Macro Uncertainty and the Term Structure of Risk Premium

Jules Tinang*

January 15, 2018

Abstract

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.

*I am very grateful to Marianne Andries, Christian Gollier and Nour Meddahi for their advices, guidance and encouragements. I also thank, Patrick Fève, René Garcia, Sophie Moinas, Sébastien Pouget, Roméo Tédongap, participants to TSE Econometrics, Macro and Finance workshops, seminar participants at IAAE 2017, EEA-ESEM 2017 for helpful comments. I gratefully acknowledge the financial support of the ERC grant n° 295208. All errors are mine.

Toulouse School of Economics, UT1 Capitole, Manufacture des Tabacs, 21 Allée de Brienne 31000 Toulouse – France. Email: jules.tinang@tse-fr.eu. An update version of the paper available here: https://sites.google.com/view/jules-tinang/research
1 Introduction

During the past 30 years, a lot 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 [1989]. 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; we can think of a stock or a stock index as a portfolio of zero-coupon equity with maturities ranging from 1 to infinity, each zero-coupon with a given maturity paying a unit of the stock dividend only at its maturity and nothing else. Some recent empirical studies have shown 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 terms, 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 appearing in the distant future. Even though it is still controversial, reconciling asset pricing models with the downward sloping term structure of equity returns has recently become a very active area of research.

\(^1\)Those puzzles refer to the inability of the basic 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 to extract the term structure of the risk premium on this market 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.
In this paper, I argue that allowing for a negative correlation between the level of the cash flow growth and its volatility (the “leverage effect”) could enable these models to reproduce a downward sloping term structure of equity returns while still matching the historical levels of equity premium and risk free rate. I show that allowing for cash flows to be negatively correlated with stochastic volatility could imply a decreasing term structure of the cash flow volatility, which can be seen through a decreasing term structure of the variance ratio statistic. This modification brings in two additional parameters and enables to match the three 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.

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; 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 increasing its price of risk and amplifying the short term exposure of cash flows to the priced risks. Indeed, 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.\textsuperscript{5} 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

\textsuperscript{5}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.
especially during recession periods simultaneously with a drop of the cash flow as we can see from figure 1. I confirmed this visual observation by running a reduced form VAR model with VIX and cash flows (consumption and dividend) 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 theoretical unconditional moments of the variables involved in the model (see Table 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 to see whether it is possible to find a valid vector of parameters implying a decreasing term structure of the equity risk premium and also because the moment conditions, more precisely the empirical counterparts, are not informative about the term structure of the risk premium. I overcome this lack of information by using the synthetic dividend strip prices constructed by van Binsbergen et al. [2012] and adding some moment conditions on the dividend strips returns, but this came at the price of losing the closed form solutions of the theoretical moments. So, I relied on the Simulated Method of Moment (SMM) for the estimation.

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 feedback 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 comes 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]6, 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

6They only allow dividend to load on short run consumption growth shocks, while the most important loading in our case is on the volatility shocks.
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 some implications concerning the unobserved component of the wealth portfolio on 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.

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, proxies by the VIX index, and cash flows shocks.

Macroeconomic effects of uncertainty has 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]). Using a VAR model, he found that uncertainty shocks lead to a sharp decrease in output and productivity (within the 4 months).
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 shows the evolutions of the VIX index, aggregate dividend 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 decreases (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 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.

Figure 1: Cash flows growth and Stock market Implied Volatility

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 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.
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 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.

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 of the consumption growth process. 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.
3.1 Preferences and cash flow dynamics

I consider a rational representative agent embedded with Epstein and Zin [1989] recursive utility function given by equation 1 who maximizes its continuation value subject to its inter-temporal budget constraint.

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

(1)

Where \( \delta \) is the pure discount factor, \( \gamma \) is the relative risk aversion coefficient and \( \psi \) 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 \( \gamma = \frac{1}{\psi} \).

The cash flows dynamics are given by:

\[ \Delta c_{t+1} = \mu_c + x_t + \sigma_c\varepsilon_{c,t+1} + \varphi_\sigma\sigma_w\varepsilon_{w,t+1} \]  

(2)

\[ x_{t+1} = \rho x_t + \varphi_c\sigma_t \varepsilon_{x,t+1} \]  

(3)

\[ \sigma_{t+1}^2 = \nu \sigma_t^2 + (1-\nu)\sigma^2 + \sigma_w\varepsilon_{w,t+1} \]  

(4)

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

(5)

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 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 3 describes the persistent component of the consumption growth process as an AR(1) process. Equations 4 presents stochastic volatility process. Finally equation 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 2-5 embed\(^8\) the one present in the LRR model of Bansal and Yaron [2004]. \( \varepsilon_{c,t+1}, \varepsilon_{d,t+1}, \varepsilon_{x,t+1} \) and \( \varepsilon_{w,t+1} \) can be obtained by setting the restrictions: \( \varphi_\sigma = \pi_\sigma = 0 \)

---

\(^7\)The full definition of the utility function is given by equation 45.

\(^8\)The dynamics in Bansal and Yaron [2004] can be obtained by setting the restrictions: \( \varphi_\sigma = \pi_\sigma = 0 \)
Yaron [2004] as a special case. The key difference lies in the facts that I have introduced the possibility of explicit correlations between the stochastic volatility process and the cash flows dynamics\(^9\). These correlations, highlighted in red in equations 2 and 5, when negative imply a negatively skewed cash flows growth which are more consistent with observed consumption and dividend growth processes\(^10\).

### 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 \(EIS \neq 1\) are derived in the appendix. The solution to the model is standard in the literature. Assuming that the log-value consumption ratio \(\log(c_t)\) is an affine function of the state variables:

\[
\log(c_t) = A_0 + A_1 x_t + A_2 \sigma_t^2
\]

and substituting 6 into the log-price consumption ratio allows to obtain the coefficients of the affine function as follows:

\[
\begin{align*}
A_1 &= \frac{\delta}{1 - \delta \rho} \\
A_2 &= \frac{\delta(1 - \gamma)}{2(1 - \delta \nu)} \left[ 1 + \left( \frac{\delta \varphi}{1 - \delta \rho} \right)^2 \right] \\
A_0 &= \frac{\delta}{1 - \delta} \left[ \mu_c + A_2 (1 - \nu) \sigma^2 + \frac{1}{2} (A_2 + \varphi \sigma) \sigma_w^2 \right]
\end{align*}
\]

So 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).

\(^9\)This correlation matrix shows the key difference between our specification and the one made by Drechsler and Yaron [2011]. Indeed, the consider cash flows and volatility dynamics that allow the possibility of jumps and which nest the dynamics in Bansal and Yaron [2004] as a special case. But there is no correlation between the volatility process and the other processes (consumption growth, dividend growth or conditional expected consumption growth).

\(^{10}\)See the summary statistics in Table 3.
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.

\[ m_{t+1} = a_{0m} + a_{1m}x_t + a_{2m}\sigma_t^2 + \lambda_c\sigma_t\varepsilon_{c,t+1} + \lambda_x\sigma_t\varepsilon_{x,t+1} + \lambda_w\sigma_w\varepsilon_{w,t+1} \] (8)

Where

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

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 9, 10 and 11.

\[ -\lambda_c = \gamma \] (9)
\[ -\lambda_x = (\gamma - 1)\frac{\delta\varphi_e}{1 - \delta\rho} \] (10)
\[ -\lambda_w = \gamma\varphi_{\sigma} + (\gamma - 1)A_2 \] (11)

On one hand, as in standard calibrations of the LRR model, I maintain the rescaling parameter \( \varphi_e \) that governs the variance of the expected consumption growth to be positive. The prices of the short and the long run risks on consumption growth are both positive. 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 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 once the volatility shock is allowed to affect the consumption growth directly. 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:

$$r_{f,t} = -a_{0m} - \frac{1}{2} \lambda_w \sigma_w^2 - a_{1m} x_t - \left( a_{2m} + \frac{1}{2} \left( \lambda_c^2 + \lambda_x^2 \right) \right) \sigma_t^2$$  \hspace{1cm} (12)

Equation 12 shows that during bad times (recessions with higher uncertainty), the precautionary motive to save becomes higher compared to the one in the standard BY model and the risk free rate is lower because the consumer prefers to shift its wealth portfolio toward risk-less assets.

The return on the aggregate market portfolio can be obtained\textsuperscript{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:

$$\tilde{r}_{m,t+1} = \log \left( \frac{P_{t+1} + D_{t+1}}{P_t} \right) \approx k_{0,m} + k_{1,m} z_{m,t+1} - z_{m,t} + \Delta d_{t+1}$$ \hspace{1cm} (13)

The market risk premium at time $t$ and the excess return volatility are respectively given by 14 and 15.\textsuperscript{12}.

$$r_{p,m,t} = -\lambda_w \beta_{m,w} \sigma_w^2 - \left( \lambda_c \beta_{m,c} + \lambda_x \beta_{m,x} \right) \sigma_t^2$$ \hspace{1cm} (14)

$$\sigma_{m,t} = \sqrt{\beta_{m,w}^2 \sigma_w^2 + \left( \beta_{m,c}^2 + \beta_{m,x}^2 + \beta_{m,d}^2 \right) \sigma_t^2}$$ \hspace{1cm} (15)

Where $\beta_{m,c}, \beta_{m,x}, \beta_{m,w}, \beta_{m,d}$ are given by 94 in the appendix.

\textsuperscript{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 10.3 for the derivation.

\textsuperscript{12}See Appendix 10.2 for details.
4 The term-structures of equity and bond returns

4.1 The term structure of zero-coupon equity

Similarly to the zero coupon bond, the zero coupon equity or dividend strip is an asset that pays the dividend 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 expected value, under the $Q$-measure, of the one period ahead future price of the $n - 1$-periods dividend strip:

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

(16)

with the boundary condition that :

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

(17)

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 $r_{d,t+1}^{(n)} = \log P_{d,t+1}^{(n)}$. Then, we have :

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

(18)

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

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

(19)

So the risk premium is a weighted sum of the risk prices. Under preference for early resolution of uncertainty (which happens when $\gamma > 1/\psi$), 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 :

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

(20)

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. If $\pi_{\sigma} < 0$, meaning that the 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. This is not surprising since we know that in the face of more risk, investor requires 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.

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

$$S_{n,t}^{(1)} = r_p^{(n)} - r_p^{(n-1)}$$

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

$$- \lambda_t \left( \left( \phi - \frac{1}{\psi} \right) \varphi \rho^{n-2} \right) \sigma_t^2$$

where

$$A_2(1) = \frac{1}{2} \left( \pi_c - 1 \right) \left( \pi_c + 1 - 2\gamma \right) + \frac{1}{2} \varphi^2, \quad F = \left( \frac{\phi - \frac{1}{\psi}}{1 - \rho} + \frac{1 - \gamma}{1 - \delta \rho} \right) \varphi c, \quad G = -\frac{1}{2} \left( 1 + \rho \right) \left( \phi - \frac{1}{\psi} \right) \varphi^2$$

The short term spread can be seen 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 21 allows to pin down the sign of the slope of the term structure of equity risk premium.

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 ($\sigma_w$, which is constant) and the weight on the latter depends on the conditional volatility ($\sigma_t$, 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 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 $n = 2$, $S_{2,t}^{(1)} = -\lambda_w A_2(1) \sigma_w^2 - \lambda_x \left( \phi - \frac{1}{\psi} \right) \varphi \sigma_f^2$ as the prices of the volatility risk and the expected consumption growth risk are respectively negative and positive, a necessary condition for $S_{2,t}^{(1)}$ to be negative is that $A_2(1)$ should be positive or $\phi$ 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, $\phi$ 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 ($\rho$). Thus with a lower persistence parameter\textsuperscript{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 22, after the second period only depends on the price of volatility risk and thus will be negative when the volatility

\textsuperscript{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.
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.

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

(22)

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 (21) becomes\(^{14}\):

\[ S^{(1)}_{n,t|\phi=1} = -\lambda_w \left( A_2(1)\nu^{\alpha-2} \right) \sigma_w^2 \]

(23)

Equation 23 shows that the short term spread here only depends on the volatility of volatility 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 \( (\pi_c) \) “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 (\( \nu \)).

**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 24 and 25.

\[ S^{(n)}_t = r_p^{(n)} - r_p^{(1)} = -\lambda_w \left( A_2(n-1) \right) \sigma_w^2 - \lambda_x A_1(n-1) \varphi_e \sigma_t^2 \]  

(24)

\(^{14}\)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 \( \varphi_d = 0 \).

\(^{15}\)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 shock. The loading of dividend growth on consumption growth is considered “large” when \( \pi_c > 2\gamma - 1 \).
\[ S_t = \lim_{n \to \infty} S_t^{(n)} = -\lambda_w (A_2(\infty)) \sigma_w^2 - \lambda_x A_1(\infty) \varphi \sigma_t^2 \]  

(25)

where

\[
A_1(\infty) = \left( \frac{\phi - \frac{1}{\psi}}{1 - \rho} \right) \\
A_2(\infty) = \frac{1}{2(1 - \nu)} \left[ (\pi_e - 1)(\pi_e + 1 - 2\gamma) + \varphi^2_2 + \left( \frac{\phi - \frac{1}{\psi}}{1 - \rho} \right) \left( \frac{\phi - \frac{1}{\psi} + 2(1 - \gamma)\delta}{1 - \delta \rho} \right) \varphi^2 \right]
\]

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 \( |\lambda_w (A_2(\infty)) \sigma_w^2| > |\lambda_x A_1(\infty) \varphi \sigma_t^2| \).

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 26 and we can see the term spread will be positive once \( \phi > \frac{1}{\psi} \) and it increases with the horizon. This is the result found 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.

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

(26)

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. Using the no-arbitrage condition:

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

(27)

I obtain the closed form expression in equation 28 for \( P_t^{B,(n)} \) with \( P_t^{B,(0)} = 1 \) given that the zero coupon bond pays a single unit at maturity.

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

(28)
The coefficient in front of the expected consumption growth is given by:

$$B_1(n) = -\left(\frac{1 - \rho^n}{1 - \rho}\right)$$ \hspace{1cm} (29)

The coefficient in front of the volatility is given by:

$$B_2(n) = B_2(1) \left(\frac{1 - \nu^n}{1 - \nu}\right) - \frac{1}{2} \left(\frac{\varphi^2}{\nu - \rho^2}\right) \left(\frac{1 + \rho}{1 - \rho}\right) \left(\frac{1 - \nu^n}{1 - \nu} - \frac{1 - \rho^{2n}}{1 - \rho^2}\right)$$

$$+ \left(\frac{\varphi^2}{\nu - \rho}\right) \left[\frac{(\gamma - 1)\delta}{1 - \delta\rho} + \frac{1}{1 - \rho}\right] \left(\frac{1 - \nu^n}{1 - \nu} - \frac{1 - \rho^n}{1 - \rho}\right)$$ \hspace{1cm} (30)

With the starting conditions that $B_2(0) = 0$ and $B_2(1) = (-\frac{1}{2} + \gamma)$. The constant term is given by equation 51.

