TI 2015-066/VI
Tinbergen Institute Discussion Paper

Inflation, Endogenous Market Segmentation and the Term Structure of Interest Rates

Casper de Vries
Xuedong Wang

Erasmus School of Economics, Erasmus University Rotterdam, and Tinbergen Institute, the Netherlands.

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: http://www.dsf.nl/

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

Inflation, Endogenous Market Segmentation and the Term Structure of Interest Rates

Casper G. de Vries* Xuedong Wang†

May 18, 2015

Abstract

The term structure of interest rates does not adhere to the expectations hypothesis, possibly due to a risk premium. We consider the implications of a risk premium that arises from endogenous market segmentation driven by variable inflation rates. In the absence of autocorrelation in inflation, the risk premium is constant. If inflation is correlated, however, the risk premium becomes time varying and we can rationalize the failure of the expectations hypothesis. Indirect empirical tests of the model's implications are provided.

JEL Classification: E43, G12

Keywords: Expectations hypothesis; Term structure; Time-Varying Risk Premia; Segmented markets; Inflation

1 Introduction

The expectations hypothesis (EH) of the yield curve of interest rates holds that the n-period yield is a weighted average of the m period yield, $m < n$, and the expected $(n-m)$-period yield. Empirical work has consistently rejected this linear relationship. The failure of the EH in term structure of the yield curve was first observed by Fama and Bliss (1987). Many other papers have subsequently documented the same result of Famma, see

*Erasmus School of Economics, Erasmus University Rotterdam and Chapman University

†Corresponding author: Erasmus School of Economics, Erasmus University Rotterdam and Tinbergen Institute, P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands. Email: xwang@ese.eur.nl. We are grateful to Wouter den Haan who commented on a first draft of the paper.

e.g. Campbell and Shiller (1991), Hardouvelis (1994), Gerlach and Smets (1997), Bekaert and Hodrick (2001) to name a few.

Currently, the main explanation for the rejection of EH in the term structure is the existence of time varying risk premia.[1] Due to the rather small variations in aggregate consumption, a representative agent model with a standard utility function cannot generate large and variable risk premia. For this reason, the explanations based on such framework, including the work of Backus, Gregory, and Zin (1989), Donaldson, Johnsen, and Mehra (1990), den Haan (1995) and Bekaert, Hodrick, and Marshall (1997), encounter difficulties. The time varying risk premia have been explained by means of habit formation in preference functions, see e.g. Brandt and Wang (2003), Wachter (2006), or by means of the Epstein-Zin recursive recursive utility function, as in Bansal and Shaliastovich (2007), and Rudebusch and Swanson (2012). A different avenue has been followed by Buraschi and Jiltsov (2005), who rely on a tax wedge.

There is a smaller but growing literature that explains time varying risk premia from endogenous market segmentation. Limited participation implies that a subset of the population is responsible for the adjustments in financial markets. This generates larger risk premia than representative agent models in which everyone participates in (complete) financial markets. In addition to the exogenous assumption that there are different types of households, there are two effective ways to generate endogenous segmentation between households.

One assumption is that financial markets are complete coupled with limited enforcement of financial contracts.[2] Seppala (2004) compares the real term structure under the no friction Lucas economy (see Lucas (1978)) and the limited risk-sharing Alvarez-Jermann economy (Alvarez and Jermann (2001)). Seppala shows that "only the model with limited risk-sharing can generate enough variation in the term premia to account for the rejections of expectations hypothesis" (Seppala (2004), p1510).

The second avenue is the segmented market approach. An endogenous market segmentation model was developed by Alvarez, Atkeson, and Kehoe (2002) and (2009). In Al-

[1] Several other explanations have been offered. For example, Bekaert, Hodrick, and Marshall (2001) refer to the peso problem. Bansal and Zhou (2002) use a regime shifting model, while Sinha (2009) considers learning behavior.

[2] The idea of limited contract enforceability was first used by Kehoe and Levine (1993) for investigating the behavior of asset markets. Later it was used by Alvarez and Jermann (2001) for examining the implications of asset pricing and by Kehoe and Perri (2002) for exploring the implications of the international business cycle.

varez, Atkeson, and Kehoe (2009), the endogenously segmented market model generates sufficient large variation of the time-varying risk premia to resolve the forward premium puzzle of exchange rates.

Inspired by the model of Alvarez, Atkeson, and Kehoe (2009), we address the expectations puzzle with an endogenously segmented market model. We solve the model with correlated inflation and show that this is necessary for rejection of the EH. The intuition is as follows. In the absence of autocorrelation in inflation, the risk premium is constant. This is unimportant in the two period model of Alvarez, Atkeson, and Kehoe (2009). But to explain the the yield curve, one needs a multiperiod model in which pricing kernels are correlated. If inflation is correlated, however, the risk premium becomes time varying and we can rationalize the failure of the expectations hypothesis. We also provide a test for the model by relating the size of the bias to the amount of autocorrelation and the level of the variance of the money supply shocks.

The remainder of the paper is organized as follows. Section 2 reviews the empirical tests for the EH in term structure. Section 3 gives a concise description of the model. Section 4 provides the theoretical proof that the consumption-based model can reslove the expectations puzzle. Section 5 offers numerical backing. Section 6 presents the empirical evidence for our results. Section 7 concludes.

2 Empirical Evidence Review

Using postwar U.S. bond yields, Campbell and Shiller (1991) show the values of β in regressions like

$$y_{t+m,n-m} - y_{t,n} = const + \beta_{n,m}\frac{m}{n-m}(y_{t,n} - y_{t,m}) + error$$

are negative and increase in absolute value with maturity n for almost any combination of maturities between one month and ten years. The EH holds that $\beta = 1$. This evidence constitutes the expectations puzzle of the term structure, first documented by Fama and Bliss (1987) and Campbell and Shiller (1991).[3] In addition to nominal yields, the real yields also reject the EH, see Bansal and Shaliastovich (2007) and Sinha (2009).

[3] A related rejection of the EH exists in the foreign exchange markets. This was first documented by Fama (1984) and is usually referred to as the forward premium puzzle.

Backus, Foresi, Mozumdar, and Wu (2001) and Buraschi and Jiltsov (2005) use U.S. data and obtain results similar to Campbell and Shiller (1991). Empirical tests conducted on interest rates outside the US, however, often show that the EH cannot be rejected at multiple horizons.[4] More importantly, empirical results show that yields of long-term maturities are more likely to violate the EH than yields of short-term maturities. Hardouvelis (1994) tests the long-term (10-year) and short-term (three-month) bond yields in G7 countries. When the long-term yields are used, five out the seven countries give negative β. However, for the short-term yields, all countries have a positive β, though all of these are lower than 1. Gerlach and Smets (1997) test the EH with 1-, 3-, 6- and 12-month Euro-rates. Their results show that the EH cannot be rejected for 35 cases out of the total 51 cases. Longstaff (2000) tested the expectations hypothesis at the extreme short end of the term structure using overnight, weekly, and monthly repo rates. His results show the expectations hypothesis cannot be rejected at any maturity level.

Table 1 shows the results of our test using Euro-rates for the currencies of 17 countries. Our results also show that the negative $\beta_{n,m}$ is not universal. The empirical values of β are either smaller or larger than unity. For example, all the values of $\beta_{n,m}$ for French franc rates are larger than unity and increase in n. So the expectations hypothesis is rejected in two directions and one would also like to be able to account for this. Moreover, the decreasing trend of β in n is observed in only a few countries, such as Australia and the U.S. While for most countries, this trend is absent.

3 The Model with Segmented Markets

The baseline model that we use is the same with that of Alvarez, Atkeson, and Kehoe (2009). This section provides a concise description of the model.

In the economy there is an asset market and a goods market. Households can buy and sell government bonds in the asset market. The Government injects money in the asset market by paying the maturing bonds. This determines the money growth rate μ_t. In each period, the shock to money growth is the only source of uncertainty in this economy. At the end of each period, each household receives the same real endowment y.

[4]For the empirical test of EH for other countries, see Hardouvelis (1994), Evans and Lewis (1994) Gerlach and Smets (1997), Longstaff (2000), Dominguez and Novales (2000), Bekaert and Hodrick (2001), and Jongen, Verschoor, and Wolff (2011) et al.

Households sell their endowment at the current price level P_t to obtain money $P_t y$ and transfer money to the next period for consumption in the goods market or for buying bonds in the asset market. Money can be transferred between the goods market and the asset market. Households who transfer money between the two markets have to pay a real transfer cost γ. Besides the traditional interpretation for such a cost as the brokerage fee, the bid-ask spread and the transaction tax, the literature explores more motives for this cost. Chatterjee and Corbae (1992) view the transfer cost as a cost involved in writing enforceable private debt contracts. Reis (2006) and Alvarez, Guiso, and Lippi (2012) consider the costs of acquiring, absorbing and processing information. Gust and Lopez-Salido (2014) interpret the presence of the transfer cost as reflecting time spent on the activities of re-optimizing and responding to new information, and the human inertia of sticking to a predetermined plan. Thus one can consider the transfer cost as the aggregate effect of all kinds of frictions which can be tangible or intangible. The value of γ varies across households with a distribution $F(\gamma)$ and density $f(\gamma)$.

There is no storage technology available in the economy except for money, so the endowment y of each period has to be consumed within the same period. The consumption of households is subject to the cash-in-advance constraint and the transition law. In period t, given state s^t,

$$c(s^t, \gamma) = \frac{P(s^{t-1})y}{P(s^t)} + x(s^t, \gamma)z(s^t, \gamma) \tag{3.1}$$

The resource constraint is given by

$$\int \left[c(s^t, \gamma) + \gamma z(s^t, \gamma)\right] f(\gamma) d\gamma = y \tag{3.2}$$

where $c(s^t, \gamma)$ is the consumption of a household in period t and $x(s^t, \gamma)$ is the real balance that the household chooses to transfer between the two markets. If the value of $x(s^t, \gamma)$ is positive, it means that money is transferred from the asset market into the goods market, and vice versa. The indicator variable $z(s^t, \gamma)$ is equal to 0 if $x(s^t, \gamma)$ is 0, otherwise, it is equal to 1.

It is assumed that households hold their assets in interest-bearing bonds rather than cash. This assumption is intuitive because bonds have tended to dominate the zero return on cash as long as nominal interest rates are positive. With this assumption, we have that the inflation rate π_t is equal to the money growth rate μ_t, i.e. $\pi_t = \mu_t = \frac{P_t}{P_{t-1}}$. So (3.1) can

be written as

$$c(s^t, \gamma) = \frac{y}{\mu_t} + x(s^t, \gamma) z(s^t, \gamma) \tag{3.3}$$

Households are divided into two types, labeled as the active and inactive households, depending on whether they transfer assets between goods and financial markets or not. We denote $c_A(s^t, \gamma)$ as the consumption of an active household for a given s^t. The consumption of both kinds of households adds up to

$$c(s^t, \gamma) = z(s^t, \gamma) c_A(s^t, \gamma) + [1 - z(s^t, \gamma)] y/\mu_t, \text{ with } \begin{cases} z(s^t, \gamma) = 0, \text{ if } x(s^t, \gamma) = 0 \\ z(s^t, \gamma) = 1, \text{ if } x(s^t, \gamma) \neq 0 \end{cases} \tag{3.4}$$

The expression (3.4) shows the consumption for the financially inactive households is pinned down by their real money balances y/μ_t as $z(s^t, \gamma) = 0$.

Inflation reduces the consumption of the inactive households from y to y/μ_t. The higher the inflation is, the lower the consumption of the inactive households will be. This effect of inflation causes some households to choose to pay the transfer cost and transfer some assets from the asset market into the goods market, so that they can compensate their loss of consumption due to inflation. If transfers are costless, all households would choose to transfer. The difference in transfer cost leads to the segmentation between households. The cost reduces the total amount of resources available for consumption. The combination of inflation and the transfer cost forms the only distortion in the model.

With this feature and the assumption of a complete financial market, the competitive equilibrium allocations and asset prices can be found from the solution of the planner's problem. Recall there is no storage technology, so the social planner's problem reduces to a sequence of static optimization problems as

$$\max \int U(c(s^t, \gamma)) f(\gamma) d\gamma \tag{3.5}$$

subject to the constraints (3.2) and (3.4). From (3.5), we get that the planning weight for households of type γ is just the fraction of households of this type. The requirement for this simple configuration is that all the households have equal Lagrange multipliers on their period zero budget constraints.

When $z(s^t)$ is fixed at 1, the first-order condition for c_A reduces to

$$U'(c_A(s^t, \gamma)) = \lambda(s^t) \tag{3.6}$$

where $\lambda(s^t)$ is the multiplier on the resource constraint, which is identical for all households. This result implies that all active households choose the same consumption level $c_A(s^t, \gamma)$ independent of γ. The reason that c_A is independent of γ is that the transfer cost γ is charged in a lump sum way. This does not have a distorting effect on the consumption of active households. After paying their cost, active households are identical. Thus, all active households choose the same consumption level. The result that $c_A(s^t, \gamma)$ is independent of γ combined with the static nature of the problem tells us that $c_A(s^t)$ only depends on the current money growth rate μ_t, so that $c_A(s^t)$ can be denoted as $c_A(\mu_t)$.

The planner's problem thus reduces to choosing $c_A(\mu)$ and to determining the fractions of active and inactive households to maximize the social welfare. Denote $\bar{\gamma}(\mu)$ as the threshold at a given money growth rate, which separates the two types of households. The households with $\gamma \leq \bar{\gamma}(\mu)$ pay their cost and consume $c_A(\mu)$. Otherwise, the households choose to be inactive with the consumption level of y/μ. For a given μ, the planner's problem is thus reduced to choosing $c_A(\mu)$ and $\bar{\gamma}(\mu)$ to solve

$$\max U(c_A(\mu))F(\bar{\gamma}(\mu)) + U(y/\mu)[1 - F(\bar{\gamma}(\mu))] \tag{3.7}$$

subject to

$$c_A(\mu)F(\bar{\gamma}(\mu)) + \int_0^{\bar{\gamma}(\mu)} \gamma f(\gamma) d\gamma + \frac{y}{\mu}[1 - F(\bar{\gamma}(\mu))] = y \tag{3.8}$$

The first-order condition for optimal consumption and the transaction distortion reads[5]

$$U(c_A(\mu)) - U(y/\mu) + U'(c_A(\mu))\left[(y/\mu) - \bar{\gamma}(\mu) - c_A(\mu)\right] = 0 \tag{3.9}$$

The social planner's problem is consistent with the functioning of the decentralized economy. In this setup, the asset prices are given by the multipliers on the resource constraints for the planner's problem. From (3.6) these multipliers are equal to the marginal

[5]See Appendix A for the derivation of (3.9).

utility of active households. Hence, the pricing kernel for nominal bonds is

$$m(s^t, s_{t+1}) = \delta \frac{U'\left(c_A(\mu_{t+1})\right)}{U'\left(c_A(\mu_t)\right)} \frac{1}{\mu_{t+1}}, \tag{3.10}$$

and for real discounted bonds it is

$$m^*(s^t, s_{t+1}) = \delta \frac{U'\left(c_A(\mu_{t+1})\right)}{U'\left(c_A(\mu_t)\right)} \tag{3.11}$$

Define $\phi(\mu)$ as the elasticity of the marginal utility of active households to a change in the money growth rate μ. Thus,

$$\phi(\mu) \equiv -\frac{d \log U'\left(c_A(\mu)\right)}{d \log \mu} \tag{3.12}$$

Define $\eta(\mu)$ as the negative derivative of this elasticity to the log of inflation, thus

$$\eta(\mu) \equiv \frac{d^2 \log U'\left(c_A(\mu)\right)}{(d \log \mu)^2} \tag{3.13}$$

The equilibrium features of the model are captured in two Theorems. [6]

Theorem 1. As μ increase, more households become active. In particular, $\bar{\gamma}'(\mu) > 0$ for $\mu > 1$ and $\bar{\gamma}'(1) = 0$. (Alvarez, Atkeson, and Kehoe (2009), *Proposition* 2, p. 863)

Theorem 2. The log of the consumption of active households $c_A(\mu)$ is strictly increasing and strictly concave in $\log \mu$ around $\mu = 1$. In particular, $\phi(1) > 0$ and $\phi'(1) < 0$. (Alvarez, Atkeson, and Kehoe (2009), *Proposition* 3, p. 863)

Theorem 1 says that if inflation is positive, the fraction of active households is proportional to the rate of inflation. Theorem 2 implies that if inflation is low, both the elasticity ϕ and its derivative η have positive values.

