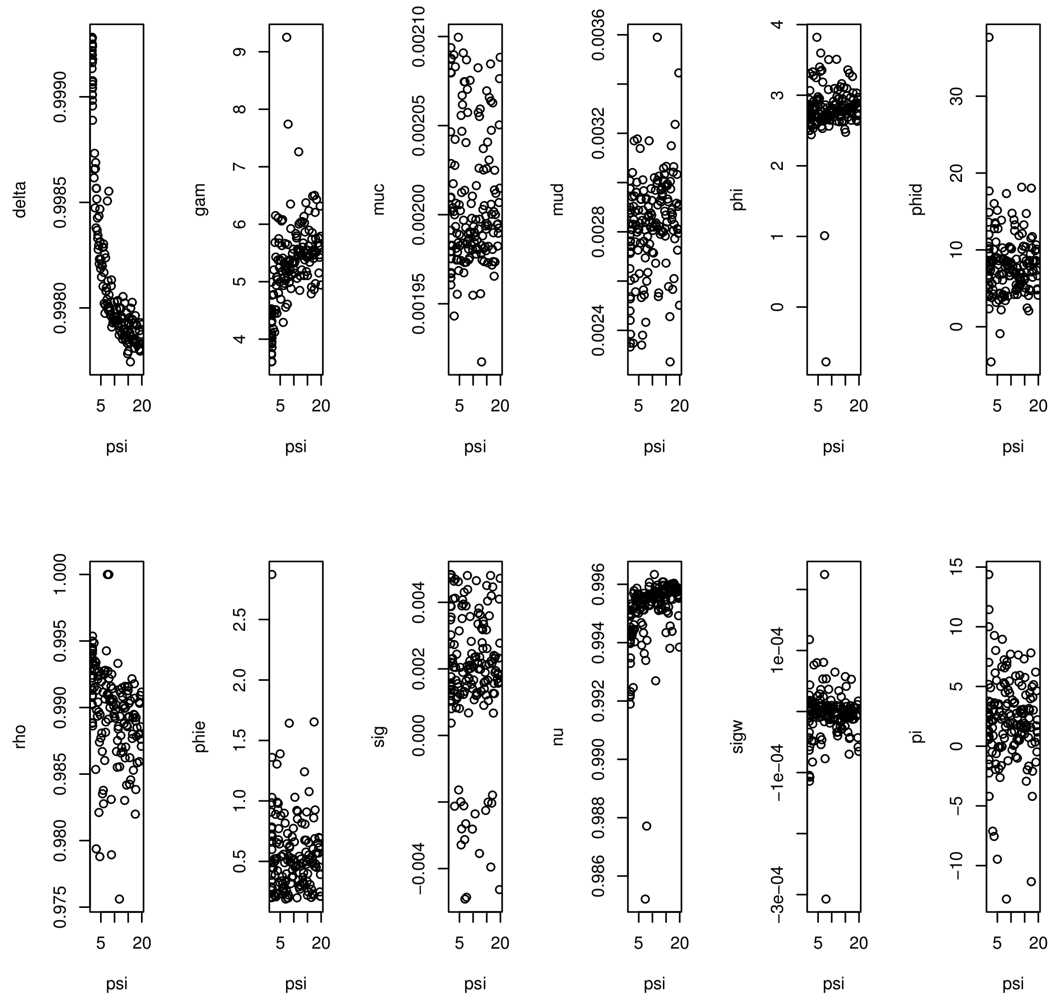