As equation 28 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 30 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. 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:

$$y_t^{(n)} = -\frac{1}{n} \log P_t^{B_t^{(n)}} = -\frac{1}{n} \left[ B_0(n) + B_1(n)x_t + B_2(n)\sigma_t^2 \right]$$ \hspace{1cm} (31)

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

---

16 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.
restrictions on the parameters making the dividend to be a random variable \((\pi_c = \varphi_d = \phi = \mu_d = 0)\).

5 Data and Estimation

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.

5.1 Imposing a negative constraint on the term-structure’s slope

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 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:

\[
\hat{\zeta} = \arg \min_{\zeta \in \Theta} T \left[ \frac{1}{T} \sum_{t=1}^{T} h(y_t, \zeta) \right] \hat{\zeta} T(\hat{\zeta})^{-1} \left[ \frac{1}{T} \sum_{s=1}^{T} h(y_s, \zeta) \right] \tag{32}
\]

s.t

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

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 5.1:

| Table 1: Structural parameters and moment conditions for the GMM |
|------------------|------------------|
| **Model’s structural parameters** | **Moment conditions** |
| \( \delta, \gamma, \psi \) | \( E(g^a_t), \text{Var}(g^a_t), \text{ACV1}(g^a_t), E(g^a_t - E(g^a_t))^3 \) |
| \( \mu_c, \mu_d, \varphi, \varphi_c, \varphi_w, \varpi, \varpi_c, \varpi_x, \sigma, \bar{\pi}, \pi_c, \pi_d, \bar{\pi} \) | Consumption growth |
| \( E(g^d_{d,t}), \text{Var}(g^d_{d,t}), \text{ACV1}(g^d_{d,t}), E(g^d_{d,t} - E(g^d_{d,t}))^3 \) | Dividend growth |
| \( E(pd^a_t), \text{Var}(pd^a_t), \text{ACV1}(pd^a_t) \) | Log-price dividend ratio |
| \( E(r^d_{f,t}), \text{Var}(r^d_{f,t}), \text{ACV1}(r^d_{f,t}) \) | Risk free rate |
| \( E(r^a_{m,t}), \text{Var}(r^a_{m,t}), \text{ACV1}(r^a_{m,t}) \) | Market return |
| \( CV(r^{\text{ex},t}_{t+1}, pd^a_t), CV(r^{\text{ex},t+3}_{t+1}, pd^a_t), CV(r^{\text{ex},t+5}_{t+1}, pd^a_t) \) | Excess return predictability |

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

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

\(^{17}\)In the application, I only put \( S^{(\infty)} \) as the constraint to account for the possibility that the average term-structure can be hump shaped.
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 2 the GMM objective function converges to the asymptotic distribution, showing that indeed, the 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.

![Figure 2: Convergence of the objective function to the asymptotic distribution](image)

**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.

Table 6 presents the results of the three groups of estimations that I have run: (i) Standard BY model without the correlations between volatility and cash flows shocks; (ii) Extended Model taking
into account the possible 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. This negative exposure to the volatility in the constrained case compensates for the lower exposure 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 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. A weak identification robust method should have been used to compute the confidence
intervals (e.g: The projection method using 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 esti-
mates 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.

5.2 Recovering the term-structure’s slope from the data

The data on the dividend strips returns with different maturities are informative about the slope of
the term-structure of the equity returns. Unfortunately they are not yet commonly available but can
be computed either from option data or from futures contracts. Dividend strips prices can also be
obtained by computing the present value of dividend futures. Dividend futures or dividend swaps are
contracts allowing the holder of a long position to obtain the dividends paid on the underlying index
over the time (n years) leading up to the settlement, in exchange of a known payment due in n years
(at the expiration). Presently they are only traded on the Over-The-Counter market in the U.S and
publicly traded in Europe and in Japan (Mixon and Onur [2017], Kragt et al. [2015]. The data-set\textsuperscript{18}
provided by van Binsbergen et al. [2012] can be used to infer the slope of the term structure of the
risk premium on U.S financial market; it contains the dividend prices\textsuperscript{19} with maturities 6 months,
12 months, 18 months and 24 months. Those prices are not directly observed but are extracted
from option prices using the put-call parity with option data or the law of one price with futures
contracts as explained by van Binsbergen et al. [2012]. I use this data-set to capture the slope of
the term structure of equity return by adding some first and second order moment conditions on the
returns of the dividend strips with the maturities available in the data. These moment on returns
are computed in the model and their empirical counterpart can be obtained using the data from van

\textsuperscript{18}Available at https://www.aeaweb.org/aer/data/june2012/20101209_data.zip
\textsuperscript{19}Notice that dividend prices are not the same as dividend strip prices. Indeed, the dividend price is the price of a
dividend swap, meaning an up-front payment made in exchange of the realized dividends that will occur between the
inception of the contract and its maturity. On the other side, as we know the dividend strip only pays dividend at
the end of the contract an nothing before that. So, the dividend contract can be seen as a portfolio of dividend strips
with maturities ranging from 1 to the maturity of the dividend contract.
Binsbergen et al. [2012] with maturities ranging from 6 months to 24 months\textsuperscript{20}.

Using the data provided by van Binsbergen et al. [2012] on dividend prices with maturities $n \in \{6 \text{ months}, 12 \text{ months}, 18 \text{ months, 24 months}\}$ and on the monthly dividends we can compute the left side of equation (73) empirically. As we can see from figure 10, the returns on the short term asset (12 month dividends contract) are higher compared to the return on the long-term asset (24 month) especially during periods of boom (first semester of 2001) or burst (June to October 2008). The expression in the right hand of equation (73) can be computed in the model by simulations. We can then match the empirical and the theoretical (obtained from simulations) first order moments in order to capture the slope of the term structure of returns.

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.

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

Figure 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\textsuperscript{21}. 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.

\textsuperscript{20}See section 10.6 in appendix for details.

\textsuperscript{21}The means and Sharpe ratio are computed from the data provided by van Binsbergen et al. [2012].
Figure 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.).

6.2 The timing of risk in the cash flow processes

The annual U.S data represented in figure 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.
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 $\tau$ for the cash flow realisations $(D_\tau)_{\tau=1..T}$ as follows.

$$Skew_t(\tau) = \frac{E_t \left[ \left( \log \left( \frac{D_{t+\tau} - D_t}{D_t} \right) - E_t \log \left( \frac{D_{t+\tau}}{D_t} \right) \right)^3 \right]}{E_t \left[ \left( \log \left( \frac{D_{t+\tau}}{D_t} \right) - E_t \log \left( \frac{D_{t+\tau}}{D_t} \right) \right)^2 \right]^\frac{3}{2}}$$

Under the dynamics assumed in equations 2-5, assuming for simplicity that $\varphi_e = 0$, the time $t$ - skewness over horizon $\tau$ of the dividend becomes:

$$Skew_t(\tau) = \frac{3\pi\sigma \left[ \left( \frac{\pi^2}{\pi} + \varphi^2 \right) \left( \tau - \frac{1}{\nu} \right) \frac{\sigma^2}{1-\nu} \right]}{\left\{ \tau \left( \frac{\pi^2}{\pi} + \varphi^2 \right) \sigma^2 + \pi^2 \sigma^2 \right\}^{\frac{3}{2}}}$$

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\textsuperscript{22}. 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\textsuperscript{23} (VR) statistics used by Belo et al. [2015], Marfe [2016] is given by:

\[ VR_t(\tau) = \frac{\sigma^2_t(\tau)}{\sigma^2_t(1)} \]  

\[ \text{Where} \]

\[ \sigma^2_t(\tau) = \frac{1}{\tau} \left[ \log E_t \left[ D^2(\tau)/D^2(0) \right] - 2 \log E_t \left[ D(\tau)/D(0) \right] \right] \]  

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:

\[ VR_t(\tau) = \frac{1}{\tau} \left[ H_{t-1} \sigma^2_t + (1 - \nu) \left( \sum_{i=0}^{\tau-2} H_i \right) \sigma^2 + \left( 2G_{t-1} - G_{t-1}' \right) \sigma^2_w \right] \]

\[ \frac{H_0 \sigma^2_t + (2G_0 - G_0') \sigma^2_w}{H_0 \sigma^2_t + (2G_0 - G_0') \sigma^2_w} \]  

\textsuperscript{22}Notice that the numerator in the expression of the skewness in 34 corresponds to the term that drives the decreasing term structure of the variance ratio.

\textsuperscript{23}The standard definition of the variance ratio is given by equation 65. When the dividend process is log normal with independent and identically distributed innovations, the two definitions coincide.
Where \( \forall j \in \{1, 2, \ldots \} \)

\[
H_j = n H_{j-1} + \varphi_d^2 + \phi \left( \frac{1 - \rho}{1 - \rho j} \right) \varphi_c^2 + \pi_c^2
\]

\[
G_j = G_{j-1} + [\pi_\sigma + H_{j-1}]^2
\]

\[
G'_j = G'_{j-1} + [\pi_\sigma + \frac{1}{2} H_{j-1}]^2
\]

and \( H_0 = \varphi_d^2 + \pi_c^2, G_0 = G'_0 = \frac{\pi_c^2}{\rho} \).

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 37 becomes:

\[
VR_t(\tau) = \frac{1}{\tau} \left[ \left( \sum_{i=0}^{\tau-1} H_i \right) \right] = \frac{1}{\tau} \sum_{i=0}^{\tau-1} \left[ \varphi_d^2 + \phi \left( \frac{1 - \rho}{1 - \rho i} \right) \varphi_c^2 + \pi_c^2 \right]
\]

(38)

We can easily see that when \( \varphi_c > 0 \), the elements in the summation in the numerator of 38 start with the same value as the denominator and then increase with the horizon. The higher the persistence of the expected consumption growth (\( \rho \)), 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 \( \varphi_c = 0 \) or \( \rho = 0 \).

We now turn to analyze the effect of stochastic volatility by shutting down the persistent component of consumption growth (\( \rho = \varphi_c = 0 \))\(^{24}\). Thus the VR statistic in 37 becomes:

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

(39)

where \( \forall j \in \{1, 2, \ldots \} \)

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

First notice that when there is no volatility persistence (\( \nu = 0 \)), the part of the numerator that depends on \( \sigma_t^2 \) and \( \sigma^2 \) is in expectation similar to the corresponding part in the denominator \((\varphi_d^2 + \pi_c^2) \sigma_t^2 \); but the part that depends on the volatility of volatility (\( \sigma_w \)) starts with the loading of

\(^{24}\) Notice that in this case, VR statistic is constant and the term structure of dividend risk is flat.
the dividend growth on volatility shock ($\pi_\sigma$) as in the denominator and then builds on by adding the term $\left[\pi_\sigma^2 + 3\pi_\sigma (\varphi_d^2 + \pi_c^2) + \frac{7}{4} (\varphi_d^2 + \pi_c^2)^2\right]$. So, if the added term is negative then the VR will decrease as the horizon increases simply because for each supplementary period, the 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 $\pi_\sigma < -\frac{7}{12} (\varphi_d^2 + \pi_c^2)$. 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 ($\pi_\sigma$) is sufficiently negative. The more the persistence of the stochastic volatility process, the lower should $\pi_\sigma$ 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 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 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).
6.3 Implications for the cross-section of returns

In the section 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 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. 9).

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 6 in the case of the LRR model with the BKY calibration, long-duration dividend strips counterfactually 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 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.

Following Lettau and Wachter [2007], we built long-lived assets through the definition of share processes that give at each date the portion of the aggregate dividend paid by the asset as its own dividend; this is done to overcome the difficulty of directly using dividend strips given that it pays dividend only once and thus the price-dividend ratio, which we would like to use to form portfolios, is not well-defined. We specify a deterministic share process that each stock $i$ has in the aggregate
To derive the price of stock $i$ at time $t$ as follows:

$$P_{it}^F = \sum_{t=1}^{\infty} s_{it+n} P_{t}^{(n)}$$

(40)

Figure 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: $r^{(n)}_{d,t} - r_{f,t} = \alpha^{(n)} + \beta^{(n)} [E_t(\sigma_t^2)]^{\frac{1}{2}} + \epsilon_t$. 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

$$^2 s_{it} = \begin{cases} 
(1 + g_s)^i \bar{s} & \text{if } 1 \leq i \leq N/2 \\
(1 + g_s)^{N-i} \bar{s} & \text{if } N/2 < i \leq N 
\end{cases}$$

where $\bar{s} = \left[ 1 + (1 + g_s)^N + 2 \sum_{i=1}^{N/2-1} (1 + g_s)^i \right]^{-1}$ and $g_s$ is the positive constant growth rate of the share. LW set $g_s$ at 5% implying an annual growth rate of 20%.
7 What drives the slope of the term structure of equity returns?

7.1 Back to the simplest case: time separable utility function

The idea that the term structure of equity risk premium is downward sloping 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, notice that with the restriction \( \gamma = \frac{1}{\psi} \) the stochastic discount factor becomes:

\[
m_{t+1} = \log \delta - \gamma \Delta c_{t+1}
\]

The risk premium on a \( n \)-period dividend strip is given by:

\[
E_t(\pi_{d,t+1} - r_{f,t}) + \frac{1}{2} \text{Var}_t(\pi_{d,t+1} - r_{f,t})
= (\gamma \varphi \sigma) \left( \pi_\sigma + \frac{1}{2} \left[ (\pi_\sigma - \gamma)^2 + \varphi_d^2 \right] \left( \frac{1 - \rho^{n-1}}{1 - \rho} \right) + (\phi - 1) \left( \frac{1 - \rho^{n-1}}{1 - \rho} - 1 \right) \right) \frac{F}{\nu - \rho} \tag{41}
\]

\[
+ \left( \frac{1 - \rho^{n-1}}{1 - \rho} - \frac{1 - \rho^{2(n-1)}}{1 - \rho^2} \right) \frac{G_{CRRA}}{\nu - \rho^2} \sigma_w^2 + \gamma \pi_c \sigma_t^2
\]

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:

\[
S_{n,1}^{(1)} = (\gamma \varphi c) \left( \frac{1}{2} \nu^{n-2} \left[ (\pi_c - \gamma)^2 + \varphi_d^2 \right] + (\phi - \gamma) \left[ \left( \frac{\nu^{n-2} - \rho^{n-2}}{\nu - \rho} \right) F_{CRRA} + \left( \frac{\nu^{n-2} - \rho^{2(n-2)}}{\nu - \rho^2} \right) G_{CRRA} \right] \right) \sigma_w^2 \tag{42}
\]

where

\[
F_{CRRA} = \left[ \left( \frac{\phi - \gamma}{1 - \rho} \right) \varphi_c^2 \right] \quad \& \quad G_{CRRA} = -\frac{1}{2} \left( \frac{\nu + 1}{\nu - \rho} \right) (\phi - \gamma) \varphi_c^2
\]
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).