4 Segmented Markets and The Expectations Puzzle

This section offers a theoretical explanation for the expectations puzzle within the realm of segmented markets. The benchmark expectations hypothesis of the term structure of interest rates holds that "the n-period interest rate equals an average of the current

[6] The proofs are in Alvarez, Atkeson, and Kehoe (2009).

short-term rate and the future short-term rates expected to hold over the n-period horizon" (Walsh (2010), p. 465). According to Campbell and Shiller (1991), the relationship between a longer-term n-period interest rate $R_t^{(n)}$ and a shorter-term m-period interest rate $R_t^{(m)}$ can be summarized as:

$$R_t^{(n)} = (1/k)\sum_{i=0}^{k-1} \mathbb{E}_t R_{t+mi}^{(m)} + c, \qquad k = n/m \tag{4.1}$$

According to this theory, the yield curve of any bond or deposit satisfies[7]

$$y_{t,n} = \frac{m}{n} y_{t,m} + \frac{n-m}{n} \mathbb{E}_t y_{t+m,n-m} + c, \qquad n \geq m \tag{4.2}$$

where $y_{i,j}$ indicates the yield (or interest rate) of a j period bond (or deposit) starting at period i.

This implies that the slope coefficient $\beta_{n,m}$ in bond yield or deposit interest rate regressions

$$y_{t+m,n-m} - y_{t,n} = const + \beta_{n,m} \frac{m}{n-m}(y_{t,n} - y_{t,m}) + error \tag{4.3}$$

equals 1 at all maturities n and time steps m. This follows since with $const = 0$,

$$\begin{aligned} \beta_{n,m}(OLS) =& 1 + \frac{Cov[y_n - y_m, (n-m)y_{m,n-m} - ny_n + my_m]}{mVar(y_n - y_m)} \\ =& 1 + \frac{Cov(y_n - y_m, 0)}{mVar(y_n - y_m)} \end{aligned} \tag{4.4}$$

However, if the time-varying risk premia are included in the interest rates, the story changes. Consider a n-period investment in a discount bond with price $P_{t,n}$. Investors may choose to buy a n-period bond with a log return of $\log(1/P_{t,n})$. Alternatively, they can buy a m-period $(m < n)$ bond first and after m periods, roll over and buy a $(n-m)$-period bond with the proceeds from the m-period bond. The log return of the second strategy is $\log[1/(P_{t,m} \cdot P_{t+m,n-m})]$. Here $P_{i,j}$ is the price of a j period bond starting at period i.

The excess log return from a m-period bond at time t plus a $(n-m)$-period bond at $t+m$ is the premium required by the investors for bearing the risk of the open position, because the return on the $(n-m)$-period bond at $t+m$ is unknown at time t. The expected

[7] If the n-period yield is the average of the expected 1-period yields $y_n = E[y_{1,1} + \cdots + y_{n,1}]/n$, then $ny_n = my_m + (n-m)y_{m,n-m}$ and where $y_{m,n-m}$ is the $n-m$ is the $n-m$ period yield starting at time m.

excess return is

$$\begin{aligned}\mathbb{E}_t rx_{t+m,n-m} = p_t =& \mathbb{E}_t \log \frac{1}{P_{t,m} \cdot P_{t+m,n-m}} - \log \frac{1}{P_{t,n}} \\ =& \log \frac{1}{P_{t,m}} + \mathbb{E}_t \log \frac{1}{P_{t+m,n-m}} - \log \frac{1}{P_{t,n}} \\ =& m y_{t,m} + (n-m)\mathbb{E}_t y_{t+m,n-m} - n y_{t,n} \end{aligned} \tag{4.5}$$

and where p_t is the risk premium at time t.

Upon rewriting (4.5), we get

$$\mathbb{E}_t y_{t+m,n-m} - y_{t,n} = \frac{m}{n-m}(y_{t,n} - y_{t,m}) + \frac{p_t}{n-m} \tag{4.6}$$

Given n and m, we can see from (4.6) that if the risk premium p_t is not a constant and correlates with $y_{t,n} - y_{t,m}$, then the slope coefficient $\beta_{n,m}$ in bond regressions (4.3) will differ from unity. The direction of the rejection of EH depends on the correlation between $y_{t,n} - y_{t,m}$ and p_t.

With rational expectations, the population value for the slope coefficient of regression (4.3) is given by[8]

$$\beta_{n,m} = 1 + \frac{Cov(\mathbb{E}_t rx_{t+m,n-m}, y_{t,n} - y_{t,m})}{mVar(y_{t,n} - y_{t,m})} \tag{4.7}$$

The segmented market model in Section 3 implies a risk premium partly driven by inflation. Moreover, the yields $y_{t,n}$ and $y_{t,m}$ are also determined by inflation, since the pricing kernel is a function of inflation. This implies that the covariance in (4.7) is non-zero under the segmented market model. So this model is a promising avenue to explore for solving the expectations puzzle.

Based on the assumption that investors act optimally, the price of nominal bonds is given by the pricing kernel defined in (3.10). This gives

$$P_{t,n} = e^{-n y_{t,n}} = \delta^n \mathbb{E}_t \frac{U'(c_A(\mu_{t+n}))}{U'(c_A(\mu_t))} \frac{1}{\pi_{t,t+n}},$$

$$P_{t,m} = e^{-m y_{t,m}} = \delta^m \mathbb{E}_t \frac{U'(c_A(\mu_{t+m}))}{U'(c_A(\mu_t))} \frac{1}{\pi_{t,t+m}}$$

[8] See Appendix B.1 for the derivation of (4.7).

and

$$P_{t+m,n-m} = e^{-(n-m)y_{t+m,n-m}} = \delta^{n-m}\mathbb{E}_{t+m}\frac{U'(c_A(\mu_{t+n}))}{U'(c_A(\mu_{t+m}))}\frac{1}{\pi_{t+m,t+n}},$$

In these price expressions $\pi_{t,t+n} = \prod_{i=1}^{n}\mu_{t+i}$ is the aggregate inflation over n periods starting from time "t". The definition of $\pi_{t,t+m}$ and $\pi_{t+m,t+n}$ is similar to $\pi_{t,t+n}$. Hence

$$ny_{t,n} = -n\log\delta - \log\mathbb{E}_t\left[\frac{U'(c_A(\mu_{t+n}))}{U'(c_A(\mu_t))}\frac{1}{\pi_{t,t+n}}\right] \tag{4.8}$$

and

$$my_{t,m} = -m\log\delta - \log\mathbb{E}_t\left[\frac{U'(c_A(\mu_{t+m}))}{U'(c_A(\mu_t))}\frac{1}{\pi_{t,t+m}}\right]$$

It follows that

$$\begin{aligned}
&(n-m)y_{t+m,n-m}\\
=&-(n-m)\log\delta - \log\mathbb{E}_{t+m}\left[\frac{U'(c_A(\mu_{t+n}))}{U'(c_A(\mu_{t+m}))}\frac{1}{\pi_{t+m,t+n}}\right]\\
=&-(n-m)\log\delta - \log\mathbb{E}_{t+m}\exp\left[\log U'(c_A(\mu_{t+n})) - \log U'(c_A(\mu_{t+m})) - \log\pi_{t+m,t+n}\right]
\end{aligned}$$

A quadratic Taylor approximation to the marginal utility of active households gives

$$\log U'\left(c_A(\mu_t)\right) = \log U'\left(c_A(\bar{\mu})\right) - \phi\hat{\mu}_t + \frac{1}{2}\eta\hat{\mu}_t^2 \tag{4.9}$$

where $\hat{\mu}_t = \log\mu_t - \log\bar{\mu}$ is the deviation of the log of money growth from its central value $\bar{\mu}$. Use the quadratic approximation in equation (4.9) to simplify the n-period yield in equation (4.8)

$$ny_{t,n} = -\log\mathbb{E}_t\exp\left[n(\log\delta - \log\bar{\mu}) - \phi(\hat{\mu}_{t+n} - \hat{\mu}_t) - \sum_{i=1}^{n}\hat{\mu}_{t+i} + \frac{1}{2}\eta(\hat{\mu}_{t+n}^2 - \hat{\mu}_t^2)\right], \tag{4.10}$$

Note that we use that the n-period inflation is equal to the sum of the one period inflation rates: $\log\pi_{t,t+n} = \sum_{i=1}^{n}\hat{\mu}_{t+i}$.

In order to solve for the values of $ny_{t,n}$, $my_{t,m}$ and $\mathbb{E}_t rx_{t+m,n-m}$, we need to calibrate the values of δ, $\bar{\mu}$, ϕ and η. More importantly, we have to assume a specific monetary policy rule that drives the process of one period inflation rates $\hat{\mu}_t$. Different monetary policy rules lead to different rates of inflation. We now show how different inflation process $\hat{\mu}_t$ affect

the yields and the risk premium. We first consider the case that the $\hat{\mu}_t$ are i.i.d. $N(0, \sigma^2)$ random variables. Given that inflation rates are well known not to be independent but positively correlated, we subsequently investigate the cases of an MA and AR process. We show that the i.i.d. case for inflation cannot explain the expectations puzzle. But once we turn to the more plausible MA and AR processes for inflation, the dependence induces a risk premium such that $\beta < 1$.[9]

This is the point where our analysis starts to differ from the analysis in Alvarez, Atkeson, and Kehoe (2009), or rather where we extend their model to the case of stochastic processes. Alvarez, Atkeson, and Kehoe (2009) do not have to consider the stochastic process implication. For their purpose, an assumption regarding the distribution of the innovation suffices, since they do not have to calculate a correlation over multiple periods.

4.1 The i.i.d-Assumption for $\hat{\mu}_{t+i}$

First consider the case in which innovations are independently and identically normally distributed. Thus assume that

$$\hat{\mu}_{t+i} = \varepsilon_{t+i}, \text{ with } \varepsilon_{t+i} \sim N(0, \sigma^2) \tag{4.11}$$

Rewriting (4.10) gives

$$n y_{t,n} = \Theta - \log \mathbb{E}_t \exp\left[\frac{\eta}{2}\hat{\mu}_t^2 - (1+\phi)\hat{\mu}_t\right], \tag{4.12}$$

where $\Theta = \log \bar{\mu} - \log \delta + \frac{n-1}{2}\sigma^2$.

To determine the expectation in equation (4.12), note that this is of the form

$$\Pi = \int_{-\infty}^{+\infty} e^{ax+bx^2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{x^2}{\sigma^2}} dx \tag{4.13}$$

To solve this integral, write the power as follows

$$ax + bx^2 - \frac{1}{2}\frac{x^2}{\sigma^2} = -c(x-m)^2, \tag{4.14}$$

where $c = (1/2\sigma^2) - b$ and $m = a/2c$. From the expectation of a non-standard normal

[9] In the text, we only show the main lines of reasoning. The details of derivations are in Appendix B.

random variable, we find

$$\Pi = \frac{1}{\sqrt{1-2b\sigma^2}} exp\left(\frac{a^2\sigma^2}{2-4b\sigma^2}\right) \tag{4.15}$$

so that

$$ny_{t,n} = \Theta + \frac{1}{2}\log(1-\eta\sigma^2) - \frac{(\phi+1)^2\sigma^2}{2(1-\eta\sigma^2)} - \phi\hat{\mu}_t + \frac{\eta}{2}\hat{\mu}_t^2 \tag{4.16}$$

provided that $1-\eta\sigma^2 > 0$. Note that the only stochastic part in $ny_{t,n}$ is $-\phi\hat{\mu}_t + \frac{\eta}{2}\hat{\mu}_t^2$.

A similar expression applies for $my_{t,m}$, so that the $y_{t,n} - y_{t,m}$ part in the covariance of equation (4.7) is readily found as

$$y_{t,n} - y_{t,m} = -\phi\left(\frac{1}{n} - \frac{1}{m}\right)\hat{\mu}_t + \frac{1}{2}\eta\left(\frac{1}{n} - \frac{1}{m}\right)\hat{\mu}_t^2, \tag{4.17}$$

since both yields have identical non-stochastic parts.

The risk premium part in the covariance of (4.7) is from (4.5)

$$\begin{aligned}\mathbb{E}_t rx_{t+m,n-m} =& (n-m)\mathbb{E}_t y_{t+m,n-m} + my_{t,m} - ny_{t,n} \\ =& \frac{1}{2}\log(1-\eta\sigma^2) - \frac{1-\eta}{2}\sigma^2 - \frac{(\phi+1)^2\sigma^2}{2(1-\eta\sigma^2)},\end{aligned} \tag{4.18}$$

which contains no random elements since these cancel from $my_{t,m} - ny_{t,n}$. Hence the covariance in equation (4.7) is zero and i.i.d. fluctuations in inflation cannot explain the expectations puzzle. This is summarized in the following.

> *Proposition* 1. If inflation is an i.i.d. normally distributed random variable, the term structure of interest rates conforms to the expectations hypothesis. As a result, the values of the Campbell-Shiller regression coefficients are equal to unity.

4.2 The MA-assumption for $\hat{\mu}_t$

Next we introduce some dependence in the inflation process. Suppose the inflation process is MA(1)

$$\hat{\mu}_{t+1} = \varepsilon_{t+1} + \theta\varepsilon_t, \qquad \varepsilon_t \sim N(0,\sigma^2) \text{ i.i.d.} \tag{4.19}$$

To explain how this changes the regression coefficient, consider the simple 2-period case with $n = 2$, $m = 1$.[10] Hence,

$$Cov(\mathbb{E}_t rx_{t+m,n-m}, y_{t,n} - y_{t,m}) = Cov(y_{t,1} + \mathbb{E}_t y_{t+1,1} - 2y_{t,2}, y_{t,n} - y_{t,m}) \tag{4.20}$$

To determine $y_{t,1}$, note that using the MA(1) scheme (4.19) gives

$$\begin{aligned} y_{t,1} =& -\log \mathbb{E}_t \text{exp}\left[\log\delta - \log\bar{\mu} - \phi\hat{\mu}_{t+1} - \hat{\mu}_{t+1} + \frac{1}{2}\eta\hat{\mu}_{t+1}^2 + \phi\hat{\mu}_t - \frac{1}{2}\eta\hat{\mu}_t^2\right] \\ =& \log\bar{\mu} - \log\delta + (\phi+1)\theta\varepsilon_t - \phi\hat{\mu}_t + \frac{1}{2}\eta\hat{\mu}_t^2 - \frac{1}{2}\eta\theta^2\varepsilon_t^2 \\ & - \log \mathbb{E}_t \text{exp}\left[(\eta\theta\varepsilon_t - \phi - 1)\varepsilon_{t+1} + \frac{1}{2}\eta\varepsilon_{t+1}^2\right] \end{aligned} \tag{4.21}$$

Using the same reasoning as we used in determining the integral Π from (4.13), we get

$$\begin{aligned} y_{t,1} =& \log\bar{\mu} - \log\delta + (\phi+1)\theta\varepsilon_t - \phi\hat{\mu}_t + \frac{1}{2}\eta\hat{\mu}_t^2 - \frac{1}{2}\eta\theta^2\varepsilon_t^2 - \frac{(\eta\theta\varepsilon_t - \phi - 1)^2\sigma^2}{2(1-\eta\sigma^2)} \\ & + \frac{1}{2}\log(1-\eta\sigma^2) \end{aligned} \tag{4.22}$$

under the same condition as before. Following the above approach one shows that

$$\begin{aligned} \mathbb{E}_t[y_{t+1,1}] =& \log\bar{\mu} - \log\delta - \phi\theta\varepsilon_t + \frac{1}{2}\eta\theta^2\varepsilon_t^2 + \frac{1}{2}\eta(1-\theta^2)\sigma^2 - \frac{(\phi+1)^2\sigma^2 + \eta^2\theta^2\sigma^4}{2(1-\eta\sigma^2)} \\ & + \frac{1}{2}\log(1-\eta\sigma^2) \end{aligned} \tag{4.23}$$

and

$$y_{t,2} = \log\bar{\mu} - \log\delta - \frac{1}{2}\phi\hat{\mu}_t + \frac{1}{4}\eta\hat{\mu}_t^2 + \frac{1}{2}\theta\varepsilon_t - c^T V^T c + \frac{1}{4}\log\left(1 - \eta\sigma^2 - \eta\sigma^2\theta^2\right), \tag{4.24}$$

where $y_{t,2}$ in (4.24) requires $1-\eta\sigma^2(1+\theta^2) > 0$. The $c^T V^T c$ part originates from the double integral involving ε_{t+1} and ε_{t+2}, but is constant. Hence the $c^T V^T c$ part plays no role in the covariance.