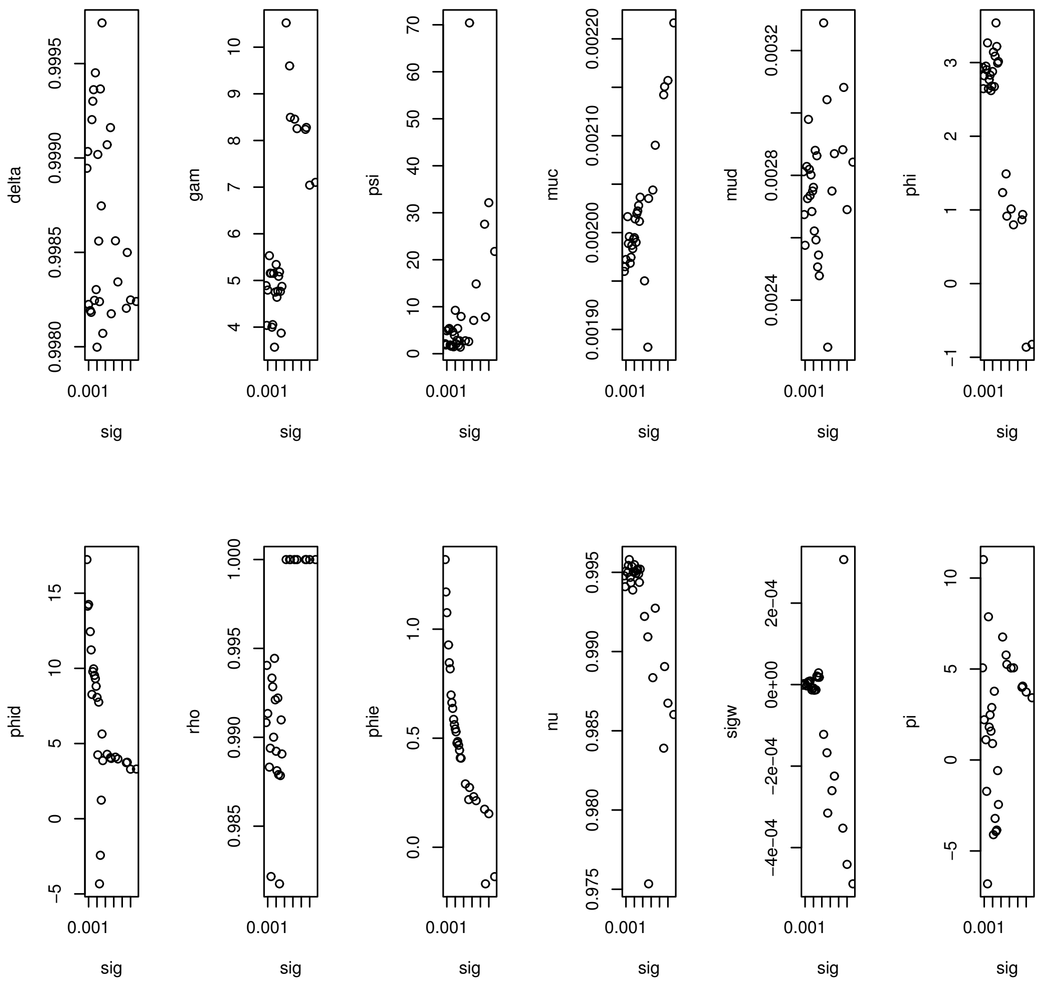