In the first case (i), the risk premium is given by \( r_p^{(n)} = \gamma \pi_c \sigma^2 \) (43), 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.

\[
\begin{align*}
\rho_p^{(n)} &= \gamma \pi_c \sigma^2
\end{align*}
\]

In the second case, the consumption growth process incorporate a stochastic volatility that changes the conditional variance of consumption growth each period. So, there 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 (44) 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 \( (\phi = 0) \), the risk premium is time varying but has the same form as in the i.i.d normal consumption growth case. When \( \phi \neq 0 \), 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.

\[
\begin{align*}
\rho_p^{(n)} &= \gamma \pi_c \sigma^2_t + (\gamma \phi) \left( \pi_\phi + \frac{1}{2} \left[ (\pi_c - \gamma)^2 + \phi^2 \right] \left( \frac{1 - \nu^{n-1}}{1 - \nu} \right) + \frac{1}{2} \left( \phi - \gamma \right)^2 \phi^2 \left( \frac{1 - \nu^{n-2}}{1 - \nu} \right) \right) \sigma_w^2 \text{ for } n \geq 2
\end{align*}
\]

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 \( \nu \), 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 \( (\varphi_\sigma) \) given that all the other terms are positive \( ((\pi_c - \gamma)^2 + \varphi_d^2) \). 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 \( (43) \) and it has a flat term structure. This happens here because the expected consumption growth risk is not priced \( (\lambda_\nu = 0) \), 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.

### 7.2 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 7 represents 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 where this happened 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.

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, $\sigma_w$) to be constant, but this can be made time varying. So the time variation in the term-structure slope could come both 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.

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.

---

See for example Bollerslev et al. [2009], Tauchen [2011].
(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 (level of the volatility vs variability of the volatility).

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.

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 9, 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. 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 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.

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.

---

27 The price of the wealth portfolio denoted by $W_t$ can be expressed as: $W_t = \lim_{T \to \infty} E_t (\sum_{i=1}^{T} M_{t+i} + C_{t+i}) = \lim_{T \to \infty} E_t (\sum_{i=1}^{T} M_{t+i} + D_{t+i} + L_{t+i}) = P^M_t + P^H_t$ where $P^M_t$ and $P^H_t$ stand respectively for the price of market portfolio (including real estate and financial assets) and the price of human capital.

28 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.
that will be used to compute the present value of the expected cash flow. Table 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 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 2: Present value by discount rate

<table>
<thead>
<tr>
<th>Model</th>
<th>discount rate</th>
<th>present value</th>
<th>Term structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAPM</td>
<td>7.5 %</td>
<td>536 millions $</td>
<td>flat</td>
</tr>
<tr>
<td>LRR</td>
<td>11%</td>
<td>22 millions $</td>
<td>increasing</td>
</tr>
<tr>
<td>ELRR</td>
<td>4%</td>
<td>14 billions $</td>
<td>decreasing</td>
</tr>
</tbody>
</table>

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\(^29\), it is useful to have a flexible model in hand that could be able to cope with any situation.

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

\(^{29}\)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 and on long-term assets.
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 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.
References


10 Appendix

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 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 3: Summary statistics

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

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 4: Summary statistics

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

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 5: Estimation of the VAR model with logVIX and consumption growth

<table>
<thead>
<tr>
<th></th>
<th>$\Delta c_t$</th>
<th>log $VIX_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta c_{t-1}$</td>
<td>0.152</td>
<td>-8.517 *</td>
</tr>
<tr>
<td></td>
<td>(0.101)</td>
<td>(4.852)</td>
</tr>
<tr>
<td>log $VIX_{t-1}$</td>
<td>-0.004 *</td>
<td>0.775 ***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.104)</td>
</tr>
<tr>
<td>$\Delta c_{t-2}$</td>
<td>0.253 **</td>
<td>2.933</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(4.71)</td>
</tr>
<tr>
<td>log $VIX_{t-2}$</td>
<td>0.005 *</td>
<td>-0.072</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.131)</td>
</tr>
<tr>
<td>$\Delta c_{t-3}$</td>
<td>0.326 ***</td>
<td>5.851</td>
</tr>
<tr>
<td></td>
<td>(0.101)</td>
<td>(4.862)</td>
</tr>
<tr>
<td>log $VIX_{t-3}$</td>
<td>-0.001</td>
<td>0.148</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.103)</td>
</tr>
<tr>
<td>const.</td>
<td>0.001</td>
<td>0.433 *</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.211)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.31</td>
<td>0.68</td>
</tr>
</tbody>
</table>

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>
<thead>
<tr>
<th></th>
<th>Standard BY Model</th>
<th>Extended Model (EIS=1)</th>
<th>Extended Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
<td>0.9998</td>
<td>0.9983</td>
<td>0.9995</td>
</tr>
<tr>
<td></td>
<td>[0.9251;1.072]</td>
<td>[0.9931;1.004]</td>
<td>[9.985e-01;1.001]</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>10.00</td>
<td>7.296</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>[-45.550;65.550]</td>
<td>[-12.163;26.755]</td>
<td>[-1.84e+01;5.162e+01]</td>
</tr>
<tr>
<td>$\psi$</td>
<td>1.463</td>
<td>15.00</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>[-41.987;44.912]</td>
<td>[-6.175e+02;6.475e+02]</td>
<td>[-1.864e+01;3.522e+01]</td>
</tr>
<tr>
<td>$\mu_c$</td>
<td>2.480e-03</td>
<td>2.058e-03</td>
<td>0.0023</td>
</tr>
<tr>
<td></td>
<td>[2.237e-03;2.723e-03]</td>
<td>[1.806e-03;2.310e-03]</td>
<td>[2.045e-03;2.604e-03]</td>
</tr>
<tr>
<td>$\mu_d$</td>
<td>0.00</td>
<td>3.040e-03</td>
<td>0.00164</td>
</tr>
<tr>
<td></td>
<td>[-1.134e-03;1.134e-03]</td>
<td>[2.027e-03;4.053e-03]</td>
<td>[5.471e-04;2.737e-03]</td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.00</td>
<td>2.311</td>
<td>5.065</td>
</tr>
<tr>
<td></td>
<td>[-16.429;16.429]</td>
<td>[-19.795;24.418]</td>
<td>[-2.361e+04;2.362e+04]</td>
</tr>
<tr>
<td>$\varphi_d$</td>
<td>0.573</td>
<td>2.236</td>
<td>10.054</td>
</tr>
<tr>
<td></td>
<td>[-118.880;120.026]</td>
<td>[-2.121e+03;2.125e+03]</td>
<td>[-1.963e+02;2.164e+02]</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.988</td>
<td>7.947e-01</td>
<td>0.821</td>
</tr>
<tr>
<td></td>
<td>[9.433e-01;1.033]</td>
<td>[-6.608;8.197]</td>
<td>[-8.219e+02;8.236e+02]</td>
</tr>
<tr>
<td>$\varphi_c$</td>
<td>2.721e-02</td>
<td>0.3615</td>
<td>0.0155</td>
</tr>
<tr>
<td></td>
<td>[-5.184e-02;1.063e-01]</td>
<td>[-42.61;43.33]</td>
<td>[-6.242e+01;6.245e+01]</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>4.820e-03</td>
<td>2.796e-03</td>
<td>0.00115</td>
</tr>
<tr>
<td></td>
<td>[2.692e-03;6.947e-03]</td>
<td>[-0.172;0.177]</td>
<td>[-2.085e-02;2.314e-02]</td>
</tr>
<tr>
<td>$\nu$</td>
<td>0.9987</td>
<td>0.997</td>
<td>0.9989</td>
</tr>
<tr>
<td></td>
<td>[0.959;1.039]</td>
<td>[0.996;0.999]</td>
<td>[9.972e-01;1.001]</td>
</tr>
<tr>
<td>$\sigma_w$</td>
<td>3.870e-06</td>
<td>4.609e-06</td>
<td>1.201e-06</td>
</tr>
<tr>
<td></td>
<td>[3.157e-06;4.583e-06]</td>
<td>[-5.409e-04;5.501e-04]</td>
<td>[1.187e-06;2.171e-06]</td>
</tr>
<tr>
<td>$\pi_c$</td>
<td>6.35</td>
<td>10.934</td>
<td>1.612</td>
</tr>
<tr>
<td></td>
<td>[-6.468;19.171]</td>
<td>[-1.124e+03;1.146e+03]</td>
<td>[-3.222e+02;3.255e+02]</td>
</tr>
<tr>
<td>$\varphi_\sigma$</td>
<td>0</td>
<td>0</td>
<td>-3920.078</td>
</tr>
<tr>
<td></td>
<td>[-1.114e+04;3.301e+03]</td>
<td>[-9.784e+03;1.200e+04]</td>
<td>[-4816.642;2269.73]</td>
</tr>
<tr>
<td>$\pi_\sigma$</td>
<td>0</td>
<td>0</td>
<td>-2.854e-16</td>
</tr>
<tr>
<td></td>
<td>[-2.624e+04;2.624e+04]</td>
<td>[-8.037e+04;6.209e+04]</td>
<td>[-1.611e+05;1472.09]</td>
</tr>
<tr>
<td>$T_{JF}$</td>
<td>122.854</td>
<td>41.486</td>
<td>118.447</td>
</tr>
</tbody>
</table>

This table shows the results of the GAM estimations. The first two columns correspond to the standard LRR model, the next two columns are for the extended model but with EIS=1 used along the paper and the last two columns are for the extended model letting the EIS free. The confidence intervals for the first two columns are made using a modified projection method following the procedure explained in 10.10. This will be made for the four first columns later but now it is the Delta method that is used for them.
| par | Standard BY Model | | | Extended Model (EIS=1) | | | Extended Model |
|-----|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| δ   | 0.9988            | 0.9983            | 0.99995           | 1.000             | 0.9995            | 0.9993            |
|     | [0.925;1.072]     | [0.993;1.004]     | [9.985e-01;1.001] |                  | [9.987e-01;1.002] |                  |
| γ   | 10.00             | 7.296             | 10                | 8.286             | 9.969             | 3.302             |
|     | [-45.550;65.550]  | [-12.163;26.755]  | [-3.16e+01;5.16e+01] | [-1.86e+01;3.52e+01] | [8.193;13.35] | [1.569;5.236] |
| ψ   | 1.463             | 15.00             | 1                  | 1                 | 1.178             | 1.724             |
|     | [-41.987;44.912]  |                  | -                 |                  | [1.086;1.156] | [1.010;2.621] |
| μ   | 2.480e-03         | 2.05e-03          | 0.0023            | 1.976e-03         | 2.33e-03          | 2.290e-03         |
|     | [2.237e-03;2.723e-03] | [1.806e-03;2.310e-03] | [2.045e-03;2.604e-03] | [1.737e-03;2.216e-03] | [1.538e-03;3.25e-03] | [1.242e-03;2.762e-03] |
| μ   | [-1.134e-03;1.134e-03] | 2.027e-03;4.053e-03 | 5.41e-04;2.737e-03 | 8.875e-04;3.690e-03 | [-1.74e-03;3.10e-03] | [-2.127e-03;2.516e-03] |
| φ   | 0.00              | 2.311             | 5.065             | 2.758             | 1.794             | 3.605             |
|     | [-16.429;16.429]  |                  |                  |                  | [1.221;6.855] | [1.625;7.325] |
| φ   | 0.573             | 2.236             | 10.054            | 1.957             | 2.942             | 5.351             |
|     | [-118.880;120.026] | [-12.11e+02;2.125e+03] | [-1.963e+02;2.164e+02] | [-1.165e+03;1.169e+03] | [5.04e-01;6.31 | [2.849;8.145] |
| ρ   | 0.988             | 7.94e-01          | 0.821             | 9.83e-01          | 0.827             | 9.900e-01         |
|     | [9.43e-01;1.033]  | [-6.608;8.197]    | [-8.219e+02;8.236e+02] | [8.92e-01;1.075] | [5.46e-01;9.91e-01] | [0.986;9.95e-01] |
| φ   | 0.00              | 0.3615            | 0.0155            | 2.09e-01          | 0.190             | 5.614e-02         |
|     | [-5.184e-02;1.063e-01] | [-4.21e+01;6.245e+01] | [-8.84e-01;1.308] |                  |                  |                  |
| φ   | 0.00              | 2.796e-03         | 0.0115            | 9.94e-04          | 4.15e-03          | 2.896e-03         |
| φ   | [-2.692e-03;6.947e-03] | [-1.72e+02;0.177] | [-2.085e-02;2.314e-02] |                  | [2.26e-03;1.33e-02] | [3.375e-04;1.284e-02] |
| ν   | 0.9987            | 0.997             | 0.9989            | 9.214e-01         | 9.99e-01          | 9.899e-01         |
|     | [0.995;1.039]     | [0.996;0.999]     | [9.97e-01;1.001]  |                  | [9.81e-01;9.99e-01] | [9.87e-01;9.99e-01] |
| σ   | 3.87e-06          | 4.609e-06         | 1.201e-06         | 1.703e-06         | 4.86e-06          | 5.00e-06          |
|     | [3.157e-06;4.583e-06] | [-5.409e-04;5.501e-04] | [1.187e-06;2.121e-06] |                  | [3.738e-06;4.994e-06] | [2.152e-06;4.025e-05] |
| π   | 6.35              | 10.934            | 1.612             | 5.958             | 0.173             | 1.170             |
|     | [-6.468;19.171]   | [-1.124e+03;1.146e+03] | [-3.22e+02;3.255e+02] |                  | [1.44e-01;3.286] | [1.21e-01;2.327] |
| π   | 0                 | 0                 | -3920.078         | 1.107e+03         | -377.981          | 1.617e-03         |
|     |                   |                   | [-1.114e+04;3.301e+03] |                  | [-8.261e+02;6.517e-01] | [-2.016e+02;2.016e+02] |
| π   | 0                 | 0                 | -2.85e-16         | -9.14e+03         | -3.01e+03         | 3.007             |
|     |                   |                   | [-2.62e+04;2.62e+04] |                  | [-6032.854;1.268] | [-2.846e-03;2.846e+03] |
| τ   | 122.854           | 41.486            | 118.447           | 57.265            | 109.215           | 39.516            |

This table shows the results of the GMM estimations. The difference with table 6 is only that the confidence intervals in the last two columns have been computed using parametric bootstrap.
Table 8: Standardized errors