For the covariance (4.7), we only need to consider the stochastic parts from (4.22), (4.23) and (4.24). Doing this gives the following simplified expressions:

$$\mathbb{E}_t\widehat{rx}_{t+1,1} = \frac{2\eta\theta(\phi+1)\sigma^2\varepsilon_t - \eta^2\theta^2\sigma^2\varepsilon_t^2}{2(1-\eta\sigma^2)} + \Lambda \tag{4.25}$$

[10] In Appendix B, we provide a complete proof for the general case of n and m.

$$\hat{y}_{t,2} - \hat{y}_{t,1} = \frac{\phi}{2}\hat{\mu}_t - \frac{1}{4}\eta\hat{\mu}_t^2 - \left(\phi + \frac{1}{2}\right)\theta\varepsilon_t + \frac{1}{2}\eta\theta^2\varepsilon_t^2 + \frac{\eta^2\theta^2\sigma^2\varepsilon_t^2 - 2\eta\theta(\phi+1)\sigma^2\varepsilon_t}{2(1-\eta\sigma^2)} + \Delta \tag{4.26}$$

where Λ and Δ stand for the constant parts. Since ε_t and ε_t^2 are both part of $\mathbb{E}_t\widehat{rx}_{t+1,1}$ in (4.25) and $\hat{y}_{t,2} - \hat{y}_{t,1}$ in (4.26), it follows that the covariance between the two parts is non-zero. Therefore we get that

$$\frac{Cov(\mathbb{E}_t rx_{t+1,1}, y_{t,2} - y_{t,1})}{Var(y_{t,2} - y_{t,1})} = \frac{Cov(\mathbb{E}_t\widehat{rx}_{t+1,1}, \hat{y}_{t,2} - \hat{y}_{t,1})}{Var(\hat{y}_{t,2} - \hat{y}_{t,1})} \neq 0.$$

Hence the regression coefficient $\hat{\beta}$ in (4.7) will differ from unity.

For the results of longer horizons, see the Appendix. In summary, we find, with the assumption that $\hat{\mu}_t$ is a MA(1) process, the slope coefficient $\beta_{n,(m=1)}$ in regression (4.3) deviates from 1. For any of the cases $m \geq 2$, the slope coefficient $\beta_{n,(m\geq 2)}$ is still equal to unity, however.[11] We capture this result in Proposition 2.

> *Proposition* 2. If the inflation is an MA(1) process, the values of the Campbell-Shiller regression coefficients are no longer equal to unity for $m = 1$. If m is equal or larger than two, the values of the Campbell-Shiller regression coefficients are equal to unity.

The intuition for this result follows from the feature of the MA(1) process. We know that the expected excess returns $\mathbb{E}_t rx_{t+m,n-m}$, i.e. the risk premium, depends on the expected aggregate inflation over the period that starts at $t+m$ and ends at the maturity date $t+n$. If $\mathbb{E}_t\hat{\mu}_{t+1} = \theta\varepsilon_t$ and $\mathbb{E}_t\hat{\mu}_{t+i} = 0, i \geq 2$, the expected one period ahead inflation becomes $\mathbb{E}_t\mu_{t+1} = \bar{\mu} + \theta\varepsilon_t$ and $\mathbb{E}_t\mu_{t+i} = \bar{\mu}$, for $i \geq 2$. So the expected one period ahead excess return $\mathbb{E}_t rx_{t+1,n-1}$ becomes time varying, because $\mathbb{E}_t\hat{\mu}_{t+1} \neq \mathbb{E}_{t'}\hat{\mu}_{t'+1}$, as long as $\varepsilon_t \neq \varepsilon_{t'}$. The $\mathbb{E}_t rx_{t+1,n-1}$ and $y_{t,n} - y_{t,1}$ depend on the state of the starting date "t", and both contain the current innovation ε_t. From an unconditional perspective the two terms are therefore correlated with each other, so that $\beta_{n,1} \neq 1$. However, for $m \geq 2$, the expected inflations from period $t+m$ to $t+n$ are constants ($\mathbb{E}_t\mu_{t+i} = \bar{\mu}$, for $i \geq 2$), hence the expected excess return $\mathbb{E}_t rx_{t+m,n-m}$ is a constant and independent of the starting date t. The $\mathbb{E}_t rx_{t+m,n-m}$ is therefore not correlated with $y_{t,n} - y_{t,1}$, and hence $\beta_{n,(m\geq 2)} = 1$.

[11] For the derivations, see Appendix B

In order to provide values of $\beta_{n,1}$ under the MA(1)-inflation, we need to calibrate the parameter values. The parameters we need are η, ϕ, θ and σ. Since ϕ and η are directly unobservable, the starting values of ϕ and η that we use are borrowed from Alvarez, Atkeson, and Kehoe (2009).[12] Table 2 shows values of $\beta_{n,1}$ under different values of θ (from 0.1-0.9) with σ equal to 0.0033 and 0.0115 respectively.[13] The value of 0.0033 is the average sample standard deviation for monthly inflation for the countries used in Section 1 from Feb 1990 to Dec 1999 and the value of 0.0115 is the average sample standard deviation for yearly inflation in these countries during the same time period. One sees from Table 2 that if the volatility is small ($\sigma = 0.0033$), the values of $\beta_{n,1}$ are around unity with tiny deviations, no matter what the value of θ is. So the expectations hypothesis holds for a low volatility process. While an increase in money growth volatility ($\sigma = 0.0115$) also increases deviations of $\beta_{n,1}$ from unity, these deviations are still too small to match the data. The results, however, do show deviations of $\beta_{n,1}$ take two directions. The $\beta_{n,1}$ first increases with θ and peaks at $\theta = 0.4$ and then decreases with θ.

In summary, the assumption of an MA(1) process for inflation implies deviations of $\beta_{n,1}$ that are too small in comparison with the deviations in the data. More importantly, the MA(1)-assumption for inflation can at most account for the deviation from unity of $\beta_{n,1}$. It is unable to account for the deviation of any $\beta_{n,m}$ with $m \geq 2$.

Even though the MA(1) process is unable to match the data, the analysis provides the theoretical intuition for deviations of $\hat{\beta}$ from unity. Continuing along these lines, one shows that an MA(2) process can explain deviations from of $\beta_{n,1}$ and $\beta_{n,2}$. Because the data reveal deviations at higher orders and because of the magnitude mismatch, we turn to AR processes.

4.3 The AR(1)-assumption for $\hat{\mu}_t$

Consider the case when $\hat{\mu}_t$ is an AR(1) process, i.e. $\hat{\mu}_{t+1} = \rho\hat{\mu}_t + \varepsilon_{t+1}$. Since the expectations hypothesis is also rejected at the real level, we check the behavior of both nominal yields and real yields under AR(1) inflation.[14] Since higher order AR process are

[12] Alvarez, Atkeson, and Kehoe (2009) find the values of ϕ and η by solving (3.8) and (3.9) with the assumption that the endowment is 1.

[13] When choosing the values for σ and η, one constraint is that $\eta\sigma^2 < 1$. See Appendix B.2 for an explanation.

[14] All the proofs and calculations are shown in Appendix B. The parameters we need for the calculation are σ ρ, ϕ and η. In the case of MA(1)-inflation, we drop the effects of ϕ and η for simplicity. Here, in addition to σ and ρ, we also discuss the effect of ϕ. It is clear that η is also important. However, according to the

analytically intractable, we resort to numerical procedure for analyzing these cases. But we do provide explicit expression for the AR(1) case.

Table 3 gives the values of slope coefficients and shows how these values vary with σ, ρ and ϕ. $Panel\ A$ gives the values of slope coefficient for nominal yields and $Panel\ B$ gives the values of slope coefficient for real yields. The coefficients for both the nominal and the real yields have similar trends.

Low money growth volatility ($\sigma = 0.0033$) generates slope coefficients around 1. With larger volatility ($\sigma = 0.0115$ or 0.0180), deviations of $\beta_{2,1}$ (and $\beta^*_{2,1}$) from 1 increase. At the lower value of volatility ($\sigma = 0.0033$), deviations of $\beta_{2,1}$ (and $\beta^*_{2,1}$) are only upward, i.e. $\beta_{2,1}$ (and $\beta^*_{2,1}$) > 1. These deviations increase with ρ and ϕ. At the higher values of volatility ($\sigma = 0.0115$ or 0.0180), the deviations start to take two directions. At the lower values of ρ, $\beta_{2,1}$ (and $\beta^*_{2,1}$) are biased upwards. With higher values of autocorrelation ρ, the $\beta_{2,1}$ (and $\beta^*_{2,1}$) falls below unity and can even become negative. If σ takes the higher values, the effect of ϕ is ambiguous . When $\rho \leq 0.4$, increasing the value of the elasticity ϕ may strengthen the deviation, however, when $\rho \geq 0.5$, the effect is reversed.

The relation between slope coefficients and parameters is thus quite complex. One clear message from this table is that large deviations for β from unity occur only if the variation of the risk premium is sizeable.

Intuitively, the greater the volatility, the higher the risk and hence the risk premium is larger. A larger value of ρ means the effect of a shock persists longer, so the risk is also greater. As a result, a higher risk premium is required. Recall that ϕ is the elasticity of the marginal utility of active households to the change in money growth. Intuitively, the larger the value of ϕ is, the more sensitive the active households are to a change in the money growth rate and hence they ask for higher risk premia. This intuition becomes clear, if we consider a specific functional form for the preference of the agents. In the case of constant relative risk aversion preferences are $U(c) = c^{1-\tau}/(1-\tau)$ and ϕ takes the form

$$\phi = \tau \frac{d \log c_A(\mu)}{d \log \mu} \tag{4.27}$$

Thus ϕ is proportional to the risk aversion coefficient τ. Therefore, the more the households are averse to risk, the larger the value of ϕ will be and the higher the risk premium

definition of η and ϕ, i.e. $\eta = -\partial\phi/\partial\mu$, the economic implications of η should be covered by ϕ, so we drop the discussion for η and only discuss the effects of σ ρ, and ϕ for the value of β.

will be.

So the size of the risk premium is proportional to the values of σ, ρ and ϕ. Large risk premia guarantee that the variation of risk premia is larger and hence the deviation of β from unity may also be large. But this does not mean that the deviation is linearly proportional to the size of the risk premia. Equation (4.7) tells us that the value of β depends on the covariance of the time-varying risk premia and the spread $y_{t,n} - y_{t,m}$, and the variance of $y_{t,n} - y_{t,m}$. The sign of the covariance decides the direction of the deviation, while the ratio of the covariance and the variance decides the magnitude of the deviation. We cannot tell exactly how the variation of the risk premia affects the direction and the magnitude of the deviation. According to Table 3, one thing is clear that to get a negative β, both σ and ρ must be large enough. The 3D plots of Figure 1 and Figure 2 show how the regression slope coefficients for the nominal and the real yields vary with the σ and ρ. The plots clearly show that as the values of σ and ρ increase, the values of $\beta_{2,1}$ and $\beta^*_{2,1}$ first increase slightly and then decrease rapidly and turn negative.

The above analysis is based on the explicit analytical results for $\beta_{2,1}$ ($\beta^*_{2,1}$) and AR(1). For $\beta_{n,m}$ in general the analytical solutions become unwieldy, but can be easily analyzed with simulated data. In the next section, we will use numerical methods to generate the $\beta_{n,m}$ and $\beta^*_{n,m}$ with $n \geq 3$ and $m \geq 2$ to check our intuition.

5 Numerical Analysis

In this section, we use simulated yield data to validate the analytical results for the cases that inflation is an i.i.d. process or follows an MA(1) process. For the AR(1) process the analytical solutions become unwieldy, but can be easily tracked numerically. By varying the parameter values we can investigate the conditions under which the $\hat{\beta}$ is considerably less than one or is even negative.

Expression for the bond yields are given in (4.12), (B.6) and (B.9) and (B.21) for the case that inflation is an i.i.d. process, a MA(1) or an AR(1) process, respectively. Based on these yield equations, we can simulate the bond yield for any length of period and maturity. The data is generated under the assumption that money growth follows an i.i.d variable, MA(1) and AR(1) process respectively. For each specific process, we conducted 200 simulations and each of these simulations run for 200 time periods.

The calibration of the parameters is as in Section 4. We choose $\bar{\mu} = 1.05$, which says that the mean of the one-period inflation rate is 5%. The scale of one period is a year.

5.1 Money growth is i.i.d. variable

Table 4 shows the average values of β for nominal yields over 200 simulations based on the assumption that inflation is an i.i.d. variable. Figure 3 and Figure 4 plot the values of $\beta_{10,1}$ and $\beta_{10,2}$ for the 200 simulations. The results show that the average values of simulated β are quite close to 1. This is consistent with the theoretical results in case inflation is an i.i.d. random variable. In that case the term structure of interest rates conforms to the expectations hypothesis and $\hat{\beta}$ is expected to be 1.

5.2 Money growth follows a MA(1) process

Table 5 shows the average values of β for nominal yields over 200 simulations based on the assumption that inflation follows a MA(1) process. Figure 5 and Figure 6 plot the values of $\beta_{10,1}$ and $\beta_{10,2}$ for the 200 simulations.

We showed that if money growth follows a MA(1) process, the regression coefficient $\beta_{n,1}$ can deviate from unity due to the time-varying risk. The magnitude of deviation depends on the values of σ, θ and ϕ. The simulation results of $\beta_{n,1}$ are consistent with the theoretical results. Both show a tiny but clearly discernable deviation from unity. The MA(1) inflation cannot, however, account for large deviations of $\beta_{1,n}$ from unity for given variations of the input parameters. The determinant for the time-varying risk is the aggregate expected inflation from period $t+1$ to $t+n$. If inflation is a MA(1) process, the expected inflation is only time varying in period $t+1$, since a shock dies out after two periods. In comparison to the AR(1) case, the MA(1) case can only generate small variations in risk. As a result, the deviation of the regression coefficients from unity is then smaller. Theoretically, the value of any $\beta_{n,m}$ with $m \geq 2$ is 1 and the average values of the simulated $\beta_{n,2}$ are closer to 1 than those of $\beta_{n,1}$.