Figure 4.9: Confidence region for ψ with respect to the other parameters



the vectors of parameters for which the CUGmm objective function value is below the 95 % quantile of $\chi^2(7)$. In the CUGmm objective function, we have 19 moment conditions, ψ is fixed here and the remaining parameters are estimated.

Figure 4.10: Confidence region for σ with respect to the other parameters



This plot shows the joint 95 % confidence region for σ and others parameters. It has been constructed by selecting all the vectors of parameters for which the CUGmm objective function value is below the 95 % quantile of $\chi^2(7)$. In the CUGmm objective function, we have 19 moment conditions, σ is fixed here and the remaining parameters are estimated.

Conclusion

This thesis is a contribution to the understanding of asset prices formation and behavior within a consumption based asset pricing model with a fully rational representative agent. Asset prices are the product of the interaction of consumer preferences and the underlying economy shocks. A risk averse investor will ask a higher return as a compensation of the risk taking by holding assets whose value will drop when its marginal utility is high. This principle has faced some difficulties to be verified empirically in the US economy since the observed consumption growth happened to be too smooth to rationalize the high observed level of the equity premium; and it creates the equity premium puzzle (Mehra and Prescott [1985]). As a response, some models where proposed to overcome this apparent contradiction. One of the prominent models is the Long Run Risks model by Bansal and Yaron [2004] which solved the puzzle by highlighting the presence of smooth predictable and slowly moving component in the consumption which is very important for asset prices formation. Indeed, negative shocks to that expected consumption growth component have long lasting effects on the consumer utility (which is recursive) and because of that she might be very reluctant to hold assets which values are positively correlated with the expected consumption growth. The LRR model has mainly being calibrated to match some key macro-finance variables and it has been very successful in that. Even though, this procedure does not allow to statistically test the model and to fully explore its richness. We provide an estimation and inference about the structural parameters of the model using the Generalized Method of Moments. Our inference procedure is robust to weak identification. We find that there is a vast range of admissible parameters that can be used to successfully calibrate the model since it is not statistically rejected. However, the model still face some tension is matching the predictability moments.

Until recently, the models that where designed to solve the equity premium puzzle were agnostic about the term structure of equity returns which means how the average equity premium is decomposed among zero coupon equities with various maturities. van Binsbergen et al. [2012]

pointed out that the equity returns term structure might be downward sloping, meaning that short-term assets earn a higher return compared to long term assets. Their finding where counter intuitive and in contradiction with the leading CCAPM models. Indeed, how to understand the fact that an asset which is expected to pay-off in the short run (and thus look more easily predictable) appears to be riskier than a similar asset that is expected to pay-off more far in the future (and by consequence less predictable) ? We provide an explanation to this puzzle by advocating the fact that cash flows are negatively affected by economic uncertainty which has a higher impact in the short run than in the long run. Our model is able to generate an average downward sloping term structure of equity return while still solving the equity premium puzzle.

The term structure of equity returns is not only upward or downward but is is time varying. More specifically, it looks on average upward sloping during normal times but on average downward sloping during recession period. We propose a regimes switching model in order to capture the changing behavior of the term structure. We show that such a model can deliver various shapes of the term structure by changing the transition probability matrix.

Finally, we test the expectation hypothesis on the equity market. This test has mainly be investigated on the bond market. We provide a counterpart to this test for the equity market. We find that, contrary to the bonds market, the EH is not rejected on the equity market. However the future excess returns on dividends strips are predictable by a combination of forward rate similarly as it was found by Cochrane and Piazzesi [2005] for the bonds market.

As future research, we planned to estimate the regime switching model that we designed to capture the changing behavior of the equity term structure slope. We planned to test of the expectation hypothesis on equity markets internationally.

Bibliography

A. Abel. Asset prices under habit formation and catching up with the joneses. *American Economic Review*, 80(2):38–42, 1990.

D. W. Andrews and X. Cheng. Gmm estimation and uniform subvector inference with possible identification failure. *Econometric Theory*, 30:287–333, 4 2014. ISSN 1469-4360.

D. W. K. Andrews. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):pp. 817–858, 1991.

D. W. K. Andrews and X. Cheng. Estimation and inference with weak, semi-strong, and strong identification. *Econometrica*, 80(5):2153–2211, 2012. ISSN 1468-0262.

M. Andries, T. Eisenbach, and M. C. Schmalz. Horizon-dependent risk aversion and the timing and pricing of uncertainty. (703), 2017.

A. Ang, R. J. Hodrick, Y. Xing, and X. Zhang. The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1):259–299, 2006. ISSN 1540-6261.

P. Augustin and R. Tedongap. Disappointment aversion, term structure, and predictability puzzles in bond markets. *Paris December 2016 Finance Meeting EUROFIDAI - AFFI.*, June 2016.

D. Backus, N. Boyarchenko, and M. Chernov. Term structures of asset prices and returns. (22162), April 2016.

R. Bansal and A. Yaron. Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance*, 59(4):1481–1509, 2004. ISSN 1540-6261.