<table>
<thead>
<tr>
<th>Moment</th>
<th>Standard BY Model</th>
<th>Extended Model (EIS=1)</th>
<th>Extended Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(g_t^a)$</td>
<td>2.394</td>
<td>0.250</td>
<td>1.602</td>
</tr>
<tr>
<td>$\text{Var}(g_t^a)$</td>
<td>-1.716</td>
<td>-2.162</td>
<td>-2.362</td>
</tr>
<tr>
<td>ACV1($g_t^a$)</td>
<td>0.960</td>
<td>-0.109</td>
<td>-0.647</td>
</tr>
<tr>
<td>$E(g_t^a - E(g_t^a))^3$</td>
<td>-1.329</td>
<td>-1.329</td>
<td>-1.380</td>
</tr>
<tr>
<td>$E(g_{d,t}^a)$</td>
<td>-1.932</td>
<td>1.788</td>
<td>0.077</td>
</tr>
<tr>
<td>$\text{Var}(g_{d,t}^a)$</td>
<td>-0.309</td>
<td>-0.097</td>
<td>-1.598</td>
</tr>
<tr>
<td>ACV1($g_{d,t}^a$)</td>
<td>1.073</td>
<td>1.388</td>
<td>-0.232</td>
</tr>
<tr>
<td>$E(g_{d,t}^a - E(g_{d,t}^a))^3$</td>
<td>0.367</td>
<td>0.367</td>
<td>0.367</td>
</tr>
<tr>
<td>$\text{CV}(g_t^a, g_{d,t}^a)$</td>
<td>0.336</td>
<td>0.140</td>
<td>-1.331</td>
</tr>
<tr>
<td>$E(pd_t^a)$</td>
<td>0.190</td>
<td>2.326</td>
<td>0.822</td>
</tr>
<tr>
<td>$\text{Var}(pd_t^a)$</td>
<td>-2.483</td>
<td>1.323</td>
<td>-1.329</td>
</tr>
<tr>
<td>ACV1($pd_t^a$)</td>
<td>-2.379</td>
<td>1.336</td>
<td>-1.183</td>
</tr>
<tr>
<td>$E(r_{e,t}^a)$</td>
<td>2.081</td>
<td>1.228</td>
<td>3.193</td>
</tr>
<tr>
<td>$\text{Var}(r_{e,t}^a)$</td>
<td>-1.760</td>
<td>-1.965</td>
<td>-2.140</td>
</tr>
<tr>
<td>ACV1($r_{e,t}^a$)</td>
<td>-1.359</td>
<td>-1.71</td>
<td>-2.008</td>
</tr>
<tr>
<td>$E(r_{m,t}^a)$</td>
<td>-2.280</td>
<td>-0.665</td>
<td>-1.309</td>
</tr>
<tr>
<td>$\text{Var}(r_{m,t}^a)$</td>
<td>-4.500</td>
<td>0.121</td>
<td>-4.854</td>
</tr>
<tr>
<td>ACV1($r_{m,t}^a$)</td>
<td>0.287</td>
<td>0.258</td>
<td>0.317</td>
</tr>
<tr>
<td>$\text{CV}(r_{e,t}^{ext}, pd_t^a)$</td>
<td>2.225</td>
<td>-0.090</td>
<td>2.019</td>
</tr>
<tr>
<td>$\text{CV}(r_{e,t}^{ext,1,5}, pd_t^a)$</td>
<td>2.315</td>
<td>-0.586</td>
<td>2.050</td>
</tr>
<tr>
<td>$\text{CV}(r_{e,t}^{ext,1,5}, pd_t^a)$</td>
<td>2.468</td>
<td>-1.120</td>
<td>2.132</td>
</tr>
</tbody>
</table>

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 9: Returns regressions on VIX and 3 Fama-French factors

<table>
<thead>
<tr>
<th>B/M portfolios deciles</th>
<th>Growth</th>
<th>Dec2</th>
<th>Dec3</th>
<th>Dec4</th>
<th>Dec5</th>
<th>Dec6</th>
<th>Dec7</th>
<th>Dec8</th>
<th>Dec9</th>
<th>Value</th>
<th>V-G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: $R_{i,t} - R_{f,t} = \alpha_i + \beta \Delta VIX_t + \varepsilon_{i,t}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_i$</td>
<td>0,52</td>
<td>0,55</td>
<td>0,60</td>
<td>0,50</td>
<td>0,50</td>
<td>0,57</td>
<td>0,35</td>
<td>0,44</td>
<td>0,67</td>
<td>0,63</td>
<td>0,11</td>
</tr>
<tr>
<td></td>
<td>(2,29)</td>
<td>(2,75)</td>
<td>(3,23)</td>
<td>(2,66)</td>
<td>(2,58)</td>
<td>(2,91)</td>
<td>(1,24)</td>
<td>(1,83)</td>
<td>(2,71)</td>
<td>(1,96)</td>
<td>(0,34)</td>
</tr>
<tr>
<td>$\beta (VIX)$</td>
<td>-0,67</td>
<td>-0,65</td>
<td>-0,63</td>
<td>-0,71</td>
<td>-0,68</td>
<td>-0,61</td>
<td>-0,73</td>
<td>-0,65</td>
<td>-0,72</td>
<td>-0,97</td>
<td>-0,30</td>
</tr>
<tr>
<td></td>
<td>(-10,74)</td>
<td>(-10,09)</td>
<td>(-8,42)</td>
<td>(-10,32)</td>
<td>(-9,56)</td>
<td>(-9,33)</td>
<td>(-12,84)</td>
<td>(-9,80)</td>
<td>(-8,65)</td>
<td>(-9,20)</td>
<td>(-4,34)</td>
</tr>
<tr>
<td>Adj.R$^2$</td>
<td>0,26</td>
<td>0,30</td>
<td>0,29</td>
<td>0,35</td>
<td>0,32</td>
<td>0,28</td>
<td>0,33</td>
<td>0,26</td>
<td>0,28</td>
<td>0,30</td>
<td>0,05</td>
</tr>
</tbody>
</table>

| Panel B: $R_{i,t} - R_{f,t} = \alpha_i + \beta_1 Rm-Rf_t + \beta_2 SMB_t + \beta_3 HLM_t + \beta_4 \Delta VIX_t + \varepsilon_{i,t}$ |        |      |      |      |      |      |      |      |      |       |     |
| $\alpha_i$             | 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)|
| $\beta (Rm-Rf)$        | 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)|
| $\beta (SMB)$          | -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)|
| $\beta (HML)$          | -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.
Figure 8: VR simulation of cash flow with and without correlation with stochastic volatility

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 from 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)
Figure 9: Loadings implied by the standard LRR model and our extended LRR model

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.)
Notes: This graph shows the 6 month holding period return on dividends contract with maturities of 12 month (blue line), 18 month (red line) and 24 month (green line). We can see that the return on the short term asset (12 month dividends contract) are higher compared to the return on the long-term asset (24 month).

10.2 Model solution when EIS≠1

The LRR model assumes a rational representative agent embedded with Epstein and Zin [1989] recursive utility function given by 45 who maximizes its continuation value subject to its inter-temporal budget constraint.

\[
V_t = \begin{cases} 
(1 - \delta)C_t^{1 - \psi} + \delta \left( E_t \left( V_{t+1}^{1 - \gamma} \right)^{\frac{1 - \psi}{1 - \gamma}} \right) \frac{1}{1 - \psi} & \text{if } \psi \neq 1, \gamma \neq 1 \\
C_t^{1 - \delta} \left( E_t \left( V_{t+1}^{1 - \gamma} \right)^{\frac{1}{1 - \delta}} \right) & \text{if } \psi = 1, \gamma \neq 1 \\
(1 - \delta)C_t^{1 - \psi} + \delta \exp \left( E_t \left( \log V_{t+1} \right) \right)^{1 - \psi} & \text{if } \psi \neq 1, \gamma = 1 \\
C_t^{1 - \delta} \exp \left( E_t \left( \log V_{t+1} \right) \right)^\delta & \text{if } \psi = 1, \gamma = 1 
\end{cases}
\] (45)
The SDF is given by:

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

\[ = \theta \log \delta - \frac{\theta}{\psi} \mu + x_t + \sigma_t \epsilon_{c,t+1} + \varphi_w \sigma_w \epsilon_{w,t+1} + (\theta - 1)r_{c,t+1} \]

\[ = \theta \log \delta - \frac{\theta}{\psi} \mu + (\theta - 1) \left[k_0 + (k_1 - 1)A_0 + \mu_c + k_1 A_2(1 - \nu)\sigma^2\right] + \left[-\frac{\theta}{\psi} + (\theta - 1) ((k_1 \rho - 1)A_1 + 1)\right]x_t \]

\[ + [(\theta - 1)(k_1 \nu - 1)A_2] \sigma^2 + [(\theta - 1) (k_1 A_2 \varphi_c)] \sigma_t \epsilon_{x,t+1} + \left[-\frac{\theta}{\psi} + (\theta - 1)\right] \sigma_t \epsilon_{c,t+1} \]

\[ + \left[-\frac{\theta}{\psi} \varphi_\sigma + (\theta - 1) (k_1 A_2 + \varphi_\sigma)\right] \sigma_w \epsilon_{w,t+1} \]

\[ = \theta \log \delta - \gamma \mu + (\theta - 1) \left[k_0 + (k_1 - 1)A_0 + \mu_c + k_1 A_2(1 - \nu)\sigma^2\right] - \frac{1}{\psi} x_t + (\theta - 1)(k_1 \nu - 1)A_2 \sigma^2 \]

\[ + [(\theta - 1)k_1 A_1 \varphi_c] \sigma_t \epsilon_{x,t+1} + [-\gamma] \sigma_t \epsilon_{c,t+1} + [-\gamma \varphi_\sigma + (\theta - 1)k_1 A_2] \sigma_w \epsilon_{w,t+1} \]

So,

\[ m_{t+1} = a_{0m} + a_{1m} x_t + a_{2m} \sigma^2 + \lambda_c \sigma_t \epsilon_{c,t+1} + \lambda_x \sigma_t \epsilon_{x,t+1} + \lambda_w \sigma_w \epsilon_{w,t+1} \quad (46) \]

Where

\[ a_{0m} = \log \delta - \frac{1}{\psi} \mu_c + \frac{1}{2} \theta (1 - \theta) \left[(1 - \frac{1}{\psi}) \varphi_\sigma + k_1 A_2\right] \sigma^2 \]

\[ a_{1m} = -\frac{1}{\psi} \]

\[ a_{2m} = (\theta - 1)(k_1 \nu - 1)A_2 \]

and the prices of risk are given by:

\[ -\lambda_c = \gamma \quad (47) \]

\[ -\lambda_x = (1 - \theta)k_1 A_1 \varphi_c \quad (48) \]

\[ -\lambda_w = \gamma \varphi_\sigma + (1 - \theta)k_1 A_2 \quad (49) \]

47 is the price of the short run consumption risk, 48 is the price of the long run consumption risk and 49 is the price of the volatility risk).

We have that:
$$m_{t+1} + r_{c,t+1} = \theta \log \delta + (1 - \gamma) \mu_c + \theta \left[ k_0 + (k_1 - 1)A_0 + k_1A_2(1 - \nu)\bar{\sigma}^2 \right] + \left[ 1 - \frac{1}{\psi} + (k_1\rho - 1)A_1 \right] x_t + \theta(k_1\nu - 1)A_2\sigma_d^2 + \left[ \theta k_1\varphi \right] \sigma_t \varepsilon_{x,t+1} + \left[ (1 - \gamma) \right] \sigma_t \varepsilon_{c,t+1} + \left[ (1 - \gamma)\varphi + \theta k_1A_2 \right] \sigma_w \varepsilon_{w,t+1}$$

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

$$A_0 = \frac{1 - \frac{1}{k_0}}{1 - k_1\rho} \left[ \log \delta + (1 - \frac{1}{\psi}) \mu_c + k_0 + k_1A_2(1 - \nu)\bar{\sigma}^2 + \frac{1}{2}\theta \left( (1 - \frac{1}{\psi})\varphi + k_1A_2 \right)^2 \sigma_w^2 \right]$$

$$A_1 = \frac{1 - \frac{1}{k_0}}{1 - k_1\rho}$$

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

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

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

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

$$pd_{t}^{(n)} = \log E_t \left[ \exp \left( m_{t+1} + \Delta d_{t+1} + pd_{t+1}^{(n-1)} \right) \right] = [a_{0m} + \mu_d + A_0(n - 1) + A_2(n - 1)(1 - \nu)\bar{\sigma}^2] + [a_{1m} + \phi + A_1(n - 1)\rho] x_t + [a_{2m} + A_2(n - 1)\nu + \frac{1}{2} (\lambda_c + \pi_c)^2 + \frac{1}{2} (\lambda_x + \varphi_c A_1(n - 1))^2 + \frac{1}{2} \varphi_d^2] \sigma_t^2 + [\lambda_w + \pi_\sigma + A_2(n - 1)]^2 \sigma_w^2$$

55
Implying that \( n \geq 1 \):

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

\[
A_1(n) = \phi + a_{1m} + A_1(n-1)\rho = \left(\phi - \frac{1}{\psi}\right) \left(\frac{1-\rho^n}{1-\rho}\right)
\]

\[
A_2(n) = a_{2m} + A_2(n-1)\nu + \frac{1}{2} (\lambda_c + \pi_c)^2 + \frac{1}{2} (\lambda_x + \varphi_c A_1(n-1))^2 + \frac{1}{2} \varphi_d^2
\]

Which is equivalent to:

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

\[
A_1(n) = \left(\phi - \frac{1}{\psi}\right) \left(\frac{1-\rho^n}{1-\rho}\right)
\]  \( (50) \)

\[
A_2(n) = A_2(n-1)\nu + (\theta - 1) (k_1 \nu - 1) A_2 + \frac{1}{2} (\pi_c - \gamma)^2 + \frac{1}{2} (\pi_c - \gamma)^2
\]

For the zero coupon bond,

\[
P_1^{(n)} = \exp \left( B_0(n) + B_1(n)x_t + B_2(n)\sigma_t^2 \right)
\]

where

\[
B_0(n) = B_0(n-1) + a_{0m} + B_2(n-1)(1 - \nu)\bar{\sigma}^2 + \frac{1}{2} [\lambda_w + B_2(n-1)]^2 \sigma_w^2
\]

\[
B_1(n) = -\frac{1}{\psi} \left(\frac{1-\rho^n}{1-\rho}\right)
\]  \( (51) \)

\[
B_2(n) = B_2(n-1)\nu + (\theta - 1)(k_1 \nu - 1) A_2 + \frac{1}{2} \lambda_c^2 + \frac{1}{2} (\lambda_x + B_1(n-1)\varphi_c)^2
\]

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

The return on the risk-free asset is given by:

\( A_0(0) = 0, A_1(0) = 0, A_2(0) = 0 \)
The return on the n-period dividend strip is given by:

\[
r_{f,t} = -E_t(m_{t+1}) - \frac{1}{2} \text{Var}_t(m_{t+1}) = -a_{0m} - \frac{1}{2} \lambda_w^2 \sigma_w^2 - a_{1m} x_t - \left( a_{2m} + \frac{1}{2} (\lambda_e^2 + \lambda_x^2) \right) \sigma_t^2
\]