5.3 Money growth follows an AR(1) process

In Section 4, we prove that the expectations hypothesis may not hold if inflation follows an AR(1) process. In the case of AR(1) inflation, the theoretical values of $\beta_{2,1}$ for both

the nominal and the real yields are close in magnitude to those in the data. We do not calculate the theoretical values of $\beta_{n,m}$ and $\beta^*_{n,m}$ with $n \geq 3$ and $m \geq 2$ due to the complexity of the calculations. However, it is intuitive that the rejection of EH is not limited to $n = 2$ and $m = 1$ in the case of AR(1) inflation. We now simulate the nominal and the real yields with the yield equations (B.21) and (B.39) and run regressions (4.3) with the simulated yields to obtain values for the slope coefficients. In order to check the difference of slope coefficients in the case of lower and higher risk, we use two combinations of values for σ and ρ, with $\sigma = 0.0080$ and $\rho = 0.6$ as the lower values and $\sigma = 0.0115$ and $\rho = 0.9$ as the higher values.

Table 6 shows the average values of $\beta_{n,m}$ and $\beta^*_{n,m}$ from 200 simulations. $Panel\ A$ shows the values of $\bar{\beta}_{n,m}$ and $Panel\ B$ shows the values of $\bar{\beta}^*_{n,m}$. The results are consistent with the theoretical calculations. The volume of the risk may change the direction of the rejection of the EH. When both σ and ρ or either one of these takes the lower value, the risk is lower and the average values of $\beta_{n,m}$ are larger than unity. When both σ and ρ take on higher values, the risk is higher and the average values of $\beta_{n,m}$ are less than unity and can even turn negative. The magnitude of the deviation from the expectations hypothesis matches the magnitude observed in the data. In the case of higher risk, we observe that β is decreasing in maturity n.

Figure 7 and Figure 8 plot $\beta_{10,1}$ and $\beta_{10,2}$ for all the 200 simulations under $\sigma = 0.0115$ and $\rho = 0.9$. Due to the larger variation in risk, the simulated $\hat{\beta}$ is more dispersed under AR(1)-inflation. If we suppose that each single simulation represents a specific economy, we see that the yield regression coefficients may be quite different for different economies, even if these share the same inflation processes. Nevertheless, the average values for $\beta_{10,1}$ and $\beta_{10,2}$ are -0.3352 and -0.2235 and most of the simulations give negative $\beta_{10,1}$ and $\beta_{10,2}$ values. The dispersion is large since the maximum and minimum values for $\beta_{10,1}$ and $\beta_{10,2}$ are approximately 2 and -2. Figure 9 and Figure 10 plot $\beta^*_{10,1}$ and $\beta^*_{10,2}$ for all the 200 simulations under $\sigma = 0.0115$ and $\rho = 0.9$. The plots of $\beta^*_{10,1}$ and $\beta^*_{10,2}$ show similar distributions for $\beta_{10,1}$ and $\beta_{10,2}$, but with lower mean values. The largest and least values for $\beta^*_{10,1}$ and $\beta^*_{10,2}$ are around 0.5 and -3 respectively.

Next, we compare the yield curves that we simulated with the real data. Figure 11 shows the average Euro-rate curves for 5 countries with maturities up to 60 months from

January 1995 to December 1999.[15] We see that all the yield curves slope upward with yields increasing with maturity.

Figure 12 shows the average nominal yield curves we simulated with different values for σ and ρ. The figure also shows the yield curve based on the i.i.d.-inflation that is used as a control to show the term structure with constant risk premium under which the expectations hypothesis holds. Like in the real data, all the simulated yield curves have typical upward slopes. It appears that the yield curve based on the i.i.d.-inflation shows the largest curvature and the yield curves based on the inflation with large value of the autoregressive coefficient (L,H and H,H) are relatively flatter. Figure 13 shows the average real yield curves. The real yield curves show similar patterns to the nominal ones, but are about 5% lower than the nominal yields. The reason is that we choose 5% as the mean of annual inflation.

Combining both the analytical and numerical results, we have the following result for AR(1)-inflation in Proposition 3.

> *Proposition* 3. If inflation is an AR(1) process, the values of the Campbell-Shiller regression coefficients can differ considerably from unity. The values can be higher or lower than unity and can be even negative. The values depend on the autocorrelation and volatility of inflation.

5.4 Robustness check

In Section 4, we solve a consumption-based asset pricing model and obtain the term structure of interest rates in the case that inflation is an i.i.d. random variable, or follows an MA(1) or an AR(1) processes. Obviously, the i.i.d. variable, MA(1) and AR(1) processes cannot represent all types of inflation in reality. Besides the AR(1) process, the AR(4) and the AR(12) processes are often applied to quarterly and monthly inflation. Cecchetti and Debelle (2006) indicate "the AR coefficient is often close to one in a large number of countries when estimated on inflation data over the past twenty years." In order to check whether the model has the power to explain the expectations puzzle if inflation is more like an AR(4) or AR(12) process, we also simulate the nominal and the real yields based on equations (4.10) and (B.38). According to the law of large numbers, the expected values in

[15] Only 5 of the countries from Table 4.3 have yields with maturities up to 60 months.

equations (4.10) and (B.38) can be approximated by the average values of a large number of random values in the brackets of equations (4.10) and (B.38). In order to ensure that the numbers of random values are large enough, so that the averages of these random values converge to the means of the random values, we calculate 200000 times the random values in the brackets of equations (4.10) and (B.38) and then take the average of these values as the approximation of the expected values.

The results are given in Table 7 and Table 8. The Tables show that the yields simulated based on AR(4) or AR(12) inflation have negative regression coefficients for most of maturities. These results indicate that the model that we use is robust to other AR processes for inflation in accounting for the rejection of the EH. Actually, with equations (4.10) and (B.38) and reliance on the law of large numbers, we can simulate the nominal and the real yields based on any inflation process, not only the AR processes, but also more complicated processes. According to our analysis, these more complicated processes are promising for generating yields with term structures that reject EH as long as they can provide large and persistent variation in the risk premium. But the AR(1) assumption for inflation suffices to explain the expectations puzzle and the simulated yield curves based on AR(1) inflation also match the data.

6 Empirical Test

Both the theoretical and numerical results in this paper show that when either the volatility (σ) or the autoregressive coefficient for AR(1) inflation (ρ) or both take on small values, the yield regression coefficients are larger than unity, while if both have large values, the yield regression coefficients are smaller than unity and may even be negative. If this holds in the data, we should expect that when we cross sectionally regress $\hat{\beta}$ onto $\hat{\sigma}$ and $\hat{\rho}$, the coefficients for σ and ρ should at least not both be positive. If both coefficients are negative, this would provide clear empirical support for the inverse relation between $\hat{\beta}$ and $(\hat{\sigma}, \hat{\rho})$. In this section, we conduct such regressions as an indirect test of the above theoretically deduced relation between $\hat{\beta}$ on the one hand and σ and ρ on the other hand, so as to see whether our theory matches the data in this subtle detail.

We collected monthly inflation rates for 15 countries from Feb 1990 to Dec 1999.[16]

[16]The 15 countries are the same as those listed in Table 1, with the exception of Aus and Nzl. Australia and New Zealand were excluded because the monthly inflation rates for these two countries are unavailable.

We fitted an AR(1) process to obtain the sample autoregression coefficient ($\hat{\rho}$) for each country.[17] The time period we use for inflation is 5 years prior to the time period that we use for the interest rates. This is done on the grounds that investors look back to determine their expectations regarding future inflation.

Table 9 shows the values of the sample autoregression coefficient ($\hat{\rho}$) and the sample standard deviation ($\hat{\sigma}$). The values of $\hat{\rho}$ for Belgium and Norway are very small ($\hat{\rho} < 0.1$) and the probabilities for $\hat{\rho} = 0$ are higher than 40%. This means that the goodness-of-fit of the two countries is very low. We decided that is inappropriate to model the inflation processes of Belgium and Norway as an AR(1) process, so we exclude the two observations in the subsequent regressions.

We regress $\beta_{n,m}$ onto $\hat{\rho}$ and $\hat{\sigma}$ cross sectionally over the different countries for given n, m as in

$$\beta_{n,m} = const + b_{n,m}\hat{\rho} + d_{n,m}\hat{\sigma} + error \tag{6.1}$$

The regression results are shown in Table 10. We can see that almost all the regression coefficients are negative, except $d_{6,3}$ and $d_{12,3}$. Most of the regression coefficients are not statistically significant, though, only $b_{3,1}$ and $d_{3,1}$ are statistically significant. The F-tests nevertheless say that the R^2 for many regressions are not low. Taking the extremely small sample sizes into account (only 11-13 observations), the results seem not disappointing. It is reasonable to believe that the statistical significance would improve if we had more observations for the regressions. So we conclude that the negative coefficients demonstrate, to some extent, the negative correlations between β and ρ and σ in the data. At least, the evidence does not go in the other direction.

7 Conclusions

We build on the endogenously segmented market model of Alvarez, Atkeson, and Kehoe (2009) to explain the expectations puzzle by introducing autocorrelated inflation. We formulate a consumption-based asset pricing model in which the risk premium for both the nominal and the real yields can vary in response to expected changes in inflation. In the theoretical part, we analyze three types of inflation, i.e. inflation is an i.i.d. random

[17] Obviously, AR(1) is not suitable for all inflation process, but we can not find a process which fits all inflations equally well. However, in order to be consistent with our theoretical analysis, we only use the AR(1) process to fit all inflations here.

variable, or follows an MA(1) or an AR(1) process. The analytical solutions show that the i.i.d-process for inflation cannot solve the expectations puzzle because the risk premium is constant under the i.i.d-inflation. If inflation follows an MA(1) process, the risk premia become time varying. Due to the rather small variation in risk premia, the MA(1) process for inflation can only generate tiny deviation of β from unity. Only the AR(1) process for inflation can generate enough variation in risk premia to account for the rejection of the expectations hypothesis. The numerical results show that the rejection of EH is robust to the AR(4) and AR(12) processes for inflation.

Our empirical tests for the EH with the Euro-rates of 17 countries show that the rejection of the EH goes in two directions. For some countries, such as Australia, we get negative regression coefficients. For other countries, such as France, the regression coefficients are significantly larger than unity. This phenomenon can be addressed by the segmented market model. Both the theoretical and the numerical results show that the regression coefficients would first increase ($\beta > 1$) and then decrease ($\beta < 1$) following an increase of the risk premium.

The result that the negative regression coefficients appear only if inflation has a large volatility and a high autoregressive coefficient imply that there exists a negative relation between β, σ and ρ. To test this prediction we regressed $\beta_{n,m}$ for 13 countries onto the sample standard deviations and the sample autoregressive coefficients for their inflations. The regression coefficients for most of the $\beta_{n,m}$ are negative. Even though most of the coefficients are not statistically significant due to the extremely small sample size, the results do not run counter to the predictions. According to our segmented market model, the simulated yields curves have the typical upward sloping term structure for both the nominal and the real yields and their magnitudes match the data. All of these results show that the endogenously segmented market model provides a reliable framework for exploring the underlying mechanisms that determine the character of the term structure.

The model also has its limitations. The only source for time-varying risk in this model is inflation. Duffee (2011) and Chernov and Mueller (2012) find there may be other latent factor(s) hidden in the yield curve besides inflation. Bansal and Shaliastovich (2012) and Rudebusch and Wu (2008) show that both output and inflation are sources of time-vary risk in the term structure of interest rates. To keep the analysis simple, the endogenously segmented market model that we use adopts an endowment economy with constant peri-

od by period income, so we do not investigate the effect of variation in output. The effect of output on the term structure of interest rates seems feasible within this model, and is of interest for future work. While it is not easy to reconcile all the factors in one model, this should be a subject of future research concerning the term structure of interest rates.

APPENDIX

A Proofs and derivations of section 2

A.1 The derivation of equation (3.9)

The Lagrangian is given by

$$\begin{aligned} L =& U(c_A(\mu))F(\bar{\gamma}(\mu),\mu) + U\left(y/\mu\right)\left[1 - F(\bar{\gamma}(\mu),\mu)\right] \\ &+ \lambda\left\{y - c_A(\mu)F(\bar{\gamma}(\mu),\mu) - \int_0^{\bar{\gamma}(\mu)} \gamma f(\gamma,\mu)d\gamma - \left(y/\mu\right)\left[1 - F(\bar{\gamma}(\mu),\mu)\right]\right\} \end{aligned} \tag{A.1}$$

The relevant F.O.C. are

$$\frac{\partial L}{\partial c_A(\mu)} = 0 : U'(c_A(\mu)) = \lambda \tag{A.2}$$

$$\frac{\partial L}{\partial \bar{\gamma}(\mu)} = 0 : U(c_A(\mu)) - U(y/\mu) + \lambda\left[(y/\mu) - \bar{\gamma}(\mu) - c_A(\mu)\right] = 0 \tag{A.3}$$

Plug (A.2) into (A.3), to derive (3.9). Q.E.D.

When (3.9) is rewritten as (A.4), we can see that the marginal utility of active households is equal to the ratio of the utility difference and the consumption difference minus the cost.

$$U'(c_A(\mu)) = \frac{U(c_A(\mu)) - U(y/\mu)}{c_A(\mu) - (y/\mu) - \bar{\gamma}(\mu)} \tag{A.4}$$

B Proofs and derivations of section 3

B.1 Derivation of the slope coefficient $\beta_{n,m}$

Note

$$\begin{aligned} Cov(x,y) =& Cov(x, y-x) + Cov(x,x) \\ =& Cov(x, y-x) + Var(x) \end{aligned}$$

$$Cov(x+z, y) = Cov(x,y) + Cov(z,y)$$

$$Cov(kx, y) = kCov(x,y)$$

$$Cov(x,y) = Cov(y,x)$$

$$\begin{aligned} y_i =& \beta_0 + \beta_i x_i + \varepsilon \\ \beta_i =& \frac{\Sigma_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\Sigma_{i=1}^n (x_i - \bar{x})^2} = \frac{Cov(x,y)}{Var(x)} \end{aligned}$$

In our case,

$$\begin{aligned} x =& \frac{m}{n-m}(y_{t,n} - y_{t,m}) \\ y =& y_{t+m,n-m} - y_{t,n} \\ Var(x) =& \frac{m^2}{(n-m)^2} Var(y_{t,n} - y_{t,m}) \end{aligned}$$

$$\begin{aligned} \beta_{n,m} =& \frac{Cov(y_{t+m,n-m} - y_{t,n}, \frac{m}{(n-m)}(y_{t,n} - y_{t,m}))}{\frac{m^2}{(n-m)^2} Var(y_{t,n} - y_{t,m})} \\ =& \frac{Cov((n-m)(y_{t+m,n-m} - y_{t,n}), m(y_{t,n} - y_{t,m}))}{m^2 Var(y_{t,n} - y_{t,m})} \\ =& \frac{Cov((n-m)y_{t+m,n-m} - ny_{t,n} + my_{t,m} + m(y_{t,n} - y_{t,m}), m(y_{t,n} - y_{t,m}))}{m^2 Var(y_{t,n} - y_{t,m})} \\ =& \frac{Cov(\mathbb{E}_t rx_{t+m,n-m}, y_{t,n} - y_{t,m})}{mVar(y_{t,n} - y_{t,m})} + 1 \qquad \text{Q.E.D.} \end{aligned}$$

B.2 Derivation of the results under the i.i.d-Assumption for $\hat{\mu}_{t+i}$

With the quadratic approximation and the definition of the pricing kernel, we have:

$$\begin{aligned} ny_{t,n} =& -\log \mathbb{E}_t \text{exp}\left[n(\log\delta - \log\bar{\mu}) - (\phi+1)\hat{\mu}_{t+n} - \sum_{i=1}^{n-1} \hat{\mu}_{t+i} + \frac{1}{2}\eta\hat{\mu}_{t+n}^2 + \phi\hat{\mu}_t - \frac{1}{2}\eta\hat{\mu}_t^2 \right] \\ =& n(\log\bar{\mu} - \log\delta) - \phi\hat{\mu}_t + \frac{1}{2}\eta\hat{\mu}_t^2 - \log \mathbb{E}_t \text{exp}\left[-\sum_{i=1}^{n-1} \hat{\mu}_{t+i} - (\phi+1)\hat{\mu}_{t+n} + \frac{1}{2}\eta\hat{\mu}_{t+n}^2 \right] \\ =& n(\log\bar{\mu} - \log\delta) - \phi\hat{\mu}_t + \frac{\eta}{2}\hat{\mu}_t^2 + \frac{n-1}{2}\sigma^2 - \log \mathbb{E}_t \text{exp}\left[-(\phi+1)\varepsilon_{t+n} + \frac{\eta}{2}\varepsilon_{t+n}^2 \right] \end{aligned}$$

It's known that if x is normally distributed with mean zero and variance σ^2 and satisfies $1 - 2b\sigma^2 > 0$, then

$$\mathbb{E}\text{exp}\left(ax + bx^2\right) = \text{exp}\left(\frac{1}{2}\frac{a^2\sigma^2}{1-2b\sigma^2}\right)\left(\frac{1}{1-2b\sigma^2}\right)^{\frac{1}{2}}$$

In our case, $x = \varepsilon_{t+n}$, $a = -(\phi+1)$ and $b = \frac{\eta}{2}$.