R. Bansal, A. R. Gallant, and G. Tauchen. Rational pessimism, rational exuberance, and asset pricing models. (13107), May 2007a.

R. Bansal, D. Kiku, and A. Yaron. Risks for the long run: Estimation and inference. *Working paper*, 2007b.

R. Bansal, D. Kiku, and A. Yaron. An empirical evaluation of the long-run risks model for asset prices. *Critical Finance Review*, 1(1):183–221, 2012a. ISSN 2164-5744.

R. Bansal, D. Kiku, and A. Yaron. Risks for the long run: Estimation with time aggregation. (18305), August 2012b.

R. Bansal, D. Kiku, and A. Yaron. Risks for the long run: Estimation with time aggregation. *Journal of Monetary Economics*, 82:52 – 69, 2016.

R. Bansal, S. Miller, and A. Yaron. Is the term structure of equity risk premia upward sloping? *Working paper*, September 2017.

R. J. Barro. Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics*, 121(3):823–866, 2006.

R. J. Barro. Rare disasters, asset prices, and welfare costs. *American Economic Review*, 99(1): 243–64, 2009.

J. Beeler and J. Y. Campbell. Appendix for "the long-run risks model and aggregate asset prices: An empirical assessment". *Critical Finance Review*, 2012.

F. Belo, P. Collin-Dufresne, and R. S. Goldstein. Dividend dynamics and the term structure of dividend strips. *The Journal of Finance*, 70(3):1115–1160, 2015.

N. Bloom. The impact of uncertainty shocks. *Econometrica*, 77(3):623–685, 2009. ISSN 1468-0262.

N. Bloom. Fluctuations in uncertainty. *Journal of Economic Perspectives*, 28(2):153–76, May 2014.

N. Bloom, S. Bond, and J. V. Reenen. Uncertainty and investment dynamics. *Review of Economic Studies*, pages 391–415, 2007.

O. Boguth and L.-A. Kuehn. Consumption volatility risk. *The Journal of Finance*, 68(6): 2589–2615, 2013.

T. Bollerslev, G. Tauchen, and H. Zhou. Expected stock returns and variance risk premia. *Review of Financial Studies*, 22(11):4463–4492, 2009.

C. Burnside. Solving asset pricing models with gaussian shocks. *Journal of Economic Dynamics and Control*, 22(3):329 – 340, 1998. ISSN 0165-1889.

L. E. Calvet and V. Czellar. Through the looking glass: Indirect inference via simple equilibria. *Journal of Econometrics*, 185(2):343 – 358, 2015.

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

J. Y. Campbell and R. J. Shiller. Yield spreads and interest rate movements: a bird's eye view. *Review of Economic Studies*, (58):495–514, 1991.

CBOE. The cboe volatility index-vix. *White paper*, 2015.

J. Cochrane. *ASSET PRICING Revisited Edition*. Princeton University Press, 41 William street Princeton, New Jersey 08540, 2001.

J. H. Cochrane. Macro-finance*. *Review of Finance*, 21(3):945–985, 2017.

J. H. Cochrane and J. Campbell. By force of habit: A consumption-based explanation of aggregate stock market behavior. (3119444), 1999.

J. H. Cochrane and M. Piazzesi. Bond risk premia. *American Economic Review*, 95(1):138–160, March 2005.

G. Constantinides. Habit formation: a resolution of the equity premium puzzle. 1990.

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

D. D. Creal and J. C. Wu. Bond risk premia in consumption-based models. *Working paper*, 4 2015.

M. M. Croce, M. Lettau, and S. C. Ludvigson. Investor information, long-run risk, and the term structure of equity. *Review of Financial Studies*, 2014.

O. de Groot. Solving asset pricing models with stochastic volatility. *Journal of Economic Dynamics and Control*, 52:308 – 321, 2015. ISSN 0165-1889.