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

\[
r_{d,t+1}^{(n)} = \Delta d_{t+1} + pd_{t+1}^{(n-1)} - pd_t^{(n)} = \left[ \mu_d + A_0(n-1) - A_0(n) + A_2(n-1)(1-\nu) \hat{\sigma}^2 \right] + (\phi + A_1(n-1) \rho - A_1(n)) x_t + (A_2(n-1) \nu - A_2(n)) \sigma_t^2 + (A_1(n-1) \varphi_e) \sigma_t \epsilon_{x,t+1} + \pi_c \sigma_t \epsilon_{c,t+1}
\]

So, the excess return is given by:

\[
r_{d,t+1}^{(n)} - r_{f,t} = \left[ a_{0m} + \frac{1}{2} \lambda_w^2 \sigma_w^2 + \mu_d + (A_0(n-1) - A_0(n)) + A_2(n-1)(1-\nu) \hat{\sigma}^2 \right] + \left( \phi - \frac{1}{\psi} + A_1(n-1) \rho - A_1(n) \right) x_t + \left( A_2(n-1) \nu - A_2(n) + a_{2m} + \frac{1}{2} (\lambda_e^2 + \lambda_x^2) \right) \sigma_t^2 + (A_1(n-1) \varphi_e) \sigma_t \epsilon_{x,t+1} + \pi_c \sigma_t \epsilon_{c,t+1}
\]

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

\[
r_{p}^{(n)} = E_t(r_{d,t+1}^{(n)} - r_{f,t}) + \frac{1}{2} \text{Var}_t(r_{d,t+1}^{(n)} - r_{f,t})
\]

\[
= -\lambda_w (\pi_\sigma + A_2(n-1)) \sigma_w^2 + \left[ -\lambda_e (\pi_e) - \lambda_x (A_1(n-1) \varphi_e) \right] \sigma_t^2
\]
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:

\[
\begin{align*}
\rho p_t^{(1)} & = E_t(r_{d,t+1}^{(1)} - r_{f,t}) + \frac{1}{2} \text{Var}_t(r_{d,t+1}^{(1)} - r_{f,t}) \\
& = -\lambda w \pi \sigma^2_w + [-\lambda c \pi c] \sigma^2_t \tag{54}
\end{align*}
\]

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:

\[
S^{(1)}_{n,t} = \rho p_t^{(n)} - \rho p_t^{(n-1)} = -\lambda w ([A_2(n-1) - A_2(n-2)]) \sigma^2_w - \lambda x [A_1(n-1) - A_1(n-2)] \varphi \sigma^2_t
\]

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

\[
F = \left[ \left( \phi - \frac{1}{\psi} \right) + (\theta - 1) k_1 A_1 \right] \varphi e^2 \\
G = -\frac{1}{2} \left( \frac{1 + \rho}{1 - \rho} \right) \left( \phi - \frac{1}{\psi} \right) \varphi e^2
\]
Therefore

\[ S^{(1)}_{n,t} = -\lambda_w \left( A_2(1) \nu^{n-2} + \left( \phi - \frac{1}{\psi} \right) \left( \frac{\nu^{n-2} - \rho^{n-2}}{\nu - \rho} \right) F + \left( \frac{\nu^{n-2} - \rho^{2(n-2)}}{\nu - \rho^2} \right) G \right) \sigma_w^2 \]  

(55)

- \left( \lambda_x \varphi \right) \left( \phi - \frac{1}{\psi} \right) \rho^{n-2} \sigma_t^2

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

\[ S^{(n)}_{t} = r_{p_t^{(n)}} - r_{p_t^{(1)}} = \sum_{i=2}^{n} S^{(i)}_{i,t} \]  

(56)

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

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

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

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

\[ S^{(n)}_{t} = r_{p_t^{(n)}} - r_{p_t^{(1)}} = -\lambda_w \left( A_2(n-1) \right) \sigma_w^2 + \left[ -\lambda_x A_1(n-1) \varphi \right] \sigma_t^2 \]  

(57)

\[ S_t = \lim_{n \to \infty} S^{(n)}_{t} = -\lambda_w \left( A_2(\infty) \right) \sigma_w^2 + \left[ -\lambda_x A_1(\infty) \varphi \right] \sigma_t^2 \]  

(58)

Where

\[ A_1(\infty) = \left( \frac{\phi - \frac{1}{\psi}}{1 - \rho} \right) \]  

(59)

\[ A_2(\infty) = \frac{1}{(1 - \nu)} \left[ (\theta - 1)(k_1 \nu - 1)A_2 + \frac{1}{2} (\pi_e - \gamma)^2 + \frac{1}{2} ((\theta - 1)k_1 A_1 + A_1(\infty)) \varphi \right] + \frac{1}{\rho^2} \]  

(60)

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:
\[ \frac{P_t}{D_t} = \sum_{n=1}^{\infty} \frac{P_t^{(n)}}{D_t} = \sum_{n=1}^{\infty} \exp\left(A_0(n) + A_1(n)x_t + A_2(n)\sigma_t^2\right) \tag{61} \]

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:

\[ R_m^{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{(P_{t+1}/D_{t+1}) + 1}{P_t/D_t} D_{t+1} \tag{62} \]

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 \( \sigma_t^2 \), 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

\[ |\rho| < 1 \tag{63} \]

Assuming that 63 holds, \( A_2(n) \) converges if and only if

\[ |\nu| < 1 \tag{64} \]

Given 63 and 64, \( \lim_{n \to \infty} A_1(n) = A_1(\infty) \) and \( \lim_{n \to \infty} A_2(n) = A_2(\infty) \).

Let denote

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

It follows from the recursion in 50 for \( A_0(n) \) that for \( n > N \) with \( N \) sufficiently large,

\[ A_0(n) \approx nA_0(\infty) + \text{constant} \]

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

Therefore the necessary and sufficient conditions for convergence of the aggregate price dividend ratio are 63, 64 and \( A_0(\infty) < 0 \).
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 $\tau$ to the variance of one period cash flow growth:

$$ VR_{1,t}(\tau) = \frac{\frac{1}{\tau} \text{Var}_t \left[ \log \left( \frac{D_{t+\tau}}{D_t} \right) \right]}{\text{Var}_t \left[ \log \left( \frac{D_{t+1}}{D_t} \right) \right]} $$

(65)

Under the dynamics assumed in 2-5, given the conditional normality, $36$ is the conditional variance at time $t$ of the sum of dividend growth over the horizon $\tau$ (from $t+1$ to $t+\tau$) divided by $\tau$. The first definition of the variance ratio provides the following analytical formula for the dividend:

$$ VR_{1,t}(\tau) = \frac{1}{\tau} \left\{ \left( \pi_c^2 + \phi_d^2 \right) \left( \frac{1-\nu^\tau}{1-\nu} \right) + \left( \frac{\phi_c \phi_d}{\tau} \right)^2 \left( \left( \frac{1-\nu^{\tau-1}}{1-\nu} \right) - 2\rho \left( \frac{\nu^{\tau-1}-\nu^{-1}}{\rho-\nu} \right) + \rho^2 \left( \frac{\nu^{2(\tau-1)}-\nu^{-1}}{\rho^2-\nu^2} \right) \right) \right\} \sigma_t^2 $$

$$ + \frac{1}{\tau} \left\{ \left( \frac{\phi_c \phi_d}{\tau} \right)^2 \left( \tau - 1 - \frac{1-\nu^{\tau-1}}{\rho-\nu} \right) + \left( \pi_c^2 + \phi_d^2 \right) \left( \tau - \frac{1-\nu^\tau}{1-\nu} \right) \right\} \bar{\sigma}_2^2 + \frac{\pi_c^2 \sigma_w^2}{\left( \pi_c^2 + \phi_d^2 \right) \sigma_t^2 + \pi_c^2 \bar{\sigma}_2^2} $$

(66)

Interestingly we can see that at the steady state ($\sigma_t^2 = \bar{\sigma}_2^2$), VR1 equals $1$ plus a positive term$^{31}$ 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 $35$ and $37$ are computed using the following intermediary calculus:

$^{31}$This term is given by: $\frac{1}{\tau} \left( \frac{\phi_c \phi_d}{\tau} \right)^2 \left( \tau - 1 - 2\rho \left( \frac{\nu^{\tau-1}-\nu^{-1}}{\rho-\nu} \right) + \rho^2 \left( \frac{\nu^{2(\tau-1)}-\nu^{-1}}{\rho^2-\nu^2} \right) \right) \bar{\sigma}_2^2 / \left( \pi_c^2 + \phi_d^2 \right) \sigma_t^2 + \pi_c^2 \bar{\sigma}_2^2$
\[
\log E_t \left[ D(\tau)/D(0) \right] = \log E_t \left[ \exp \left( \sum_{i=1}^{\tau} \Delta d_{t+i} \right) \right]
\]

\[
= \log E_t \exp \sum_{i=1}^{\tau} \left( \mu_d + \phi x_{t+i-1} + \pi_c s_{t+i} + \pi_x v_{t+i} + \pi_\sigma w_{t+i} + \varphi_d u_{d,t+i} \right)
\]

\[
= \log E_t \exp \sum_{i=1}^{\tau} \left( \mu_d + \phi x_{t+i-1} + \sum_{r=0}^{i-2} \left[ \rho^r \varphi_c v_{t+i-r-1} \right] + \pi_c s_{t+i} + \pi_x v_{t+i} + \pi_\sigma w_{t+i} + \varphi_d u_{d,t+i} \right)
\]

\[
= \log E_t \exp \left( \tau \mu_d + \phi \left[ \frac{1 - \rho^\tau}{1 - \rho} \right] x_t + \sum_{i=1}^{\tau} \sum_{i=1}^{\tau} \left[ \rho^r \varphi_c v_{t+i-r-1} \right] + \pi_c s_{t+i} + \pi_x v_{t+i} + \pi_\sigma w_{t+i} + \varphi_d u_{d,t+i} \right)
\]

\[
= \tau \mu_d + \phi \left[ \frac{1 - \rho^\tau}{1 - \rho} \right] x_t + \log E_t \exp \left( \sum_{i=1}^{\tau} \pi_c s_{t+i} + \sum_{i=1}^{\tau} \left( \frac{1 - \rho^{-i}}{1 - \rho} \right) \varphi_c v_{t+i} + \sum_{i=1}^{\tau} \pi_\sigma w_{t+i} + \sum_{i=1}^{\tau} \varphi_d u_{d,t+i} \right)
\]

\[
= \tau \mu_d + \phi \left[ \frac{1 - \rho^\tau}{1 - \rho} \right] x_t + \log E_t \exp \left( \sum_{i=1}^{\tau} \left[ \pi_c s_{t+i} + \left( \frac{1 - \rho^{-i}}{1 - \rho} \right) \varphi_c v_{t+i} \right] \right)
\]

\[
+ \pi_\sigma w_{t+i} + \varphi_d u_{d,t+i} \times E_{t+\tau} \left[ \pi_c s_{t+i} + \pi_\sigma w_{t+i} + \varphi_d u_{d,t+i} \right] \exp \left[ \frac{1}{2} \nu H_0 \sigma_\nu^2 \right]
\]

\[
= \tau \mu_d + \phi \left[ \frac{1 - \rho^\tau}{1 - \rho} \right] x_t + \log E_t \exp \left( \sum_{i=1}^{\tau} \left[ \pi_c s_{t+i} + \left( \frac{1 - \rho^{-i}}{1 - \rho} \right) \varphi_c v_{t+i} \right] \right)
\]

\[
+ \pi_\sigma w_{t+i} + \varphi_d u_{d,t+i} \times E_{t+\tau} \exp \left[ \frac{1}{2} H_0 \sigma^2 \right]
\]

\[
\text{The same way, we obtain that:}
\]

\[
\log E_t \left[ D^2(\tau)/D^2(0) \right] = 2\tau \mu_d + 2\phi \left[ \frac{1 - \rho^\tau}{1 - \rho} \right] x_t + 2H_{\tau-1} \sigma_\tau^2 + \frac{1}{2} G_{\tau-1} \sigma_\nu^2 + \frac{1}{2} (1 - \nu) \left[ \sum_{i=0}^{\tau-2} H_i \right] \sigma^2
\]
Where $H_0 = \varphi^2_d + \pi^2_c + \pi^2_v$, $G_0 = G_0' = \pi^2_v$ and $\forall j \in \{1, 2, \ldots\}$,

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

### 10.5 Term structure of interest rate

#### 10.5.1 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 ($\pi_t$) and to compute the nominal log stochastic discount factor ($m^S_{t+1}$) as follows:

$$
m^S_{t+1} = m_{t+1} - \pi_{t+1}
$$

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.

$$
\begin{align*}
\pi_{t+1} &= \mu\pi_t + z_{t+1} + \nu_{\pi}\sigma_t \varepsilon_{c,t+1} + \pi_{\sigma}\sigma_{w,t+1} + \nu_t \varepsilon_{\pi,t+1} \\
z_{t+1} &= \phi_z z_t + \nu_z (\nu_{\sigma}\sigma_t \varepsilon_{c,t+1} + \nu_t \varepsilon_{\pi,t+1}) \\
v_{t+1}^2 &= (1 - \phi_v) \bar{v} + \nu_v \varepsilon_{v,t+1} \\
(\varepsilon_{\pi,t+1}, \varepsilon_{w,t+1}, \varepsilon_{c,t+1}, \varepsilon_{v,t+1}) &\sim N.i.id(0, I)
\end{align*}
$$

Where $\nu_{\pi}$ captures the effect of consumption growth shock on inflation, while $\pi_{\sigma}$ 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:

$$
P^S_{n,t} = \exp \left(B^S_{0}(n) + B^S_{x}(n)x_t + B^S_{z}(n)z_t + B^S_{\sigma}(n)\sigma_t^2 + B^S_{v}(n)v_t^2\right)
$$

and using the bond pricing equation 27, we can deduce the coefficient of the affine function as
follows:

\[ B_s^x(n) = a_{1m} + \rho B_s^x(n - 1) \]
\[ B_s^z(n) = -1 + \phi_z B_s^z(n - 1) \]
\[ B_s^\sigma(n) = a_{2m} + \nu B_s^\sigma(n - 1) + \frac{1}{2} \left( \lambda_c - \nu \pi + \nu_\pi B_s^\sigma(n - 1) \right)^2 + \frac{1}{2} \left( \lambda_\sigma + \varphi_\sigma B_s^\sigma(n - 1) \right)^2 \]
\[ B_s^\nu(n) = \phi_\nu B_s^\nu(n - 1) + \frac{1}{2} \left( -1 + \nu B_s^\nu(n - 1) \right)^2 \]
\[ B_s^\sigma_0(n) = a_{0m} - \mu_\pi + B_s^\sigma_0(n - 1) + (1 - \nu) B_s^\sigma_0(n - 1) \sigma^2 + (1 - \phi_\nu) B_s^\nu(n - 1) \bar{v} + \frac{1}{2} \left( \nu B_s^\nu(n - 1) \right)^2 + \frac{1}{2} \left( \lambda_\sigma - \pi \pi + B_s^\sigma_0(n - 1) \right)^2 \sigma_w^2 \]

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 (70) only shows that there is enough flexibility to match the nominal bond term structure without compromising the other achievements of the model.