Hence, under the assumption $1 - \eta\sigma^2 > 0$

$$\log \mathbb{E}_t \exp\left[-(\phi+1)\varepsilon_{t+n} + \frac{\eta}{2}\varepsilon_{t+n}^2\right] = \frac{(\phi+1)^2\sigma^2}{2(1-\eta\sigma^2)} - \frac{1}{2}\log(1-\eta\sigma^2)$$

So

$$ny_{t,n} = n(\log\bar{\mu} - \log\delta) - \phi\hat{\mu}_t + \frac{\eta}{2}\hat{\mu}_t^2 + \frac{n-1}{2}\sigma^2 - \frac{(\phi+1)^2\sigma^2}{2(1-\eta\sigma^2)} + \frac{1}{2}\log(1-\eta\sigma^2) \tag{B.1}$$

With a similar derivation, we get

$$my_{t,m} = m(\log\bar{\mu} - \log\delta) - \phi\hat{\mu}_t + \frac{\eta}{2}\hat{\mu}_t^2 + \frac{m-1}{2}\sigma^2 - \frac{(\phi+1)^2\sigma^2}{2(1-\eta\sigma^2)} + \frac{1}{2}\log(1-\eta\sigma^2) \tag{B.2}$$

$$\begin{aligned}
&(n-m)y_{t+m,n-m} \\
=&-\log\mathbb{E}_{t+m}\exp\left[(n-m)(\log\delta - \log\bar{\mu}) - (\phi+1)\hat{\mu}_{t+n} - \sum_{i=1}^{n-m-1}\hat{\mu}_{t+m+i} + \frac{1}{2}\eta\hat{\mu}_{t+n}^2 + \phi\hat{\mu}_{t+m} - \frac{1}{2}\eta\hat{\mu}_{t+m}^2\right] \\
=&(n-m)(\log\bar{\mu} - \log\delta) - \phi\hat{\mu}_{t+m} + \frac{1}{2}\eta\hat{\mu}_{t+m}^2 - \log\mathbb{E}_{t+m}\exp\left[-(\phi+1)\hat{\mu}_{t+n} - \sum_{i=1}^{n-m-1}\hat{\mu}_{t+m+i} + \frac{1}{2}\eta\hat{\mu}_{t+n}^2\right] \\
=&(n-m)(\log\bar{\mu} - \log\delta) - \phi\hat{\mu}_{t+m} + \frac{1}{2}\eta\hat{\mu}_{t+m}^2 + \frac{n-m-1}{2}\sigma^2 - \frac{(\phi+1)^2\sigma^2}{2(1-\eta\sigma^2)} + \frac{1}{2}\log(1-\eta\sigma^2)
\end{aligned}$$

So

$$\begin{aligned}
\mathbb{E}_t rx_{t+m,n-m} =& \mathbb{E}_t[(n-m)y_{t+m,n-m} + my_{t,m} - ny_{t,n}] \\
=&(n-m)\mathbb{E}_t y_{t+m,n-m} + my_{t,m} - ny_{t,n}
\end{aligned}$$

$$my_{t,m} - ny_{t,n} = -(n-m)(\log\bar{\mu} - \log\delta) + \frac{n-m}{2}\sigma^2$$

$$\begin{aligned}
(n-m)\mathbb{E}_t y_{t+m,n-m} =&(n-m)(\log\bar{\mu} - \log\delta) + \frac{\eta+n-m-1}{2}\sigma^2 - \frac{(\phi+1)^2\sigma^2}{2(1-\eta\sigma^2)} \\
&+ \frac{1}{2}\log(1-\eta\sigma^2)
\end{aligned} \tag{B.3}$$

$$\mathbb{E}_t rx_{t+m,n-m} = \frac{1}{2}\log(1-\eta\sigma^2) - \frac{1-\eta}{2}\sigma^2 - \frac{(\phi+1)^2\sigma^2}{2(1-\eta\sigma^2)}$$

So If the money growth deviation $\hat{\mu}_{t+i}$ are i.i.d random variables, then $\mathbb{E}_t rx_{t+m,n-m}$ is a constant. Hence,

$$\frac{Cov(\mathbb{E}_t rx_{t+m,n-m}, y_{t,n} - y_{t,m})}{mVar(y_{t,n} - y_{t,m})} = 0, \qquad \beta = 1$$

B.3 Derivation of the results under the MA-Assumption for $\hat{\mu}_{t+i}$

By the definition, $n \geq 2$, so

$$\begin{aligned}
&ny_{t,n}(n \geq 2) \\
=&-\log \mathbb{E}_t \exp\left[n(\log\delta - \log\bar{\mu}) - \phi\hat{\mu}_{t+n} - \sum_{i=1}^{n}\hat{\mu}_{t+i} + \frac{1}{2}\eta\hat{\mu}_{t+n}^2 + \phi\hat{\mu}_t - \frac{1}{2}\eta\hat{\mu}_t^2\right] \\
=&n(\log\bar{\mu} - \log\delta) - \phi\hat{\mu}_t + \frac{1}{2}\eta\hat{\mu}_t^2 - \log\mathbb{E}_t\exp\left(-\sum_{i=1}^{n}\hat{\mu}_{t+i} - \phi\hat{\mu}_{t+n} + \frac{1}{2}\eta\hat{\mu}_{t+n}^2\right) \\
=&n(\log\bar{\mu} - \log\delta) - \phi\hat{\mu}_t + \frac{1}{2}\eta\hat{\mu}_t^2 + \theta\varepsilon_t \\
&-\log\mathbb{E}_t\exp\left\{-(\theta+1)\sum_{i=1}^{n-2}\varepsilon_{t+i} - [\theta(\phi+1)+1]\varepsilon_{t+n-1} - (\phi+1)\varepsilon_{t+n} + \frac{1}{2}\eta\hat{\mu}_{t+n}^2\right\} \\
=&n(\log\bar{\mu} - \log\delta) - \phi\hat{\mu}_t + \frac{1}{2}\eta\hat{\mu}_t^2 + \theta\varepsilon_t + \frac{1}{2}(n-2)(\theta+1)^2\sigma^2 - \\
&\log\mathbb{E}_t\exp\left\{-[\theta(\phi+1)+1]\varepsilon_{t+n-1} - (\phi+1)\varepsilon_{t+n} + \frac{1}{2}\eta\theta^2\varepsilon_{t+n-1}^2 + \eta\theta\varepsilon_{t+n-1}\varepsilon_{t+n} + \frac{1}{2}\eta\varepsilon_{t+n}^2\right\}
\end{aligned} \tag{B.4}$$

Let

$$\varepsilon = \begin{pmatrix} \varepsilon_{t+n-1} \\ \varepsilon_{t+n} \end{pmatrix} \sim N(0,\Sigma), \text{ and } \Sigma = \sigma^2 I_2, \;\; B^T B = \frac{1}{2}\eta\begin{pmatrix} \theta^2 & \theta \\ \theta & 1 \end{pmatrix}$$

and

$$c = \begin{pmatrix} -\frac{1}{2}[\theta(\phi+1)+1] \\ -\frac{1}{2}(\phi+1) \end{pmatrix}$$

the equation can be wrriten as a matrix form

$$\begin{aligned}
&ny_{t,n} \\
=&n(\log\bar{\mu} - \log\delta) - \phi\hat{\mu}_t + \frac{1}{2}\eta\hat{\mu}_t^2 + \theta\varepsilon_t + \frac{1}{2}(n-2)(\theta+1)^2\sigma^2 \\
&-\log E_t\exp\left(\varepsilon^T B^T B\varepsilon + \varepsilon^T c + c^T\varepsilon\right)
\end{aligned} \tag{B.5}$$

$$
\begin{aligned}
=&n(\log\bar{\mu}-\log\delta)-\phi\hat{\mu}_t+\frac{1}{2}\eta\hat{\mu}_t^2+\theta\varepsilon_t+\frac{1}{2}(n-2)(\theta+1)^2\sigma^2-2c^TV^Tc\\
&+\frac{1}{2}\log\left(1-\eta\sigma^2-\eta\sigma^2\theta^2\right)
\end{aligned}
\tag{B.6}
$$

where $V=\left(\Sigma^{-1}-2B^TB\right)^{-1}$.

The following is the general rule for multivariate integration we used in this paper:

$$
\begin{aligned}
\hat{\mu}_t=&A_0+A_1\varepsilon\\
nY=&B_0+B_1\hat{\mu}_t+\hat{\mu}_tB_2\hat{\mu}_t\\
Y=&B_0+B_1A_0+B_1A_1\varepsilon
\end{aligned}
$$

$$
B\varepsilon+\varepsilon^Tc+c^T\varepsilon\qquad\varepsilon\sim N(0,\Sigma),\qquad\Sigma=\sigma^2I_n
$$

$$
\begin{aligned}
&\mathbb{E}\exp\left(\varepsilon^TB^TB\varepsilon+\varepsilon^Tc+c^T\varepsilon\right)\\
=&\oint_{R^n}\frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left(\varepsilon^TB^TB\varepsilon+\varepsilon^Tc+c^T\varepsilon\right)\exp\left(-\frac{1}{2}\varepsilon^T\Sigma^{-1}\varepsilon\right)d\varepsilon\\
=&\oint_{R^n}\frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left[-\varepsilon^T\left(\frac{1}{2}\Sigma^{-1}-B^TB\right)\varepsilon+\varepsilon^Tc+c^T\varepsilon\right]d\varepsilon
\end{aligned}
$$

We denote

$$
\frac{1}{2}\Sigma^{-1}-B^TB=\frac{1}{2}V^{-1}\ i.e.\ V=\left(\Sigma^{-1}-2B^TB\right)^{-1}\qquad\text{and}\qquad c=\frac{1}{2}V^{-1}\nu,\ i.e.\ \nu=2Vc
$$

then

$$
\begin{aligned}
&-\varepsilon^T\left(\Sigma^{-1}-B^TB\right)\varepsilon+\varepsilon^Tc+c^T\varepsilon\\
=&-\varepsilon^T\left(\Sigma^{-1}-B^TB\right)\varepsilon+\varepsilon^Tc+c^T\varepsilon-\frac{1}{2}\nu^TV^{-1}\nu+\frac{1}{2}\nu^TV^{-1}\nu\\
=&-\frac{1}{2}(\varepsilon-\nu)^TV^{-1}(\varepsilon-\nu)+\frac{1}{2}\nu^TV^{-1}\nu
\end{aligned}
$$

so

$$
\begin{aligned}
&\mathbb{E}\exp\left(\varepsilon^TB^TB\varepsilon+\varepsilon^Tc+c^T\varepsilon\right)\\
=&\oint_{R^n}\frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}(\varepsilon-\nu)^TV^{-1}(\varepsilon-\nu)+\frac{1}{2}\nu^TV^{-1}\nu\right]d\varepsilon\\
=&\oint_{R^n}\frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}(\varepsilon-\nu)^TV^{-1}(\varepsilon-\nu)\right]d\varepsilon\exp\left(\frac{1}{2}\nu^TV^{-1}\nu\right)
\end{aligned}
$$

$$
\begin{aligned}
=&\frac{1}{|\Sigma|^{\frac{1}{2}}}\frac{1}{|\Sigma^{-1}-2B^TB|^{\frac{1}{2}}}\exp\left(\frac{1}{2}\nu^TV^{-1}\nu\right)\\
=&\frac{1}{|\Sigma|^{\frac{1}{2}}}\frac{1}{|\Sigma^{-1}-2B^TB|^{\frac{1}{2}}}\exp\left(c^TV^Tc\right)\\
=&\frac{\exp\left(c^TV^Tc\right)}{|I_n-2I_n\sigma^2B^TB|^{\frac{1}{2}}}
\end{aligned}
$$

According to (4.19), we derive the expressions of $my_{t,m}$ in the cases of $m=1$ and $m\geq 2$ separately.

If $m\geq 2$, with similar derivation, we have:

$$
\begin{aligned}
&my_{t,m}\\
=&m(\log\bar{\mu}-\log\delta)-\phi\hat{\mu}_t+\frac{1}{2}\eta\hat{\mu}_t^2+\theta\varepsilon_t+\frac{1}{2}(m-2)(\theta+1)^2\sigma^2 \qquad \text{(B.7)}\\
&-\log\mathbb{E}_t\exp\left(\hat{\varepsilon}^TB^TB\hat{\varepsilon}+\hat{\varepsilon}^Tc+c^T\hat{\varepsilon}\right)\\
=&m(\log\bar{\mu}-\log\delta)-\phi\hat{\mu}_t+\frac{1}{2}\eta\hat{\mu}_t^2+\theta\varepsilon_t+\frac{1}{2}(m-2)(\theta+1)^2\sigma^2-2c^TV^Tc\\
&+\frac{1}{2}\log\left(1-\eta\sigma^2-\eta\sigma^2\theta^2\right) \qquad \text{(B.8)}
\end{aligned}
$$

where

$$
\hat{\varepsilon}=\begin{pmatrix}\varepsilon_{t+m-1}\\ \varepsilon_{t+m}\end{pmatrix}\sim N(0,\Sigma),\ \text{and}\ \Sigma=\sigma^2I_2
$$

If m=1, $my_{t,m}=y_{t,1}$, and

$$
\begin{aligned}
y_{t,1}=&-\log\mathbb{E}_t\exp\left[\log\delta-\log\bar{\mu}-\phi\hat{\mu}_{t+1}-\hat{\mu}_{t+1}+\frac{1}{2}\eta\hat{\mu}_{t+1}^2+\phi\hat{\mu}_t-\frac{1}{2}\eta\hat{\mu}_t^2\right]\\
=&\log\bar{\mu}-\log\delta+(\phi+1)\theta\varepsilon_t-\phi\hat{\mu}_t+\frac{1}{2}\eta\hat{\mu}_t^2-\frac{1}{2}\eta\theta^2\varepsilon_t^2\\
&-\log\mathbb{E}_t\exp\left[(\eta\theta\varepsilon_t-\phi-1)\varepsilon_{t+1}+\frac{1}{2}\eta\varepsilon_{t+1}^2\right]\\
=&\log\bar{\mu}-\log\delta+(\phi+1)\theta\varepsilon_t-\phi\hat{\mu}_t+\frac{1}{2}\eta\hat{\mu}_t^2-\frac{1}{2}\eta\theta^2\varepsilon_t^2-\frac{(\eta\theta\varepsilon_t-\phi-1)^2\sigma^2}{2(1-\eta\sigma^2)}\\
&+\frac{1}{2}\log(1-\eta\sigma^2) \qquad \text{(B.9)}
\end{aligned}
$$

For the same reason, we should derive the expressions of $(n-m)y_{t+m,n-m}$ in the cases of $n-m=1$ and $n-m\geq 2$ separately.