A. K. Dixit and R. S. Pindyck. *Investment under Uncertainty*. Princeton University Press, 1994. ISBN 9780691034102.

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

J.-m. Dufour. Identification, weak instruments, and statistical inference in econometrics. *Journal of Econometrics*, pages 767–808, 2003.

J.-m. Dufour. Comments on "weak instrument robust tests in gmm and the new keynesian phillips curve" by f. kleibergen and s. mavroeidis. *Journal of Business & Economic Statistics*, 27(3):318–321, 2009.

L. G. Epstein and S. E. Zin. *Substitution, Risk A version, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework*, volume 57. July 1989.

E. Fama and R. R. Bliss. The information in long-maturity forward rates. *American Economic Review*, 77(4):680–92, 1987.

E. F. Fama and K. R. French. Dividend yields and expected stock returns. *Journal of Financial Economics*, 22(1):3 – 25, 1988. ISSN 0304-405X.

R. Garcia, N. Meddahi, and R. Tédongap. An analytical framework for assessing asset pricing models and predictability. *Unpublished Paper, Edhec Business School, Imperial College London, and Stockholm School of Economics*, 2008. ISSN 0165-1889.

I. Ghattassi and N. Meddahi. Time-aggregation effects on estimating asset pricing models. Working paper, 2012.

L. R. Glosten, R. Jagannathan, and D. E. Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5): 1779–1801, 1993. doi: 10.1111/j.1540-6261.1993.tb05128.x.

C. Gollier. Stochastic volatility implies fourth-degree risk dominance: Applications to asset pricing. *Working paper*, 2017.

N. J. Gormsen. Time variation of the equity term structure. *Working paper*, October 2017.

C. Gourieroux, A. Monfort, F. Pegoraro, and J.-P. Renne. Regime switching and bond pricing. *Journal of Financial Econometrics*, 12(2):237–277, 2014.

L. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–54, 1982.

L. Hansen, J. Heaton, and A. Yaron. Finite-sample properties of some alternative gmm estimators. *Journal of Business & Economic Statistics*, 14(3):262–80, 1996.

M. Hasler and R. Marfé. Disaster recovery and the term structure of dividend strips. *Journal of Financial Economics*, 122(1):116 – 134, 2016. ISSN 0304-405X.

R. J. Hodrick. Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *Review of Financial studies*, 5(3):357–386, 1992.

A. Holly, A. Monfort, and M. Rockinger. Fourth order pseudo maximum likelihood methods. *Journal of Econometrics*, 162(2):278–293, June 2011.

C. Julliard and A. Ghosh. Can rare events explain the equity premium puzzle? *Review of Financial Studies*, 2012.

K. Jurado, S. C. Ludvigson, and S. Ng. Measuring uncertainty. *American Economic Review*, 105(3):1177–1216, March 2015.

B. Kelly, L. Pastor, and P. Veronesi. The price of political uncertainty: Theory and evidence from the option market. *The Journal of Finance*, 71(5):2417–2480, 2016. ISSN 1540-6261.

F. Kleibergen. Testing parameters in gmm without assuming that they are identified. *Econometrica*, 73(4):1103–1123, 2005. ISSN 1468-0262.

F. Kleibergen and S. Mavroeidis. Inference on subsets of parameters in gmm without assuming identification. 2009.

S. Leduc and Z. Liu. Uncertainty shocks are aggregate demand shocks. *Journal of Monetary Economics*, 82:20 – 35, 2016. ISSN 0304-3932.

M. Lettau and J. A. Wachter. Why Is Long-Horizon Equity Less Risky? A Duration-Based Explanation of the Value Premium. *Journal of Finance*, 62(1):55–92, 02 2007.

M. Lettau and J. A. Wachter. The term structures of equity and interest rates. *Journal of Financial Economics*, 101(1):90–113, July 2011.

J. Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, 47(1):13–37, 1965.