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:
\[ r^{(n)}_{t+6} = \log \left( \frac{F_{t+6}^{(n-6)} + \sum_{i=1}^{6} D_{t+i}}{F_t^{(n)}} \right) \]

\[ = \log \left( 1 + \frac{F_{t+6}^{(n-6)}}{\sum_{i=0}^{5} D_{t+6-i}} \right) - \log \left( \frac{F_t^{(n)}}{\sum_{i=0}^{5} D_{t-i}} \right) + \log \left( \frac{\sum_{i=0}^{5} D_{t+6-i}}{\sum_{i=0}^{5} D_{t-i}} \right) \]

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

\[
\log \left( \frac{F_t^{(n)}}{D_{0.5y}^{(n)}} \right) = \log \left( \sum_{i=1}^{n} \frac{p_{i}}{\sum_{i=0}^{5} D_{t-i}} \right) \approx \log \left[ \sum_{i=1}^{n} \exp \left( pd_{t}^{(i)} \right) - \log 6 - \sum_{k=0}^{5} \left( k + 1 \right) \right] \]

(72)

Where \( pd_{t}^{(i)} \) is the log-price dividend ratio of the \( i \)-period dividend strip\(^{32}\).

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

\[
r^{(n)}_{6,t+6} = \log D_{t+6}^{0.5y} - \log D_{t}^{0.5y} + \log \left[ 1 + \exp \log \left( \frac{F_{t+6}^{(n-6)}}{D_{t+6}^{0.5y}} \right) \right] - \log \left( \frac{F_t^{(n)}}{D_{t}^{0.5y}} \right) \]

(73)

There is no closed form solution for \( r^{(n)}_{6,t+6} \) even though there are closed form for \( \Delta d_{12t} \) 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.

### 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 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

\(^{32}\)Equation (107) in 10.8 is used for the proof.
(meaning that investor would like to hedge themselves again 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 (84) 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 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 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 lifetime utility given by:

$$\mathbb{E} \sum_{t=0}^{\infty} \delta^t \frac{(C_t - X_t)^{1-\gamma} - 1}{1 - \gamma} \tag{74}$$

Where $X_t$ denotes the habit level which is defined 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 ($s_t = \log(S_t)$):

$$\Delta c_{t+1} = \mu_c + \sigma_c \varepsilon_{c,t+1} + \varphi \sigma_w \varepsilon_{w,t+1} \tag{75}$$

$$\sigma_{c,t+1}^2 = \nu \sigma_{c,t}^2 + (1 - \nu) \bar{\sigma}^2 + \sigma_w \varepsilon_{w,t+1} \tag{76}$$

$$s_{t+1} = (1 - \phi) \bar{s} + \phi s_t + \lambda(s_t)(\Delta c_{t+1} - \mathbb{E}_t \Delta c_{t+1}) \tag{77}$$

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 assume that

---

33This self 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.
\( \lambda(s_t) = \lambda \) is a constant\(^{34}\). The log-pricing kernel is given by:

\[
m_{t+1} = \log(\delta) - \gamma \mu_c + \gamma (1 - \phi)(s_t - \bar{s}) - \gamma (1 + \lambda) \sigma_c \varepsilon_{c,t+1} - \gamma (1 + \lambda) \varphi_\sigma \sigma_w \varepsilon_{w,t+1}
\]

\[
= \log(\delta) - \gamma \mu_c + \gamma (1 - \phi)(s_t - \bar{s}) - \lambda_c \sigma_c \varepsilon_{c,t+1} - \lambda_w \sigma_w \varepsilon_{w,t+1}
\]

(78)

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:

\[
r_{f,t} = -E_t(m_{t+1}) - \text{Var}_t(m_{t+1})
\]

\[
= \log \delta + \gamma \mu_c - \gamma (1 - \phi)(s_t - \bar{s}) - \frac{1}{2} \gamma^2 (1 + \lambda)^2 \sigma_t^2 - \frac{1}{2} \gamma^2 (1 + \lambda)^2 \varphi_\sigma^2 \sigma_w^2
\]

(79)

The dividend growth process is defined by:

\[
\Delta d_{t+1} = \mu_d + \pi_c \sigma_c \varepsilon_{c,t+1} + \varphi_d \sigma_d \varepsilon_{d,t+1} + \pi_\sigma \sigma_w \varepsilon_{w,t+1}
\]

(80)

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

\[
pd_{t}^{(n)} = A_0(n) + A_1(n)(s_t - \bar{s}) + A_2(n)\sigma_t^2
\]

(81)

The coefficient of the affine function can be computed recursively thanks 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

\[
pd_{t}^{(n)} = \log E_t \left( \exp \left( m_{t+1} + \Delta d_{t+1} + pd_{t+1}^{(n-1)} \right) \right),
\]

implying that:

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

\[
A_1(n) = A_1(n - 1)\phi + \gamma (1 - \phi)
\]

\[
A_2(n) = \nu A_2(n - 1) + \frac{1}{2} \left( \varphi_\sigma^2 + (\pi_\sigma + A_1(n - 1)\lambda - \gamma (1 + \lambda))^2 \right)
\]

(82)

\(^{34}\)In order to obtain the risk free rate as a linear function of the log-surplus, we can define the sensitivity function as follow: \(\lambda(s_t, \sigma_t^2) = \frac{1}{\sigma_t^2} \sqrt{1 - 2(s_t - \bar{s}) - 1}\) where \(S_t = \sqrt{(\sigma_t^2 + \varphi_\sigma^2 \sigma_w^2) \frac{1}{1 - \phi - \pi}}\) and to be sure that \(\lambda(s_t, \sigma_t^2)\) is always defined, we set \(\lambda(s_t, \sigma_t^2) = 0\) when \(s_t > s_{t,max} = \bar{s} + \frac{1}{\lambda}(1 - S_t^2)\)
The log-return on the $n$-period dividend strip is given by $r_{d,t+1}^{(n)} = \Delta d_{t+1} + p d_{t+1}^{(n-1)} - p d_{t}^{(n)}$, from which we derive its risk premium as follows:

$$rp_t^{(n)} = E_t(r_{d,t+1}^{(n)} - r_{f,t}) + \frac{1}{2} \text{Var}_t(r_{d,t+1}^{(n)} - r_{f,t}) = \lambda w (\pi_\sigma + A_2(n - 1) + A_1(n - 1)\lambda \varphi_\sigma) \sigma_w^2 + \lambda c (\pi_c + A_1(n - 1)\lambda) \sigma_t^2$$

(83)

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:

$$S_{n,t}^{(1)} = r_{p_t}^{(n)} - r_{p_t}^{(n-1)}$$

$$= -\lambda w ([A_2(n - 1) - A_2(n - 2)] + [A_1(n - 1) - A_1(n - 2)] \lambda \varphi_c) \sigma_w^2 - \lambda c ([A_1(n - 1) - A_1(n - 2)] \lambda) \sigma_t^2$$

(84)

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 (84) 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.

10.8 Temporal aggregation and moment conditions

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

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

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 $t$

$$\tau_j = \frac{j + 1}{h} \quad \text{if} \quad j < h \quad \text{and} \quad \tau_j = \frac{2h - j - 1}{h} \quad \text{if} \quad j \geq h$$
\[
\Delta c_{t+1}^a = \log \frac{\sum_{j=0}^{h-1} C_{h(t+1)-j}}{\sum_{j=0}^{h-1} C_{ht-j}} \approx \sum_{j=0}^{2h-2} \tau_j \Delta c_{h(t+1)-j} \\
= h \mu_c + b_{h-1} x_{h(t-1)} + \sum_{j=0}^{h-1} a_j \varphi_c \sigma_{h(t+1)-j} - 2 \varepsilon_x, h(t+1) - j - 1 + \sum_{j=0}^{h-2} b_j \varphi_c \sigma_{ht-j} - 2 \varepsilon_{x,ht-j-1} \\
+ \sum_{j=0}^{2(h-1)} \tau_j \sigma_{h(t+1)-j} - 1 \varepsilon_c, h(t+1) - j + \sum_{j=0}^{2(h-1)} \tau_j \varphi_x \sigma_{h(t+1)-j} - 1 \varepsilon_{x, h(t+1) - j} \\
+ \sum_{j=0}^{2(h-1)} \tau_j \varphi \sigma w \varepsilon_{w, h(t+1) - j}
\]

Where

\[
a_j = \sum_{k=0}^{j} \left( \frac{k+1}{h} \right) \rho^{j-k} = \frac{\rho^{j+2} - (j+2) \rho + j + 1}{h(1-\rho)^2} \quad \text{if } j < h
\]

\[
b_j = \sum_{k=0}^{j+h} \tau_k \rho^{j+h-k} = \frac{\rho^{j+h+2} - 2 \rho^{j+1} + (h-2-j)(1-\rho) + 1}{h(1-\rho)^2} \quad \text{if } j \leq h - 1
\]

Let us denote \( \nu_{h(t+1) - j} = \sigma_{h(t+1) - j} - 1 \varepsilon_{x, h(t+1) - j} \), \( \sigma_{h(t+1) - j} = \sigma_{h(t+1) - j} - 1 \varepsilon_c, h(t+1) - j \) and \( w_{h(t+1) - j} = \sigma_{w \varepsilon_{w, h(t+1) - j}} \).

Then,

\[
\Delta c_{t+1}^a = h \mu_c + b_{h-1} x_{h(t-1)} + \sum_{j=0}^{h} \hat{a}_j \nu_{h(t+1) - j} + \sum_{j=0}^{h-2} \hat{b}_j \nu_{ht-j-1} + \sum_{j=0}^{h} c_j \sigma_{h(t+1) - j} + \sum_{j=0}^{h-2} d_j \sigma_{ht-j-1} \quad (85)
\]

\[
+ \sum_{j=0}^{h} \tilde{f}_j w_{h(t+1) - j} + \sum_{j=0}^{h-2} \tilde{g}_j w_{ht-j-1}
\]

Where

\[
\begin{align*}
\hat{a}_j &= \varphi_c a_{j-1} & e_j &= \tau_j & f_j &= \varphi_c \tau_j & 0 \leq j \leq h \\
\hat{b}_j &= \varphi_c b_j & d_j &= \tau_{h+j+1} & g_j &= \varphi_c \tau_{h+j+1} & 0 \leq j \leq h - 2
\end{align*}
\]

Therefore,

\[
E(\Delta c_{t+1}^a) = E(\sum_{j=0}^{2(h-1)} \tau_j \mu_c) = h \mu_c \quad (86)
\]
\[
\text{Var}(\Delta e_{t+1}^a) = b_{h-1}^2 \text{Var} \left( x_{h(t-1)} \right) + \left( \sum_{j=0}^{h} a_j^2 + \sum_{j=0}^{h-2} b_j^2 \right) \psi^2 \sigma^2 + \left( \sum_{j=0}^{h} c_j^2 + \sum_{j=0}^{h-2} d_j^2 \right) \tilde{\sigma}^2
\]

\[
\text{Cov}(\Delta e_t^a, \Delta e_{t+1}^a) = \text{Cov} \left( h \mu_c + b_{h-1} x_{h(t-2)} + \sum_{j=0}^{h} \bar{a}_j v_{ht-j} + \sum_{j=0}^{h-2} \bar{b}_j v_{h(t-1)-j-1} + \sum_{j=0}^{h} c_j s_{ht-j} \right)
\]

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

\[
+ \sum_{j=0}^{h} \bar{a}_j v_{h(t+1)-j} + \sum_{j=0}^{h-2} \bar{b}_j v_{h(t+1)-j-1} + \sum_{j=0}^{h} c_j s_{h(t+1)-j} + \sum_{j=0}^{h-2} d_j s_{ht-j-1}
\]

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

\[
= \rho^h b_{h-1}^2 \text{Var} \left( x_{h(t-2)} \right) + \text{Cov} \left( \sum_{j=0}^{h} \bar{a}_j v_{ht-j} + \sum_{j=0}^{h-2} \bar{b}_j v_{h(t-1)-j-1} + \sum_{j=0}^{h} c_j s_{ht-j} \right)
\]

\[
+ \sum_{j=0}^{h-2} d_j s_{h(t-1)-j-1} + \sum_{j=0}^{h} f_j w_{ht-j} + \sum_{j=0}^{h-2} g_j w_{h(t-1)-j-1}, \sum_{j=0}^{h} \bar{a}_j v_{h(t+1)-j}
\]

\[
+ b_{h-1} \left( \sum_{k=0}^{h-1} \rho^k (\varphi v_{h(t-1)-k}) \right) + \sum_{j=0}^{h-2} \bar{b}_j v_{ht-j-1}
\]

\[
+ \sum_{j=0}^{h} c_j s_{h(t+1)-j} + \sum_{j=0}^{h-2} d_j s_{ht-j-1} + \sum_{j=0}^{h} f_j w_{h(t+1)-j} + \sum_{j=0}^{h-2} g_j w_{ht-j-1}
\]

\[
\text{Cov}(\Delta e_t^a, \Delta e_{t+1}^a) = \rho^h b_{h-2}^2 \text{Var} \left( x_{h(t-2)} \right) + \sum_{j=0}^{h-2} (\bar{a}_j + \rho b_{h-2} \rho^j \psi \varphi) \bar{b}_j \tilde{\sigma}^2 + \sum_{j=0}^{h-2} (c_j) d_j \sigma^2 + \sum_{j=0}^{h-2} (f_j + g_j) \sigma^2
\]

\[
+ \bar{a}_h (\bar{a}_0 + b_{h-1} \psi) \tilde{\sigma}^2 + c_h (c_0) \tilde{\sigma}^2 + f_h (f_0) \sigma^2
\]

Where \( \text{Var}(x_t) = \psi^2 \sigma^2 / (1 + \rho^2) \)
\[ E \left[ (\Delta a_{t+1}^2 - E(\Delta a_{t+1}^2))^3 \right] = 3E \left[ \left( \sum_{j=0}^{h} a_j v_{ht(t+1)-j} + \sum_{j=0}^{h-2} b_j v_{ht-j-1} \right)^2 + \left( \sum_{j=0}^{h} c_j s_{ht(t+1)-j} + \sum_{j=0}^{h-2} d_j s_{ht-j-1} \right)^2 \right] \times \left( \sum_{j=0}^{h-1} f_j w_{ht(t+1)-j} + \sum_{j=0}^{h-2} g_j w_{ht-j-1} \right) \]

\[ = 3 \left( \sum_{j=0}^{h-1} a_j^2 + \sum_{j=0}^{h-2} c_j^2 \right) f_k \nu^{k-j+1} + \sum_{j=0}^{h} \sum_{k=0}^{h-2} (a_j^2 + c_j^2) g_k \nu^{k-j+1} + \sum_{j=0}^{h-2} \sum_{k=0}^{h-2} (d_j^2 + d_k^2) g_k \nu^{k-j+1} \right) \sigma_w^2 \]