B.3.1 Results of n-m=1

If $n-m=1$, $(n-m)y_{t+m,n-m} = y_{t+m,1}$,

$$\begin{aligned}
y_{t+m,1} =& -\log \mathbb{E}_{t+m} \exp\left[\log\delta - \log\bar{\mu} - \phi\hat{\mu}_{t+m+1} - \hat{\mu}_{t+m+1} + \frac{1}{2}\eta\hat{\mu}_{t+m+1}^2 + \phi\hat{\mu}_t - \frac{1}{2}\eta\hat{\mu}_t^2\right] \\
=& \log\bar{\mu} - \log\delta + (\phi+1)\theta\varepsilon_{t+m} - \phi\hat{\mu}_{t+m} + \frac{1}{2}\eta\hat{\mu}_{t+m}^2 - \frac{1}{2}\eta\theta^2\varepsilon_{t+m}^2 \\
& - \log \mathbb{E}_{t+m} \exp\left[(\eta\theta\varepsilon_{t+m} - \phi - 1)\varepsilon_{t+m+1} + \frac{1}{2}\eta\varepsilon_{t+m+1}^2\right] \\
=& \log\bar{\mu} - \log\delta + (\phi+1)\theta\varepsilon_{t+m} - \phi\hat{\mu}_{t+m} + \frac{1}{2}\eta\hat{\mu}_{t+m}^2 - \frac{1}{2}\eta\theta^2\varepsilon_{t+m}^2 \\
& - \frac{(\eta\theta\varepsilon_{t+m} - \phi - 1)^2\sigma^2}{2(1-\eta\sigma^2)} + \frac{1}{2}\log(1-\eta\sigma^2)
\end{aligned}$$

a. If $m \geq 2$, then

$$\mathbb{E}_t[y_{t+m,1}] = \log\bar{\mu} - \log\delta + \frac{1}{2}\eta(\theta^2+1)\sigma^2 - \frac{(\phi-1)^2\sigma^2 + \eta^2\theta^2\sigma^4}{2(1-\eta\sigma^2)} + \frac{1}{2}\log(1-\eta\sigma^2)$$

is a constant.

Moreover, for $m \geq 2$,

$$my_{t,m} - ny_{t,n} = (m-n)(\log\bar{\mu} - \log\delta) + \frac{1}{2}(m-n)(\theta+1)^2\sigma^2 \tag{B.10}$$

is a constant too.

The value of $\mathbb{E}_t rx_{t+m,n-m} = \mathbb{E}_t[y_{t+m,1}] + my_{t,m} - ny_{t,n}$ is the difference of two constants, which must be a constant. Hence

$$\frac{Cov(\mathbb{E}_t rx_{t+m,n-m}, y_{t,n} - y_{t,m})}{mVar(y_{t,n} - y_{t,m})} = 0, \qquad \beta = 1$$

b. If $m=1$, then $y_{t+m,1} = y_{t+1,1}$, and

$$\begin{aligned}
\mathbb{E}_t[y_{t+1,1}] =& \log\bar{\mu} - \log\delta - \phi\theta\varepsilon_t + \frac{1}{2}\eta\theta^2\varepsilon_t^2 + \frac{1}{2}\eta(1-\theta^2)\sigma^2 - \frac{(\phi+1)^2\sigma^2 + \eta^2\theta^2\sigma^4}{2(1-\eta\sigma^2)} \\
& + \frac{1}{2}\log(1-\eta\sigma^2)
\end{aligned}$$

For $m=1$, since $n-m=1$, so $n=2$, and then

$$y_{t,1} - 2y_{t,2} = 2c^T V^T c - \log\bar{\mu} + \log\delta + \phi\theta\varepsilon_t - \frac{1}{2}\log\left(\frac{1-\eta\sigma^2-\eta\sigma^2\theta^2}{1-\eta\sigma^2}\right) - \frac{1}{2}\eta\theta^2\varepsilon_t^2$$

$$- \frac{(\eta\theta\varepsilon_t - \phi - 1)^2\sigma^2}{2(1-\eta\sigma^2)} \tag{B.11}$$

$$\begin{aligned}\mathbb{E}_t rx_{t+1,1} =& \mathbb{E}_t[y_{t+1,1}] + y_{t,1} - 2y_{t,2} \\ =& 2c^T V^T c - \frac{1}{2}\log\left(\frac{1-\eta\sigma^2-\eta\sigma^2\theta^2}{1-\eta\sigma^2}\right) - \frac{(\eta\theta\varepsilon_t-\phi-1)^2\sigma^2}{2(1-\eta\sigma^2)} + \log(1-\eta\sigma^2) \\ &+ \frac{1}{2}\eta(1-\theta^2)\sigma^2 - \frac{(\phi+1)^2\sigma^2+\eta^2\theta^2\sigma^4}{2(1-\eta\sigma^2)}\end{aligned} \tag{B.12}$$

$$\begin{aligned}y_{t,2} - y_{t,1} =& \frac{\phi}{2}\hat{\mu}_t - \frac{1}{4}\eta\hat{\mu}_t^2 - \left(\phi+\frac{1}{2}\right)\theta\varepsilon_t + \frac{1}{2}\eta\theta^2\varepsilon_t^2 + \frac{(\eta\theta\varepsilon_t-\phi-1)^2\sigma^2}{2(1-\eta\sigma^2)} - c^T V^T c \\ &+ \frac{1}{4}\log\left(1-\eta\sigma^2-\eta\sigma^2\theta^2\right) - \frac{1}{2}\log(1-\eta\sigma^2)\end{aligned} \tag{B.13}$$

Dropping the constant part of $\mathbb{E}_t rx_{t+1,1}$ and $y_{t,2} - y_{t,1}$, we get

$$\mathbb{E}_t\widehat{rx}_{t+1,1} = \frac{2\eta\theta(\phi+1)\sigma^2\varepsilon_t - \eta^2\theta^2\sigma^2\varepsilon_t^2}{2(1-\eta\sigma^2)} \tag{B.14}$$

$$\hat{y}_{t,2} - \hat{y}_{t,1} = \frac{\phi}{2}\hat{\mu}_t - \frac{1}{4}\eta\hat{\mu}_t^2 - \left(\phi+\frac{1}{2}\right)\theta\varepsilon_t + \frac{1}{2}\eta\theta^2\varepsilon_t^2 + \frac{\eta^2\theta^2\sigma^2\varepsilon_t^2 - 2\eta\theta(\phi+1)\sigma^2\varepsilon_t}{2(1-\eta\sigma^2)} \tag{B.15}$$

$$\frac{Cov(\mathbb{E}_t rx_{t+1,1}, y_{t,2}-y_{t,1})}{Var(y_{t,2}-y_{t,1})} = \frac{Cov(\mathbb{E}_t\widehat{rx}_{t+1,1}, \hat{y}_{t,2}-\hat{y}_{t,1})}{Var(\hat{y}_{t,2}-\hat{y}_{t,1})} \neq 0 \Rightarrow \beta \neq 1$$

B.3.2 Results of $n - m \geq 2$

If $n - m \geq 2$, then

$$\begin{aligned}&(n-m)y_{t+m,n-m} \\ =& -\log\mathbb{E}_{t+m}\exp\left[(n-m)(\log\delta - \log\bar{\mu}) - \phi\hat{\mu}_{t+n} - \sum_{i=1}^{n-m}\hat{\mu}_{t+m+i} + \frac{1}{2}\eta\hat{\mu}_{t+n}^2 + \phi\hat{\mu}_{t+m} - \frac{1}{2}\eta\hat{\mu}_{t+m}^2\right] \\ =& (n-m)(\log\bar{\mu} - \log\delta) - \phi\hat{\mu}_{t+m} + \frac{1}{2}\eta\hat{\mu}_{t+m}^2 + \theta\varepsilon_{t+m} + \frac{1}{2}(n-m-2)(\theta+1)^2\sigma^2 \\ &- \log\mathbb{E}_{t+m}\exp\left(\varepsilon^T B^T B\varepsilon + \varepsilon^T c + c^T\varepsilon\right) \\ =& (n-m)(\log\bar{\mu} - \log\delta) - \phi\hat{\mu}_{t+m} + \frac{1}{2}\eta\hat{\mu}_{t+m}^2 + \theta\varepsilon_{t+m} + \frac{1}{2}(n-m-2)(\theta+1)^2\sigma^2 \\ &- 2c^T V^T c + \frac{1}{2}\log\left(1-\eta\sigma^2-\eta\sigma^2\theta^2\right)\end{aligned}$$

where

$$\varepsilon = \begin{pmatrix} \varepsilon_{t+n-1} \\ \varepsilon_{t+n} \end{pmatrix} \sim N(0, \Sigma), \text{ and } \Sigma = \sigma^2 I_2$$

a. if $m \geq 2$,

$$\begin{aligned}
&\mathbb{E}_t[(n-m)y_{t+m,n-m}] \\
=&(n-m)(\log\bar{\mu} - \log\delta) + \frac{1}{2}(n-m-2)(\theta+1)^2\sigma^2 - 2c^T V^T c + \frac{1}{2}\log\left(1-\eta\sigma^2-\eta\sigma^2\theta^2\right) \\
&+ \mathbb{E}_t\left[-\phi\hat{\mu}_{t+m} + \frac{1}{2}\eta\hat{\mu}_{t+m}^2 + \theta\varepsilon_{t+m}\right] \\
=&(n-m)(\log\bar{\mu} - \log\delta) + \frac{1}{2}(n-m-2)(\theta+1)^2\sigma^2 - 2c^T V^T c + \frac{1}{2}\log\left(1-\eta\sigma^2-\eta\sigma^2\theta^2\right) \\
&+ \mathbb{E}_t\left[(\theta-\phi)\varepsilon_{t+m} + \theta\varepsilon_{t+m-1} + \frac{\eta}{2}(\varepsilon_{t+m}^2 + \theta^2\varepsilon_{t+m-1}^2 + 2\theta\varepsilon_{t+m}\varepsilon_{t+m-1})\right] \\
=&(n-m)(\log\bar{\mu} - \log\delta) + \frac{1}{2}(n-m-2)(\theta+1)^2\sigma^2 + \frac{\eta}{2}(\theta^2+1)\sigma^2 \\
&- 2c^T V^T c + \frac{1}{2}\log\left(1-\eta\sigma^2-\eta\sigma^2\theta^2\right)
\end{aligned}$$

$\mathbb{E}_t[(n-m)y_{t+m,n-m}]$ is a constant.

$$\mathbb{E}_t rx_{t+m,n-m} = \mathbb{E}_t[(n-m)y_{t+m,n-m}] + my_{t,m} - ny_{t,n} \tag{B.16}$$

If $m \geq 2$, as (B.10) indicates, $ny_{t,n} - my_{t,m}$ is a constant too, so $\mathbb{E}_t rx_{t+m,n-m}$ is a constant.

Hence

$$\frac{Cov(\mathbb{E}_t rx_{t+m,n-m}, y_{t,n} - y_{t,m})}{mVar(y_{t,n} - y_{t,m})} = 0, \qquad \beta = 1$$

b. If $m = 1$, then $my_{t,m} - ny_{t,n} = y_{t,1} - ny_{t,n}$, and

$$\begin{aligned}
y_{t,1} - ny_{t,n} =& -(n-1)(\log\bar{\mu} - \log\delta) + \phi\theta\varepsilon_t + 2c^T V^T c - \frac{1}{2}\log\left(\frac{1-\eta\sigma^2-\eta\sigma^2\theta^2}{1-\eta\sigma^2}\right) \\
&- \frac{1}{2}\eta\theta^2\varepsilon_t^2 - \frac{(\eta\theta\varepsilon_t - \phi - 1)^2\sigma^2}{2(1-\eta\sigma^2)} - \frac{1}{2}(n-2)(\theta+1)^2\sigma^2
\end{aligned} \tag{B.17}$$

And $\mathbb{E}_t[(n-m)y_{t+m,n-m}] = \mathbb{E}_t[(n-1)y_{t+1,n-1}]$

$$\begin{aligned}
\mathbb{E}_t[(n-1)y_{t+1,n-1}] =& (n-1)(\log\bar{\mu} - \log\delta) - \phi\theta\varepsilon_t + \frac{1}{2}\eta\theta^2\varepsilon_t^2 + \frac{1}{2}(n-3)(\theta+1)^2\sigma^2 \\
&+ \frac{\eta}{2}(\theta^2+1)\sigma^2 + \frac{1}{2}\log\left(1-\eta\sigma^2-\eta\sigma^2\theta^2\right) - 2c^T V^T c
\end{aligned} \tag{B.18}$$

We denote $\tilde{\mathbb{E}}_t rx_{t+1,n-1}$ as a approximate of $\mathbb{E}_t rx_{t+1,n-1}$, which does not contain the constant terms of $\mathbb{E}_t rx_{t+1,n-1}$, then

$$\mathbb{E}_t \widetilde{rx}_{t+1,n-1} = \frac{2\eta\theta(\phi+1)\sigma^2\varepsilon_t - \eta^2\theta^2\sigma^2\varepsilon_t^2}{2(1-\eta\sigma^2)} \tag{B.19}$$

We denote $n\tilde{y}_{t,n} - \tilde{y}_{t,1}$ as a approximate of $ny_{t,n} - y_{t,1}$, which does not contain the constant terms of $ny_{t,n} - y_{t,1}$, then

$$n\tilde{y}_{t,n} - \tilde{y}_{t,1} = (1-\frac{1}{n})\phi\hat{\mu}_t + (\frac{1}{2n} - \frac{1}{2})\eta\hat{\mu}_t^2 + \left(\frac{1}{n} - \phi - 1\right)\theta\varepsilon_t + \frac{1}{2}\eta\theta^2\varepsilon_t^2 + \frac{\eta^2\theta^2\sigma^2\varepsilon_t^2 - 2\eta\theta(\phi+1)\sigma^2\varepsilon_t}{2(1-\eta\sigma^2)} \tag{B.20}$$

$$\frac{Cov(\mathbb{E}_t rx_{t+1,n-1}, y_{t,n} - y_{t,1})}{Var(y_{t,n} - y_{t,1})} = \frac{Cov(\mathbb{E}_t \widetilde{rx}_{t+1,n-1}, \tilde{y}_{t,n} - \tilde{y}_{t,1})}{Var(\tilde{y}_{t,2} - \tilde{y}_{t,1})} \neq 0 \Rightarrow \beta \neq 1$$