R. Lucas. Asset prices in an exchange economy. *Econometrica*, 46(6):1429–45, 1978.

M. Maggiori, S. Giglio, and J. Stroebel. Very long-run discount rates. *Quarterly Journal of Economics*, 130(1):1–53 (Lead Article), 2015.

R. Marfe. Income insurance and the equilibrium term-structure of equity. *Journal of Finance, Forthcoming*, 2016.

N. Meddahi and J. Tinang. Gmm estimation of the long run risks model. 2016 Meeting Papers 1107, Society for Economic Dynamics, 2016.

R. Mehra. The Equity Premium: Why is it a Puzzle? NBER Working Papers 9512, National Bureau of Economic Research, Inc, Feb. 2003.

R. Mehra and E. C. Prescott. The equity premium: A puzzle. *Journal of Monetary Economics*, 11(15):145–161, March 1985.

J. A. Nelder and R. Mead. A simplex method for function minimization. *Computer Journal*, 7 (4):308–313, 1965.

W. Newey and K. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–08, 1987.

W. K. Newey and D. McFadden. Large sample estimation and hypothesis testing. In R. F. Engle and D. McFadden, editors, *Handbook of Econometrics*, volume 4 of *Handbook of Econometrics*, chapter 36, pages 2111–2245. Elsevier, 1986.

W. K. Newey and K. D. West. Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies*, 61(4):631–53, October 1994.

T. A. Rietz. The equity risk premium a solution. *Journal of Monetary Economics*, 22(1):117 – 131, 1988. ISSN 0304-3932.

T. Santos and P. Veronesi. Labor income and predictable stock returns. *Review of Financial Studies*, 19(1):1–44, 2006.

F. Schorfheide, D. Song, and A. Yaron. Identifying long-run risks: A bayesian mixed-frequency approach. (20303), July 2014.

W. F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442, 1964.

R. Shiller. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71(3):421–36, 1981.

J. Stock, M. Yogo, and J. Wright. A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics*, 20:518 – 529, 2002.

J. H. Stock and J. H. Wright. Gmm with weak identification. *Econometrica*, 68(5):1055–1096, 2000. ISSN 1468-0262.

G. Tauchen. Stochastic volatility in general equilibrium. *Quarterly Journal of Finance*, 01(04): 707–731, 2011.

J. Tinang. Macro uncertainty and the term structure of the risk premium. *Working paper*, 2018.

J. van Binsbergen, M. Brandt, and R. Koijen. On the timing and pricing of dividends. *American Economic Review*, 102(4):1596–1618, June 2012.

J. van Binsbergen, W. Hueskes, R. Koijen, and E. Vrugt. Equity yields. *Journal of Financial Economics*, 110(3):503 – 519, 2013.

J. H. van Binsbergen and R. S. Koijen. The term structure of returns: Facts and theory. *Journal of Financial Economics*, (21234), 2016.

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

J. A. Wachter. Can time-varying risk of rare disasters explain aggregate stock market volatility? *The Journal of Finance*, 68(3):987–1035, 2013. ISSN 1540-6261.

P. Weil. The equity premium: A puzzle. *Journal of Monetary Economics*, 24(15):401–421, November 1989a.

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

R. F. Whitelaw. Time variations and covariations in the expectation and volatility of stock market returns. *The Journal of Finance*, 49(2):515–541, 1994. doi: 10.1111/j.1540-6261.1994.tb05150.x.

R. F. Whitelaw. Stock market risk and return: An equilibrium approach. *The Review of Financial Studies*, 13(3):521–547, 2000. doi: 10.1093/rfs/13.3.521.

D. Wilcox. The construction of u.s. consumption data: Some facts and their implications for empirical work. *American Economic Review*, 82(4):922–41, 1992.

S. William F. Adjusting for risk in portfolio performance measurement. *The Journal of Portfolio Management*, 1(2):29–34, Winter 1975.

G. Zhou and Y. Zhu. Macroeconomic volatilities and long-run risks of asset prices. *Management Science*, 61(2):413–430, 2015.