- **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:

\[ \Delta d_{t+1}^2 = \log \frac{\sum_{j=0}^{h-1} D_{h(t+1)-j}}{\sum_{j=0}^{h-1} D_{ht-j}} \approx 2 \sum_{j=0}^{2(h-1)} \tau_j \Delta d_{h(t+1)-j} \]

\[ = h\mu_d + \phi b_{ht-1} x_{ht(t-1)} + \phi \sum_{j=0}^{h-1} a_j \varphi_c x_{ht(t+1)-j-1} + \phi \sum_{j=0}^{h-2} b_j \varphi_c x_{ht-j-1} + \pi_c \sum_{j=0}^{2(h-1)} \tau_j s_{ht(t+1)-j} \]

\[ + \pi_x \sum_{j=0}^{2(h-1)} \tau_j t_{ht(t+1)-j} + \pi_x \sum_{j=0}^{2(h-1)} \tau_j u_{ht(t+1)-j} + \tau_j \varphi d_{ht(t+1)-j} \]

\[ \Delta d_{t+1}^2 = h\mu_d + \phi b_{ht-1} x_{ht(t-1)} + \sum_{j=0}^{h-2} [\phi a_j \varphi_c v_{ht(t+1)-j}] \]

\[ + \sum_{j=0}^{h-2} [\phi b_j \varphi_c v_{ht-j-1}] + \sum_{j=0}^{h} [\pi c \varphi x_{ht(t+1)-j}] + \sum_{j=0}^{h-2} [\pi c \varphi s_{ht-j-1}] \]

\[ + \sum_{j=0}^{h-2} [\pi x \varphi t_{ht-j-1}] + \sum_{j=0}^{h-2} [\pi x \varphi u_{ht-j-1}] + \sum_{j=0}^{2(h-1)} \tau_j \varphi d_{ht(t+1)-j} \]

Therefore,

\[ E(\Delta d_{t+1}^2) = h\mu_d \]
\begin{align*}
\text{Var}(\Delta d_{t+1}^a) &= [\phi b_{h-1}]^2 \text{Var}(x_{h(t-1)}) + \sum_{j=0}^{h} [\phi a_{j-1} \phi c_e]^2 \sigma^2 + \sum_{j=0}^{h} [\pi_e \tau_j]^2 \sigma^2 \\
&\quad + \sum_{j=0}^{h} [\pi_d \tau_j]^2 \sigma_w^2 + \sum_{j=0}^{h-2} [\phi b_j \phi c_e]^2 \sigma^2 + \sum_{j=0}^{h-2} [\pi_e \tau_{h+j+1}]^2 \sigma^2 + \sum_{j=0}^{h-2} [\pi_d \tau_{h+j+1}]^2 \sigma_w^2 \\
&\quad + \left( \sum_{j=0}^{2(h-1)} \tau_j^2 \right) \psi_d^2 \sigma^2 \\
\text{Cov}(\Delta d_t^a, \Delta d_{t+1}^a) &= \text{Cov} \left( h_{\mu_d} + \phi b_{h-1} x_{h(t-2)} + \sum_{j=0}^{h} [\phi a_{j-1} \phi c_e] v_{ht-j} \\
&\quad + \sum_{j=0}^{h-2} [\phi b_j \phi c_e] v_{h(t-1)-j-1} + \sum_{j=0}^{h} [\pi_e \tau_j] s_{ht-j} + \sum_{j=0}^{h-2} [\pi_e \tau_{h+j+1}] s_{ht(1)-j-1} \\
&\quad + \sum_{j=0}^{h} [\pi_e \tau_j] w_{ht-j} + \sum_{j=0}^{h-2} [\pi_e \tau_{h+j+1}] w_{h(t-1)-j-1} + \sum_{j=0}^{2(h-1)} \tau_j \phi d_{ht-j}, h_{\mu_d} \\
&\quad + \phi b_{h-1} x_{h(t-1)} + \sum_{j=0}^{h} [\phi a_{j-1} \phi c_e] v_{ht(1)-j} + \sum_{j=0}^{h-2} [\phi b_j \phi c_e] v_{ht-j-1} \\
&\quad + \sum_{j=0}^{h} [\pi_e \tau_j] s_{ht(1)-j} + \sum_{j=0}^{h-2} [\pi_e \tau_{h+j+1}] s_{ht-j-1} + \sum_{j=0}^{h} [\pi_d \tau_j] w_{h(t+1)-j} \\
&\quad + \sum_{j=0}^{h-2} [\pi_d \tau_{h+j+1}] w_{ht-j-1} + \sum_{j=0}^{2(h-1)} \tau_j \phi d_{ht(t+1)-j} \right) \\
\text{Cov}(\Delta d_t^a, \Delta d_{t+1}^a) &= \rho^h [\phi b_{h-1}]^2 \text{Var}(x_{h(t-2)}) + \sum_{j=0}^{h} [\phi a_{j-1} \phi c_e] [\phi b_{j-1} \phi c_e] \sigma^2 \\
&\quad + \sum_{j=0}^{h} [\pi_e \tau_j] [\pi_e \tau_{h+j}] \sigma^2 + \sum_{j=0}^{h} [\pi_d \tau_j] [\pi_d \tau_{h+j}] \sigma_w^2 \\
&\quad + \sum_{j=0}^{h-2} [\phi b_j \phi c_e] [\rho^{j+1} \phi b_{h-1} \phi c_e] \sigma^2 + \phi_d^2 \left( \sum_{j=0}^{h-2} \tau_j \tau_{h+j} \right) \sigma^2
\end{align*}
\[ E \left[ (\Delta d_{t+1}^e - E(\Delta d_{t+1}^e))^2 \right] = 3E \left\{ \left( \sum_{j=0}^{h} [\phi a_{j-1} \varphi_c] \nu_h(t+1)-j + \sum_{j=0}^{h-2} [\phi b_j \varphi_c] \nu_{ht-j-1} \right)^2 + \left( \sum_{j=0}^{h} [\pi \sigma \tau_j] s_h(t+1)-j + \sum_{j=0}^{h-2} [\pi \sigma \tau_{h+j+1}] s_{ht-j-1} \right)^2 \right\} \]

\[ = 3E \left\{ \left( \sum_{j=0}^{h} [\phi a_{j-1} \varphi_c]^2 \nu_h(t+1)-j + \sum_{j=0}^{h-2} [\phi b_j \varphi_c]^2 \nu_{ht-j-1} \right)^2 + \left( \sum_{j=0}^{h} [\pi \sigma \tau_j]^2 s_h(t+1)-j + \sum_{j=0}^{h-2} [\pi \sigma \tau_{h+j+1}]^2 s_{ht-j-1} \right)^2 \right\} \]

\[ = 3 \left( \sum_{j=0}^{h} \sum_{k=j+1}^{h} \left( [\phi a_{j-1} \varphi_c]^2 + [\pi \sigma \tau_j]^2 \right) \left[ \pi \sigma \tau_{k-j} \right] \right) \sigma_w^2 \]

\[ \text{Cov} (\Delta c_{t+1}^a, \Delta d_{t+1}^p) = \text{Cov} \left( h_{\mu_c} + b_{h-1}x_h(t-1) + \sum_{j=0}^{h-2} \tilde{d}_j \nu_h(t+1)-j + \sum_{j=0}^{h} \tilde{b}_j \nu_{ht-j-1} + \sum_{j=0}^{h} c_j s_h(t+1)-j \right) \]

\[ + \sum_{j=0}^{h-2} d_j s_{ht-j-1} + \sum_{j=0}^{h} f_j w_h(t+1)-j + \sum_{j=0}^{h-2} g_j w_{ht-j-1} + \sum_{j=0}^{h} \phi b_{h-1}x_h(t-1) \]

\[ + \sum_{j=0}^{h} [\phi a_{j-1} \varphi_c] \nu_h(t+1)-j + \sum_{j=0}^{h-2} [\phi b_j \varphi_c] \nu_{ht-j-1} + \sum_{j=0}^{h} [\pi \sigma \tau_j] s_h(t+1)-j \]

\[ + \sum_{j=0}^{h} [\pi \sigma \tau_{h+j+1}] s_{ht-j-1} + \sum_{j=0}^{h} [\pi \sigma \tau_j] w_h(t+1)-j + \sum_{j=0}^{h-2} [\pi \sigma \tau_{h+j+1}] w_{ht-j-1} \]

\[ + \sum_{j=0}^{h-2} \tau_j \varphi_d U_h(t+1)-j \]

\[ \text{Cov} (\Delta c_{t+1}^a, \Delta d_{t+1}^p) = \phi b_{h-1}^2 \text{Var} (x_{h(t-2)}) + \sum_{j=0}^{h} [\phi a_{j-1} \varphi_c] \tilde{a}_j \tilde{a}^2 + \sum_{j=0}^{h} [\pi \sigma \tau_j] c_j \tilde{a}^2 \]

\[ + \sum_{j=0}^{h} [\pi \sigma \tau_j] f_j \sigma_w^2 + \sum_{j=0}^{h-2} [\pi \sigma \tau_{h+j+1}] \tilde{d}_j \tilde{a}^2 + \sum_{j=0}^{h-2} [\phi b_j \varphi_c] \tilde{b}_j \tilde{a}^2 \]

\[ + \sum_{j=0}^{h-2} [\pi \sigma \tau_{h+j+1}] g_j \tilde{a}^2 \]

- **Annual market return**

Let's denote
\[
\Gamma_0 = k_0 m + (k_1 m - 1)A_0 m + k_1 m A_2 m (1 - \nu) \sigma^2 + \mu_d
\]
(94)

\[
\Gamma_1 m = A_1 m (k_1 m \rho - 1) + \phi; \quad \Gamma_2 m = (k_1 m \nu - 1) A_2 m
\]

\[
\beta_{m,c} = \pi_c; \quad \beta_{m,x} = k_1 m A_1 m \varphi_c; \quad \beta_{m,w} = \pi_\sigma + k_1 m A_2 m; \quad \beta_{m,d} = \varphi_d
\]

The monthly return on the market portfolio is given by:

\[
m_{m,t+1} = \Gamma_0 m + \Gamma_1 m x_t + \Gamma_2 m \sigma_t^2 + \beta_{m,c} s_{t+1} + \beta_{m,x} v_t + \beta_{m,w} w_t + \beta_{m,d} u_t
\]

So the aggregate return on the market portfolio can be obtained as:

\[
\forall j \geq 1,
\]

\[
r_{m,t+j}^a = \sum_{k=0}^{h-1} \left[ \Gamma_{0 m} + \Gamma_{1 m} x_{h(t+j)-k} + \Gamma_{2 m} \sigma_{h(t+j)-k}^2 + \beta_{m,c} s_{h(t+j)-k} + \beta_{m,x} v_{h(t+j)-k} + \beta_{m,w} w_{h(t+j)-k} + \beta_{m,d} u_{h(t+j)-k} \right]
\]

\[
+ \sum_{k=0}^{h-1} \sum_{r=0}^{h-1} \left[ \rho_{h(t+j)-k} \varphi_{h(t+j)-r} \sigma_{h(t+j)-r}^2 + \rho_{h(t+j)-k} \varphi_{h(t+j)-r} \sigma_{h(t+j)-r}^2 + \beta_{m,x} s_{h(t+j)-k} + \beta_{m,w} w_{h(t+j)-k} + \beta_{m,d} u_{h(t+j)-k} \right]
\]

\[
r_{m,t+j}^a = h \Gamma_{0 m} + \Gamma_{2 m} \left( h - \nu^h \left( 1 - \nu^h \right) \right) \sigma^2 + \Gamma_{1 m} \varphi^h \left( 1 - \nu^h \right) x_h(t-1) + \Gamma_{2 m} \varphi^h \left( 1 - \nu^h \right) \sigma_h(t-1)
\]
(95)

And
\[ r_{m,t}^a = h \Gamma_{0m} + \Gamma_{2m} \left( h - \left( 1 - \frac{\nu^h}{1 - \nu} \right) \right) \vartheta^2 + \Gamma_{1m} \left( 1 - \frac{\rho^h}{1 - \rho} \right) x_{h(t-1)} + \Gamma_{2m} \left( \frac{1 - \nu^h}{1 - \nu} \right) \sigma_{h(t-1)}^2 \] 
\[ + \sum_{k=0}^{h-1} \beta_{m,c} \epsilon_{ht-k} + \sum_{k=0}^{h-1} \beta_{m,x} + \varphi x_{m,c} + \varphi x_{1m} \left( \frac{1 - \rho^k}{1 - \rho} \right) \right] v_{ht-k} + \sum_{k=0}^{h-1} \beta_{m,d} u_{h(t+j)-k} \]
\[ + \sum_{k=0}^{h-1} \left[ \beta_{m,w} + \Gamma_{2m} \left( \frac{1 - \nu^k}{1 - \nu} \right) \right] w_{ht-k} \]

Therefore,

\[ E(r_{m,t}^a) = h \Gamma_0 + h \Gamma_{2m} \sigma^2 \] (98)

\[ \text{Var}(r_{m,t}^a) = \left( \Gamma_{1m} \left( \frac{1 - \rho^h}{1 - \rho} \right) \right)^2 \vartheta^2 + \left( \Gamma_{2m} \left( \frac{1 - \nu^h}{1 - \nu} \right) \right)^2 \vartheta^2 + h \beta_{m,w}^2 \sigma^2 \] (99)

\[ + 2 \Gamma_{1m} \Gamma_{2m} \left( \frac{1 - \rho^h}{1 - \rho} \right) \left( \frac{1 - \nu^h}{1 - \nu} \right) \text{Cov}(x_t, \sigma^2_t) + \sum_{k=0}^{h-1} \beta_{m,c}^2 \sigma^2 \]
\[ + \sum_{k=0}^{h-1} \left[ \beta_{m,x} + \varphi \epsilon_{1m} \left( \frac{1 - \rho^k}{1 - \rho} \right) \right] \sigma^2 + \sum_{k=0}^{h-1} \left[ \beta_{m,w} + \Gamma_{2m} \left( \frac{1 - \nu^k}{1 - \nu} \right) \right] \sigma^2 \]