B.4 Derivation of the results under the AR(1)-Assumption for $\hat{\mu}_{t+i}$

If $\hat{\mu}_t$ is an AR(1) process, i.e. $\hat{\mu}_{t+1} = \rho\hat{\mu}_t + \varepsilon_{t+1}$, then we have

$$\mathbb{E}_t\hat{\mu}_{t+1} = \rho\hat{\mu}_t \qquad \mathbb{E}_t\hat{\mu}_{t+m} = \rho^m\hat{\mu}_t \qquad \text{and} \qquad \mathbb{E}_t\hat{\mu}_{t+n} = \rho^n\hat{\mu}_t$$

$$\begin{aligned}
\hat{\mu}_{t+n} =& \rho^n\hat{\mu}_t + \sum_{i=1}^{n}\rho^{i-1}\varepsilon_{t+n+1-i} \\
\hat{\mu}_{t+n}^2 =& \rho^{2n}\hat{\mu}_t^2 + \sum_{i=1}^{n}\sum_{j=1}^{n}\rho^{i+j-2}\varepsilon_{t+n+1-i}\varepsilon_{t+n+1-j} + 2\rho^n\hat{\mu}_t\sum_{i=1}^{n}\rho^{i-1}\varepsilon_{t+n+1-i}
\end{aligned}$$

$$\begin{aligned}
& ny_{t,n} \\
=& -\log\mathbb{E}_t\exp\left[n(\log\delta - \log\bar{\mu}) - \phi\hat{\mu}_{t+n} - \sum_{i=1}^{n}\hat{\mu}_{t+i} + \frac{1}{2}\eta\hat{\mu}_{t+n}^2 + \phi\hat{\mu}_t - \frac{1}{2}\eta\hat{\mu}_t^2\right] \\
=& n(\log\bar{\mu} - \log\delta) - \phi\hat{\mu}_t + \frac{1}{2}\eta\hat{\mu}_t^2 - \log\mathbb{E}_t\exp\left(-\sum_{i=1}^{n}\hat{\mu}_{t+i} - \phi\hat{\mu}_{t+n} + \frac{1}{2}\eta\hat{\mu}_{t+n}^2\right) \\
=& n(\log\bar{\mu} - \log\delta) + \left(\sum_{i=1}^{n}\rho^i + \phi\rho^n - \phi\right)\hat{\mu}_t + \frac{1}{2}\eta\left(1-\rho^{2n}\right)\hat{\mu}_t^2 - \\
& \log\mathbb{E}_t\exp\left[\frac{1}{2}\eta\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\rho^{i+j-2}\varepsilon_{t+n+1-i}\varepsilon_{t+n+1-j} + 2\rho^n\hat{\mu}_t\sum_{i=1}^{n}\rho^{i-1}\varepsilon_{t+n+1-i}\right) - \sum_{i=1}^{n}\sum_{j=1}^{i}\rho^{j-1}\varepsilon_{t+i+1-j} - \phi\sum_{i=1}^{n}\rho^{i-1}\varepsilon_{t+n+1-i}\right]
\end{aligned}$$

When we rewrite the equation as a matrix form, we get

$$\begin{aligned}
& ny_{t,n} \\
=& n(\log\bar{\mu} - \log\delta) + \left(\sum_{i=1}^{n}\rho^i + \phi\rho^n - \phi\right)\hat{\mu}_t + \frac{1}{2}\eta\left(1-\rho^{2n}\right)\hat{\mu}_t^2
\end{aligned}$$

$$
\begin{aligned}
&-\log \mathbb{E}_t \exp\left(\epsilon^T A^T A \epsilon + \epsilon^T D + D^T \epsilon\right) \\
=&n(\log\bar{\mu} - \log\delta) + \left(\sum_{i=1}^{n}\rho^i + \phi\rho^n - \phi\right)\hat{\mu}_t + \frac{1}{2}\eta\left(1-\rho^{2n}\right)\hat{\mu}_t^2 - 2D^T\Omega^T D \\
&+\frac{1}{2}\log\left|I_n - 2\Sigma A^T A\right|
\end{aligned} \tag{B.21}
$$

where

$$
\epsilon = \begin{pmatrix} \varepsilon_{t+1} \\ \vdots \\ \varepsilon_{t+n} \end{pmatrix} \sim N(0,\Sigma), \text{ and } \Sigma = \sigma^2 I_n, \;\; A^T A = \frac{1}{2}\eta \begin{pmatrix} \rho^{2(n-1)} & \cdots & \rho^{(n-1)} \\ \vdots & \ddots & \vdots \\ \rho^{(n-1)} & \cdots & 1 \end{pmatrix}
$$

and

$$
D = \frac{1}{2}\begin{pmatrix} (\eta\rho^n\hat{\mu}_t - \phi)\rho^{n-1} - \sum_{i=1}^{n}\rho^{i-1} \\ (\eta\rho^n\hat{\mu}_t - \phi)\rho^{n-2} - \sum_{i=1}^{n-1}\rho^{i-1} \\ \vdots \\ (\eta\rho^n\hat{\mu}_t - \phi)\rho - (1+\rho) \\ (\eta\rho^n\hat{\mu}_t - \phi) - 1 \end{pmatrix}
$$

$$
\frac{1}{2}\Sigma^{-1} - A^T A = \frac{1}{2}\Omega^{-1} \; i.e. \; \Omega = \left(\Sigma^{-1} - 2A^T A\right)^{-1} \qquad \text{and} \qquad D = \frac{1}{2}\Omega^{-1}\nu, \; i.e. \; \nu = 2\Omega D \tag{B.22}
$$

$$
\begin{aligned}
(n-m)y_{t+m,n-m} =&(n-m)(\log\bar{\mu} - \log\delta) + \left(\sum_{i=1}^{n-m}\rho^i + \phi\rho^{n-m} - \phi\right)\hat{\mu}_{t+m} \\
&+\frac{1}{2}\eta\left(1-\rho^{2n}\right)\hat{\mu}_{t+m}^2 - 2\hat{D}^T\hat{\Omega}^T\hat{D} + \frac{1}{2}\log\left|I_{n-m} - 2\hat{\Sigma}\hat{A}^T\hat{A}\right|
\end{aligned} \tag{B.23}
$$

Clearly, it is impossible to get a general analytical solution for any number of n and m. For simplification, we try the easiest number, i.e. $n = 2$ and $m = 1$. Substituting $n = 2$ and $m = 1$ into (B.21) and (B.23) we get:

$$
\begin{aligned}
&2y_{t,2} \\
=&-\log\mathbb{E}_t\exp\left[2(\log\delta - \log\bar{\mu}) - \phi\hat{\mu}_{t+2} - \hat{\mu}_{t+1} - \hat{\mu}_{t+2} + \frac{1}{2}\eta\hat{\mu}_{t+2}^2 + \phi\hat{\mu}_t - \frac{1}{2}\eta\hat{\mu}_t^2\right] \\
=&2(\log\bar{\mu} - \log\delta) - \phi\hat{\mu}_t + \frac{1}{2}\eta\hat{\mu}_t^2 - \log\mathbb{E}_t\exp\left(-\hat{\mu}_{t+1} - \hat{\mu}_{t+2} - \phi\hat{\mu}_{t+2} + \frac{1}{2}\eta\hat{\mu}_{t+2}^2\right) \\
=&2(\log\bar{\mu} - \log\delta) + \left(\rho + \rho^2 + \phi\rho^2 - \phi\right)\hat{\mu}_t + \frac{1}{2}\eta\left(1-\rho^4\right)\hat{\mu}_t^2 - \\
&\log\mathbb{E}_t\exp\left[\frac{1}{2}\eta\left(\sum_{i=1}^{2}\sum_{j=1}^{2}\rho^{i+j-2}\varepsilon_{t+3-i}\varepsilon_{t+3-j} + 2\rho^2\hat{\mu}_t\sum_{i=1}^{2}\rho^{i-1}\varepsilon_{t+3-i}\right) - \sum_{i=1}^{2}\sum_{j=1}^{i}\rho^{j-1}\varepsilon_{t+i+1-j} - \phi\sum_{i=1}^{2}\rho^{i-1}\varepsilon_{t+3-i}\right]
\end{aligned} \tag{B.24}
$$

$$
\begin{aligned}
&\frac{1}{2}\eta\left(\sum_{i=1}^{2}\sum_{j=1}^{2}\rho^{i+j-2}\varepsilon_{t+3-i}\varepsilon_{t+3-j}+2\rho^2\hat{\mu}_t\sum_{i=1}^{2}\rho^{i-1}\varepsilon_{t+3-i}\right)-\sum_{i=1}^{2}\sum_{j=1}^{i}\rho^{j-1}\varepsilon_{t+i+1-j}\\
&-\phi\sum_{i=1}^{2}\rho^{i-1}\varepsilon_{t+3-i}\\
=&\frac{1}{2}\eta\left(\varepsilon_{t+2}^2+\rho^2\varepsilon_{t+1}^2+2\rho\varepsilon_{t+2}\varepsilon_{t+1}+2\rho^3\hat{\mu}_t\varepsilon_{t+1}+2\rho^2\hat{\mu}_t\varepsilon_{t+2}\right)-\varepsilon_{t+1}-\varepsilon_{t+2}-\rho\varepsilon_{t+1}\\
&-\phi\varepsilon_{t+2}-\phi\rho\varepsilon_{t+1}\\
=&\frac{1}{2}\eta\left(\varepsilon_{t+2}^2+\rho^2\varepsilon_{t+1}^2+2\rho\varepsilon_{t+2}\varepsilon_{t+1}\right)+\left(\eta\rho^3\hat{\mu}_t-\phi\rho-\rho-1\right)\varepsilon_{t+1}+\left(\eta\rho^2\hat{\mu}_t-\phi-1\right)\varepsilon_{t+2}
\end{aligned}
$$

When we rewrite (B.24) as a matrix form, we get

$$
\begin{aligned}
&2y_{t,2}\\
=&2(\log\bar{\mu}-\log\delta)+\left(\rho^2+\phi\rho^2+\rho-\phi\right)\hat{\mu}_t+\frac{1}{2}\eta\left(1-\rho^4\right)\hat{\mu}_t^2\\
&-\log\mathbb{E}_{t+m}\exp\left(\widetilde{\epsilon}^T\widetilde{A}^T\widetilde{A}\widetilde{\epsilon}+\widetilde{\epsilon}^T\widetilde{D}+\widetilde{D}^T\widetilde{\epsilon}\right)\\
=&2(\log\bar{\mu}-\log\delta)+\left(\rho^2+\phi\rho^2+\rho-\phi\right)\hat{\mu}_t+\frac{1}{2}\eta\left(1-\rho^4\right)\hat{\mu}_t^2\\
&-\log\left[\frac{1}{|\widetilde{\Sigma}|^{\frac{1}{2}}}\frac{1}{\left|\widetilde{\Sigma}^{-1}-2\widetilde{A}^T\widetilde{A}\right|^{\frac{1}{2}}}\exp\left(\widetilde{D}^T\widetilde{\Omega}^T\widetilde{D}\right)\right]\\
=&2(\log\bar{\mu}-\log\delta)+\left(\rho^2+\phi\rho^2+\rho-\phi\right)\hat{\mu}_t+\frac{1}{2}\eta\left(1-\rho^4\right)\hat{\mu}_t^2-2\widetilde{D}^T\widetilde{\Omega}^T\widetilde{D}\\
&+\frac{1}{2}\log\left(1-\eta\sigma^2-\eta\sigma^2\rho^2\right)
\end{aligned}
\tag{B.25}
$$

where

$$
\widetilde{\epsilon}=\begin{pmatrix}\varepsilon_{t+1}\\ \varepsilon_{t+2}\end{pmatrix}\sim N(0,\widetilde{\Sigma}),\ \text{and}\ \widetilde{\Sigma}=\sigma^2 I_2,\quad \widetilde{A}^T\widetilde{A}=\frac{1}{2}\eta\begin{pmatrix}\rho^2 & \rho\\ \rho & 1\end{pmatrix}
$$

and

$$
\widetilde{D}=\frac{1}{2}\begin{pmatrix}\eta\rho^3\hat{\mu}_t-\phi\rho-\rho-1\\ \eta\rho^2\hat{\mu}_t-\phi-1\end{pmatrix}
$$

$$
\frac{1}{2}\widetilde{\Sigma}^{-1}-\widetilde{A}^T\widetilde{A}=\frac{1}{2}\widetilde{\Omega}^{-1}\ i.e.\ \widetilde{\Omega}=\left(\widetilde{\Sigma}^{-1}-2\widetilde{A}^T\widetilde{A}\right)^{-1}\qquad\text{and}\qquad\widetilde{D}=\frac{1}{2}\widetilde{\Omega}^{-1}\widetilde{\nu},\ i.e.\ \widetilde{\nu}=2\widetilde{\Omega}\widetilde{D}
$$

$$
y_{t,1}=-\log\mathbb{E}_t\exp\left[\log\delta-\log\bar{\mu}-\phi\hat{\mu}_{t+1}-\hat{\mu}_{t+1}+\frac{1}{2}\eta\hat{\mu}_{t+1}^2+\phi\hat{\mu}_t-\frac{1}{2}\eta\hat{\mu}_t^2\right]
$$

$$
\begin{aligned}
&= \log\bar{\mu} - \log\delta - \phi\hat{\mu}_t + \frac{1}{2}\eta\hat{\mu}_t^2 - \log\mathbb{E}_t\exp\left[-(\phi+1)\hat{\mu}_{t+1} + \frac{1}{2}\eta\hat{\mu}_{t+1}^2\right] \\
&= \log\bar{\mu} - \log\delta + (\rho\phi+\rho-\phi)\hat{\mu}_t + \frac{1}{2}\eta(1-\rho^2)\hat{\mu}_t^2 \\
&\quad - \log\mathbb{E}_t\exp\left[(\eta\rho\hat{\mu}_t - \phi - 1)\varepsilon_{t+1} + \frac{1}{2}\eta\varepsilon_{t+1}^2\right] \\
&= \log\bar{\mu} - \log\delta + (\rho\phi+\rho-\phi)\hat{\mu}_t + \frac{1}{2}\eta(1-\rho^2)\hat{\mu}_t^2 - \frac{(\eta\rho\hat{\mu}_t - \phi - 1)^2\sigma^2}{2(1-\eta\sigma^2)} + \frac{1}{2}\log(1-\eta\sigma^2)
\end{aligned}
\tag{B.26}
$$

$$
\begin{aligned}
&y_{t+1,1} \\
&= -\log\mathbb{E}_{t+1}\exp\left[\log\delta - \log\bar{\mu} - \phi\hat{\mu}_{t+2} - \hat{\mu}_{t+2} + \frac{1}{2}\eta\hat{\mu}_{t+2}^2 + \phi\hat{\mu}_{t+1} - \frac{1}{2}\eta\hat{\mu}_{t+1}^2\right] \\
&= \log\bar{\mu} - \log\delta - \phi\hat{\mu}_{t+1} + \frac{1}{2}\eta\hat{\mu}_{t+1}^2 - \log\mathbb{E}_{t+1}\exp\left[-(\phi+1)\hat{\mu}_{t+2} + \frac{1}{2}\eta\hat{\mu}_{t+2}^2\right] \\
&= \log\bar{\mu} - \log\delta + (\rho\phi+\rho-\phi)\hat{\mu}_{t+1} + \frac{1}{2}\eta(1-\rho^2)\hat{\mu}_{t+1}^2 \\
&\quad - \log\mathbb{E}_{t+1}\exp\left[(\eta\rho\hat{\mu}_{t+1} - \phi - 1)\varepsilon_{t+2} + \frac{1}{2}\eta\varepsilon_{t+2}^2\right] \\
&= \log\bar{\mu} - \log\delta + (\rho\phi+\rho-\phi)\hat{\mu}_{t+1} + \frac{1}{2}\eta(1-\rho^2)\hat{\mu}_{t+1}^2 - \frac{(\eta\rho\hat{\mu}_{t+1} - \phi - 1)^2\sigma^2}{2(1-\eta\sigma^2)} \\
&\quad + \frac{1}{2}\log(1-\eta\sigma^2)
\end{aligned}
\tag{B.27}
$$

$$
\begin{aligned}
\mathbb{E}_t(y_{t+1,1}) = &\log\bar{\mu} - \log\delta + (\rho^2 + \rho^2\phi - \rho\phi)\hat{\mu}_t + \frac{1}{2}\eta(\rho^2-\rho^4)\hat{\mu}_t^2 + \frac{1}{2}\eta(1-\rho^2)\sigma^2 \\
&- \frac{(\eta\rho^2\hat{\mu}_t - \phi - 1)^2\sigma^2 + \eta^2\rho^2\sigma^4}{2(1-\eta\sigma^2)} + \frac{1}{2}\log(1-\eta\sigma^2)
\end{aligned}
\tag{B.28}
$$