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

\[ \text{Cov}(r_{m,t}^a, r_{m,t+1}^a) = \rho^h \left[ \Gamma_{1m} \left( \frac{1 - \rho^h}{1 - \rho} \right) \right]^2 \vartheta^2 + \nu^h \left[ \Gamma_{2m} \left( \frac{1 - \nu^h}{1 - \nu} \right) \right]^2 \vartheta^2 \]
\[ + \sum_{k=0}^{h-1} \left[ \beta_{m,x} + \varphi \epsilon_{1m} \left( \frac{1 - \rho^k}{1 - \rho} \right) \right] \varphi \epsilon_{1m} \left( \frac{1 - \rho^k}{1 - \rho} \right) \rho^k \sigma^2 \]
\[ + \sum_{k=0}^{h-1} \left[ \beta_{m,w} + \Gamma_{2m} \left( \frac{1 - \nu^k}{1 - \nu} \right) \right] \right] \Gamma_{2m} \left( \frac{1 - \nu^k}{1 - \nu} \right) \nu^k \sigma^2 \]
\[ = \rho \left( \frac{\varphi \sigma^2}{1 - \rho^2} \right) \left( \frac{1 - \rho^h}{1 - \rho} \right)^2 \Gamma_{1m}^2 + \nu \left( \frac{\sigma^2_w}{1 - \nu^2} \right) \left( \frac{1 - \nu^h}{1 - \nu} \right)^2 \Gamma_{2m}^2 \]
\[ + \beta_{m,x} \Gamma_{1m} \varphi \sigma^2 \left( \frac{1 - \rho^h}{1 - \rho} \right)^2 + \beta_{m,w} \Gamma_{2m} \sigma^2 \left( \frac{1 - \nu^h}{1 - \nu} \right)^2 \]

- **Annual risk free rate**

\[ \forall j \geq 1, \]
\begin{equation}
\begin{split}
r_{f,t+j}^o &= \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_{h(t+j)-r}^2 \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) \\
&+ A_{2f} \left( \nu^{h(j+1)-r} \sigma_{h(t-1)}^2 + \sum_{k=0}^{h(j+1)-r-1} (\nu^k w_{h(t+j)-r-k} + (1 - \nu) \bar{\sigma}^2 \nu^k) \right)
\end{split}
\end{equation}

\begin{equation}
\begin{split}
r_{f,t+j}^o &= hA_{0f} + A_{2f} \left[ h - \nu^{hj+1} \left( \frac{1 - \nu^h}{1 - \nu} \right) \right] \bar{\sigma}^2 + A_{1f} \nu^{hj+1} \left( \frac{1 - \nu^h}{1 - \nu} \right) x_{h(t-1)} + A_{2f} \nu^{hj+1} \left( \frac{1 - \nu^h}{1 - \nu} \right) \sigma_{h(t-1)}^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=0}^{h-1} \rho^r \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}^{h-1} \nu^r \left( \frac{1 - \nu^h}{1 - \nu} \right) w_{h(t+j)-1-r} \right]
\end{split}
\end{equation}

and

\begin{equation}
\begin{split}
r_{f,t}^o &= 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_{h(t-1)}^2 \\
&+ A_{1f} \varphi e \left[ \sum_{r=0}^{h-1} \left( \frac{1 - \rho^r + 1}{1 - \rho} \right) v_{h(t)-r} + \sum_{r=0}^{h-1} \left( \frac{1 - \nu^r + 1}{1 - \nu} \right) w_{h(t)-r} \right]
\end{split}
\end{equation}

Therefore,

\begin{equation}
\mathbb{E}(r_{f,t}^o) = hA_{0f} + hA_{2f}
\end{equation}

\begin{equation}
\text{Var}(r_{f,t}^o) = \left[ A_{1f} \rho \left( \frac{1 - \rho^h}{1 - \rho} \right) \right]^2 \text{Var}(x_t) + \left[ A_{2f} \nu \left( \frac{1 - \nu^h}{1 - \nu} \right) \right]^2 \text{Var}(\sigma_t^2) \\
+ \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_w^2
\end{equation}

\begin{equation}
\text{Cov}(r_{f,t}^o, r_{f,t+1}^o) = \rho^h \left[ A_{1f} \rho \left( \frac{1 - \rho^h}{1 - \rho} \right) \right]^2 \text{Var}(x_t) + \nu^h \left[ A_{2f} \nu \left( \frac{1 - \nu^h}{1 - \nu} \right) \right]^2 \text{Var}(\sigma_t^2) \\
+ (A_{1f} \varphi e)^2 \sum_{r=0}^{h-1} \left[ \rho^r \left( \frac{1 - \rho^r + 1}{1 - \rho} \right) \right]^2 \sigma^2 \\
+ \sum_{r=0}^{h-1} \left[ \frac{1 - \nu^r + 1}{1 - \nu} \right] A_{2f} \left[ A_{2f} \nu \left( \frac{1 - \nu^h}{1 - \nu} \right) \right]^2 \sigma_w^2
\end{equation}
**Forward annual excess return**

Using 95 and 102, the aggregate excess return on the market portfolio can be expressed as follows:

\[ r_{ex,t+j}^a = r_{m,t+j}^a - r_{f,t+j}^a \]

\[ = h (\Gamma_0m - A_{0f}) + h (\Gamma_{2m} - A_{2f}) \sigma^2 - (\Gamma_{2m} - \nu A_{2f}) \left( \nu^h \left( \frac{1 - \nu^h}{1 - \nu} \right) \right) \sigma^2 \]

\[ + (\Gamma_{1m} - \rho A_{1f}) \rho^h \left( \frac{1 - \rho^h}{1 - \rho} \right) x_{h(t-1)} + (\Gamma_{2m} - \nu A_{2f}) \nu^h \left( \frac{1 - \nu^h}{1 - \nu} \right) \sigma^2_{h(t-1)} \]

\[ + \sum_{k=0}^{h-1} \beta_{m,c} \sigma h_{t+k} + \sum_{k=0}^{h-1} [\beta_{m,x} + \varphi_e \left( \Gamma_{1m} \left( \frac{1 - \rho^h}{1 - \rho} \right) - A_{1f} \left( \frac{1 - \rho^{k+1}}{1 - \rho} \right) \right) \right] u_{h(t+k)-k} \]

\[ + \varphi_e \left( \frac{1 - \rho^h}{1 - \rho} \right) \sum_{r=0}^{h-1} \rho^r \left[ \Gamma_{1m} - A_{1f} \right] v_{h(t+k-1)-r} + \left( \frac{1 - \rho^h}{1 - \nu} \right) \sum_{k=0}^{h-1} \nu^k \left[ \Gamma_{2m} - A_{2f} \nu \right] w_{h(t+k-1)-k} \]

\[ + \sum_{k=0}^{h-1} \beta_{m,d} \sigma u_{h(t+k)-k} + \sum_{k=0}^{h-1} \left[ \beta_{m,w} + \Gamma_{2m} \left( \frac{1 - \nu^h}{1 - \nu} \right) - \left( A_{2f} \left( \frac{1 - \rho^{k+1}}{1 - \rho} \right) \right) \right] w_{h(t+k)-k} \]

In particular,

\[ r_{ex,t}^a = r_{m,t}^a - r_{f,t}^a \]

\[ = h (\Gamma_0m - A_{0f}) + h (\Gamma_{2m} - A_{2f}) - (\Gamma_{2m} - \nu A_{2f}) \left( \frac{1 - \nu^h}{1 - \nu} \right) \sigma^2 \]

\[ + (\Gamma_{1m} - \rho A_{1f}) \left( \frac{1 - \rho^h}{1 - \rho} \right) x_{h(t-1)} + (\Gamma_{2m} - \nu A_{2f}) \left( \frac{1 - \nu^h}{1 - \nu} \right) \sigma^2_{h(t-1)} \]

\[ + \sum_{k=0}^{h-1} \beta_{m,c} \sigma h_{t+k} + \sum_{k=0}^{h-1} [\beta_{m,x} + \varphi_e \left( \Gamma_{1m} \left( \frac{1 - \rho^h}{1 - \rho} \right) - A_{1f} \left( \frac{1 - \rho^{k+1}}{1 - \rho} \right) \right) \right] u_{h(t+k)-k} \]

\[ + \sum_{k=0}^{h-1} \beta_{m,d} \sigma u_{h(t+k)-k} + \sum_{k=0}^{h-1} \left[ \beta_{m,w} + \Gamma_{2m} \left( \frac{1 - \nu^h}{1 - \nu} \right) - \left( A_{2f} \left( \frac{1 - \rho^{k+1}}{1 - \rho} \right) \right) \right] w_{h(t+k)-k} \]

**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.
\[ p_t^0 - d_t^0 = \log P_{ht} - \log \sum_{j=0}^{h-1} D_{ht-j} \]
\[ = \log P_{ht} - \log D_{ht} + \sum_{j=0}^{h-1} \left( \log D_{ht-j} - \log D_{ht-j-1} \right) - \log h - \sum_{j=0}^{h-1} \left( \frac{j+1}{h} \right) \triangle d_{ht-j} \]
\[ = z_{m,ht} \sum_{j=0}^{h-1} \triangle d_{ht-j} - \log h - \sum_{j=0}^{h-1} \left( \frac{j+1}{h} \right) \triangle d_{ht-j} \]
\[ = z_{m,ht} - \log h - \sum_{j=0}^{h-1} \left( \frac{j+1}{h} - 1 \right) \triangle d_{ht-j} \] (107)

\[ p_t^0 - d_t^0 = A_{0m} + A_1 x_{ht} + A_2 \sigma_{ht}^2 - \log h - \sum_{j=0}^{h-1} \left( \frac{j+1}{h} - 1 \right) \left( \mu_d + \phi x_{ht-j-1} + \pi_c s_{ht-j} + \sigma_d u_{ht-j} + \sigma_x w_{ht-j} \right) \]
\[ + \varphi_d u_{ht-j} + \pi_x w_{ht-j} \left( \frac{1}{2} \right) (108) \]

and

\[ p_{t+1}^0 - d_{t+1}^0 = A_{0m} + A_1 x_{h(t+1)} + A_2 \sigma_{h(t+1)}^2 - \log h - \sum_{j=0}^{h-1} \left( \frac{j+1}{h} - 1 \right) \left( \mu_d + \phi x_{h(t+1)-j-1} + \pi_c s_{h(t+1)-j} + \sigma_d u_{h(t+1)-j} + \sigma_x w_{h(t+1)-j} \right) \]
\[ = \left[ A_{0m} - \log h - \frac{1}{2} \left( h - 1 \right) \mu_d + A_2 \left( 1 - \nu^h \right) \sigma^2 \right] + \rho^h \left[ A_1 \mu \phi x_{h-1} + A_2 \nu^2 \sigma_{h(t+1)}^2 \right] \]
\[ + \sum_{j=0}^{h-1} \rho^j \left( A_1 \mu \phi x_{h-1} \right) \varphi_d u_{ht-j} + \sum_{j=0}^{h-1} \left[ \left( A_1 \mu \phi x_{h-1} \right) \varphi_c s_{ht-j} + \sum_{j=0}^{h-1} \left( \frac{j+1}{h} - 1 \right) \pi_x w_{ht-j} \right] \]
\[ + \sum_{j=0}^{h-1} \sum_{j=0}^{h-1} \left( \frac{j+1}{h} - 1 \right) \pi_c s_{ht(t+1)-j} + \sum_{j=0}^{h-1} A_2 \nu^{h+j} w_{ht-j} - \sum_{j=0}^{h-1} \left( \frac{j+1}{h} - 1 \right) \varphi_d u_{ht-j} \]
\[ + \sum_{j=0}^{h-1} A_2 \nu^{h+j} \left( \frac{j+1}{h} - 1 \right) \pi_x w_{ht(t+1)-j} \]

Where
\[ a_j = a_j - \left( \frac{1 - \rho^{j+1}}{1 - \rho} \right) \] for \( j \in \{ -1, 0, \ldots, h - 1 \} \)

Therefore,

\[ E(p_t^0 - d_t^0) = A_{0m} - \log h - \frac{1}{2} \left( h - 1 \right) \mu_d + A_2 \left( 1 - \nu^h \right) \sigma^2 \] (109)
\[ \text{Var} (p_t^a - d_t^a) = \left[ A_{1m} \rho^h - \phi a'_{h-1} \right]^2 \text{Var} (x_t) + \left[ A_{2m} \nu^h \right]^2 \text{Var} (\sigma_t^a) + 2 \left[ A_{1m} \rho^h - \phi a'_{h-1} \right] A_{2m} \nu^h \text{Cov} (x_t, \sigma_t^a) \] (109)

\[ + \sum_{j=0}^{h-1} \left[ \left( A_{1m} \rho^j - \phi a'_{j-1} \right) \varphi_c \right] \sigma^2 + \sum_{j=0}^{h-1} \left[ \left( \frac{j+1}{h} - 1 \right) \pi_c \right] \sigma^2 \]

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

\[ \text{Cov} (p_t^a - d_t^a, p_{t+1}^a - d_{t+1}^a) = \rho^h \left[ A_{1m} \rho^h - \phi a'_{h-1} \right]^2 \text{Var} (x_t) + (\rho^h + \nu^h) \left[ A_{1m} \rho^h - \phi a'_{h-1} \right] A_{2m} \nu^h \text{Cov} (x_t, \sigma_t^a) \] (110)

\[ + \nu^h \left[ A_{2m} \nu^h \right]^2 \text{Var} (\sigma_t^a) + \sum_{j=0}^{h-1} \rho^j \left[ \left( A_{1m} \rho^j - \phi a'_{j-1} \right) \varphi_c \right] \left( A_{1m} \rho^h - \phi a'_{h-1} \right) \varphi_\sigma \sigma^2 \]

\[ + \sum_{j=0}^{h-1} \left[ A_{2m} \nu^j + j \right] \sigma^2 \]

10.9 Theoretical moments for the predictive regression

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

\[ \forall j \geq 1, \]

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

\[ + (\Gamma_{2m} - \nu A_{2f}) \nu^{h(j+1)} \left( \frac{1 - \nu^h}{1 - \nu} \right) A_{2m} \text{Var} (\sigma_t^a) \]

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

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

- Useful transformations

\[ x_{h(t+j)-k} = \rho^{h(j+1)-k} x_{h(t-1)} + \sum_{r=0}^{h(j+1)-k-1} (\rho^j \varphi_c \nu_{h(t+j)-k-r}) \] (111)

\[ \sum_{r=0}^{h(j+1)-1-r} \sum_{k=0}^{r} \rho^k s_{h(t+j)-r-k} = \sum_{r=0}^{h-1} \left( \frac{1 - \rho^r}{1 - \rho} \right) s_{h(t+j)-r} + \sum_{r=0}^{h(j+1)-1} \left( \frac{1 - \rho^h}{1 - \rho} \right) s_{h(t+j)-1-r} \] (112)
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. Fortunately 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. 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 bootstrap; 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.

\footnote{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 2.}