$$
\begin{aligned}
y_{t,1} - 2y_{t,2} = &2\widetilde{D}^T\widetilde{\Omega}^T\widetilde{D} - (\log\bar{\mu} - \log\delta) - (\rho^2 + \rho^2\phi - \rho\phi)\hat{\mu}_t - \frac{1}{2}\eta(\rho^2-\rho^4)\hat{\mu}_t^2 \\
&- \frac{1}{2}\log\left(1 - \eta\sigma^2 - \eta\sigma^2\rho^2\right) - \frac{(\eta\rho\hat{\mu}_t - \phi - 1)^2\sigma^2}{2(1-\eta\sigma^2)} + \frac{1}{2}\log(1-\eta\sigma^2)
\end{aligned}
\tag{B.29}
$$

$$
\begin{aligned}
&\mathbb{E}_t rx_{t+1,2} \\
=&\mathbb{E}_t(y_{t+1,1}) + y_{t,1} - 2y_{t,2} \\
=&-\frac{(\eta\rho\hat{\mu}_t - \phi - 1)^2\sigma^2 + (\eta\rho^2\hat{\mu}_t - \phi - 1)^2\sigma^2 + \eta^2\rho^2\sigma^4}{2(1-\eta\sigma^2)} + \log(1-\eta\sigma^2) + \frac{1}{2}\eta(1-\rho^2)\sigma^2 \\
&+ 2\widetilde{D}^T\Omega^T\widetilde{D} - \frac{1}{2}\log\left(1 - \eta\sigma^2 - \sigma^2\rho^2\right)
\end{aligned}
$$

$$= -\frac{\eta^2\rho^2\sigma^2(1+\rho^2)}{2(1-\eta\sigma^2)}\hat{\mu}_t^2 + \frac{2\eta\rho\sigma^2(\phi+1)(\rho+1)}{2(1-\eta\sigma^2)}\hat{\mu}_t + 2\widetilde{D}^T\Omega^T\widetilde{D} + \text{a constant term} \tag{B.30}$$

where

$$\begin{aligned}
&\widetilde{D}^T\widetilde{\Omega}^T\widetilde{D} \\
=&\left[\frac{(\eta\sigma^4-\sigma^2)(\rho+\phi\rho-\eta\rho^3\hat{\mu}_t+1)}{4(\eta\rho^2\sigma^2+\eta\sigma^2-1)} - \frac{\eta\rho\sigma^4(-\eta\rho^2\hat{\mu}_t+\phi+1)}{4(\eta\rho^2\sigma^2+\eta\sigma^2-1)}\right](\rho+\phi\rho-\eta\rho^3\hat{\mu}_t+1) \\
&-\left[\frac{(\sigma^2-\eta\rho^2\sigma^4)(-\eta\rho^2\hat{\mu}_t+\phi+1)}{4(\eta\rho^2\sigma^2+\eta\sigma^2-1)} + \frac{\eta\rho\sigma^4(\rho+\phi\rho-\eta\rho^3\hat{\mu}_t+1)}{4(\eta\rho^2\sigma^2+\eta\sigma^2-1)}\right]\left(-\eta\rho^2\hat{\mu}_t+\phi+1\right) \\
=&\frac{(\eta\sigma^4-\sigma^2)(\rho+\phi\rho-\eta\rho^3\hat{\mu}_t+1)^2 - \eta\rho\sigma^4(-\eta\rho^2\hat{\mu}_t+\phi+1)(\rho+\phi\rho-\eta\rho^3\hat{\mu}_t+1)}{4(\eta\rho^2\sigma^2+\eta\sigma^2-1)} \\
&-\frac{(\sigma^2-\eta\rho^2\sigma^4)(-\eta\rho^2\hat{\mu}_t+\phi+1)^2 + \eta\rho\sigma^4(-\eta\rho^2\hat{\mu}_t+\phi+1)(\rho+\phi\rho-\eta\rho^3\hat{\mu}_t+1)}{4(\eta\rho^2\sigma^2+\eta\sigma^2-1)} \\
=&\frac{(\eta\sigma^4-\sigma^2)(\rho+\phi\rho-\eta\rho^3\hat{\mu}_t+1)^2 - 2\eta\rho\sigma^4(-\eta\rho^2\hat{\mu}_t+\phi+1)(\rho+\phi\rho-\eta\rho^3\hat{\mu}_t+1)}{4(\eta\rho^2\sigma^2+\eta\sigma^2-1)} \\
&-\frac{(\sigma^2-\eta\rho^2\sigma^4)(-\eta\rho^2\hat{\mu}_t+\phi+1)^2}{4(\eta\rho^2\sigma^2+\eta\sigma^2-1)} \\
=&\frac{-\eta^2\rho^4\sigma^2(\rho^2+1)}{4(\eta\rho^2\sigma^2+\eta\sigma^2-1)}\hat{\mu}_t^2 + \frac{\eta\rho^2\sigma^2(\phi+1)(\rho^2+1)-\eta^2\rho^4\sigma^4+\eta\rho^3\sigma^2}{2(\eta\rho^2\sigma^2+\eta\sigma^2-1)}\hat{\mu}_t \\
&+\text{a constant term}
\end{aligned} \tag{B.31}$$

$$\begin{aligned}
y_{t,2}-y_{t,1} = &\left[\frac{\eta}{2}\left(\rho^2-\frac{\rho^4}{2}-\frac{1}{2}\right)+\frac{\eta^2\rho^2\sigma^2}{2(1-\eta\sigma^2)}\right]\hat{\mu}_t^2 + \\
&\left[\frac{1}{2}(\phi\rho^2+\rho^2-2\phi\rho+\phi-\rho)-\frac{\eta\rho(\phi+1)\sigma^2}{1-\eta\sigma^2}\right]\hat{\mu}_t - \widetilde{D}^T\widetilde{\Omega}^T\widetilde{D} + \text{a constant term}
\end{aligned} \tag{B.32}$$

where

$$\begin{aligned}
\widetilde{D}^T\widetilde{\Omega}^T\widetilde{D} =&\frac{-\eta^2\rho^4\sigma^2(\rho^2+1)}{4(\eta\rho^2\sigma^2+\eta\sigma^2-1)}\hat{\mu}_t^2 + \frac{\eta\rho^2\sigma^2(\phi+1)(\rho^2+1)-\eta^2\rho^4\sigma^4+\eta\rho^3\sigma^2}{2(\eta\rho^2\sigma^2+\eta\sigma^2-1)}\hat{\mu}_t \\
&+\text{a constant term}
\end{aligned} \tag{B.33}$$

We drop the constant parts for each equation and get

$$\mathbb{E}_t\widehat{rx}_{t+1,2} = \alpha\hat{\mu}_t^2 + \psi\hat{\mu}_t \tag{B.34}$$

$$\hat{y}_{t,2}-\hat{y}_{t,1} = \xi\hat{\mu}_t^2 + \kappa\hat{\mu}_t \tag{B.35}$$

where

$$\begin{aligned}
\alpha =& -\left[\frac{\eta^2\rho^2\sigma^2(1+\rho^2)}{2(1-\eta\sigma^2)} + \frac{\eta^2\rho^4\sigma^2(\rho^2+1)}{2(\eta\rho^2\sigma^2+\eta\sigma^2-1)}\right] \\
\psi =& \frac{2\eta\rho\sigma^2(\phi\rho+\rho+\phi+1)}{2(1-\eta\sigma^2)} + \frac{\eta\rho^2\sigma^2(\phi+1)(\rho^2+1)-\eta^2\rho^4\sigma^4+\eta\rho^3\sigma^2}{\eta\rho^2\sigma^2+\eta\sigma^2-1} \\
\xi =& \frac{\eta}{2}\left(\rho^2-\frac{\rho^4}{2}-\frac{1}{2}\right) + \frac{\eta^2\rho^2\sigma^2}{2(1-\eta\sigma^2)} + \frac{\eta^2\rho^4\sigma^2(\rho^2+1)}{4(\eta\rho^2\sigma^2+\eta\sigma^2-1)} \\
\kappa =& \frac{1}{2}\left[(\phi\rho^2+\rho^2-2\phi\rho+\phi-\rho) - \frac{2\eta\rho(\phi+1)\sigma^2}{1-\eta\sigma^2}\right] - \\
& \frac{1}{2}\left[\frac{\eta\rho^2\sigma^2(\phi+1)(\rho^2+1)-\eta^2\rho^4\sigma^4+\eta\rho^3\sigma^2}{\eta\rho^2\sigma^2+\eta\sigma^2-1}\right]
\end{aligned}$$

$$\begin{aligned}
Cov(\mathbb{E}_t\widehat{rx}_{t+1,1}, y_{t,2}-y_{t,1}) =& \mathbb{E}[(\alpha\hat{\mu}_t^2+\psi\hat{\mu}_t)(\xi\hat{\mu}_t^2+\kappa\hat{\mu}_t)] - \mathbb{E}(\alpha\hat{\mu}_t^2+\psi\hat{\mu}_t)\mathbb{E}(\xi\hat{\mu}_t^2+\kappa\hat{\mu}_t) \\
=& \frac{2\alpha\xi\sigma^4}{(1-\rho^2)^2} + \frac{\psi\kappa\sigma^2}{1-\rho^2}
\end{aligned} \tag{B.36}$$

$$Var(\hat{y}_{t,2}-\hat{y}_{t,1}) = Var(\xi\hat{\mu}_t^2+\kappa\hat{\mu}_t) = \frac{2\xi^2\sigma^4}{(1-\rho^2)^2} + \frac{\kappa^2\sigma^2}{1-\rho^2} \tag{B.37}$$

$$\begin{aligned}
\beta_{2,1} = 1 + \frac{Cov(\mathbb{E}_t rx_{t+1,1}, y_{t,2}-y_{t,1})}{Var(y_{t,2}-y_{t,1})} =& 1 + \frac{Cov(\mathbb{E}_t\widehat{rx}_{t+1,1}, \hat{y}_{t,2}-\hat{y}_{t,1})}{Var(\hat{y}_{t,2}-\hat{y}_{t,1})} \\
=& 1 + \frac{2\alpha\xi\sigma^4+\psi\kappa\sigma^2(1-\rho^2)}{2\xi^2\sigma^4+\kappa^2\sigma^2(1-\rho^2)}
\end{aligned}$$

B.5 Derivation of the real yield curves under the AR(1)-Assumption for $\hat{\mu}_{t+i}$

We use "*" to indicate anything which is specific to the real term.

$$\begin{aligned}
& ny_{t,n}^* \\
=& -\log\mathbb{E}_t\exp\left[n\log\delta - \phi\hat{\mu}_{t+n} + \frac{1}{2}\eta\hat{\mu}_{t+n}^2 + \phi\hat{\mu}_t - \frac{1}{2}\eta\hat{\mu}_t^2\right] \\
=& -n\log\delta - \phi\hat{\mu}_t + \frac{1}{2}\eta\hat{\mu}_t^2 - \log\mathbb{E}_t\exp\left(\frac{1}{2}\eta\hat{\mu}_{t+n}^2 - \phi\hat{\mu}_{t+n}\right) \\
=& -n\log\delta + \phi\left(\rho^n-1\right)\hat{\mu}_t + \frac{1}{2}\eta\left(1-\rho^{2n}\right)\hat{\mu}_t^2 - \\
& \log\mathbb{E}_t\exp\left[\frac{1}{2}\eta\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\rho^{i+j-2}\varepsilon_{t+n+1-i}\varepsilon_{t+n+1-j} + 2\rho^n\hat{\mu}_t\sum_{i=1}^{n}\rho^{i-1}\varepsilon_{t+n+1-i}\right) - \phi\sum_{i=1}^{n}\rho^{i-1}\varepsilon_{t+n+1-i}\right]
\end{aligned} \tag{B.38}$$

Rewriting the equation as a matrix form, we get

$$\begin{aligned} ny^*_{t,n} =& -n\log\delta + \phi\left(\rho^n - 1\right)\hat{\mu}_t + \frac{1}{2}\eta\left(1-\rho^{2n}\right)\hat{\mu}_t^2 - \log\mathbb{E}_t\text{exp}\left(\epsilon^T A^T A\epsilon + \epsilon^T D^* + D^{*T}\epsilon\right) \\ =& -n\log\delta + \phi\left(\rho^n - 1\right)\hat{\mu}_t + \frac{1}{2}\eta\left(1-\rho^{2n}\right)\hat{\mu}_t^2 - 2D^{*T}\Omega^T D^* + \frac{1}{2}\log\left|I_n - 2\Sigma A^T A\right| \end{aligned} \tag{B.39}$$

where

$$\epsilon = \begin{pmatrix} \varepsilon_{t+1} \\ \vdots \\ \varepsilon_{t+n} \end{pmatrix} \sim N(0,\Sigma), \text{ and } \Sigma = \sigma^2 I_n, \;\; A^T A = \frac{1}{2}\eta \begin{pmatrix} \rho^{2(n-1)} & \cdots & \rho^{(n-1)} \\ \vdots & \ddots & \vdots \\ \rho^{(n-1)} & \cdots & 1 \end{pmatrix}$$

and

$$D^* = \frac{1}{2}\begin{pmatrix} (\eta\rho^n\hat{\mu}_t - \phi)\rho^{n-1} \\ (\eta\rho^n\hat{\mu}_t - \phi)\rho^{n-2} \\ \vdots \\ (\eta\rho^n\hat{\mu}_t - \phi)\rho \\ (\eta\rho^n\hat{\mu}_t - \phi) \end{pmatrix}$$

$$\frac{1}{2}\Sigma^{-1} - A^T A = \frac{1}{2}\Omega^{-1} \; i.e. \; \Omega = \left(\Sigma^{-1} - 2A^T A\right)^{-1} \qquad \text{and} \qquad D^* = \frac{1}{2}\Omega^{-1}\nu, \; i.e. \; \nu = 2\Omega D^* \tag{B.40}$$

Substituting $n = 2$ and $m = 1$ into (B.40) we get

$$\begin{aligned} & 2y^*_{t,2} \\ =& -\log\mathbb{E}_t\text{exp}\left[2\log\delta - \phi\hat{\mu}_{t+2} + \frac{1}{2}\eta\hat{\mu}_{t+2}^2 + \phi\hat{\mu}_t - \frac{1}{2}\eta\hat{\mu}_t^2\right] \\ =& -2\log\delta - \phi\hat{\mu}_t + \frac{1}{2}\eta\hat{\mu}_t^2 - \log\mathbb{E}_t\text{exp}\left(\frac{1}{2}\eta\hat{\mu}_{t+2}^2 - \phi\hat{\mu}_{t+2}\right) \\ =& -2\log\delta + \left(\phi\rho^2 - \phi\right)\hat{\mu}_t + \frac{1}{2}\eta\left(1-\rho^4\right)\hat{\mu}_t^2 - \\ & \log\mathbb{E}_t\text{exp}\left[\frac{1}{2}\eta\left(\sum_{i=1}^{2}\sum_{j=1}^{2}\rho^{i+j-2}\varepsilon_{t+3-i}\varepsilon_{t+3-j} + 2\rho^2\hat{\mu}_t\sum_{i=1}^{2}\rho^{i-1}\varepsilon_{t+3-i}\right) - \phi\sum_{i=1}^{2}\rho^{i-1}\varepsilon_{t+3-i}\right] \end{aligned} \tag{B.41}$$

When we rewrite (B.41) as a matrix form, we get

$$
\begin{aligned}
&2y^*_{t,2}\\
=&-2\log\delta+\phi\left(\rho^2-1\right)\hat{\mu}_t+\frac{1}{2}\eta\left(1-\rho^4\right)\hat{\mu}_t^2-\log\mathbb{E}_{t+m}\exp\left(\widetilde{\epsilon}^T\widetilde{A}^T\widetilde{A}\widetilde{\epsilon}+\widetilde{\epsilon}^T\widetilde{D^*}+\widetilde{D^*}^T\widetilde{\epsilon}\right)\\
=&-2\log\delta+\phi\left(\rho^2-1\right)\hat{\mu}_t+\frac{1}{2}\eta\left(1-\rho^4\right)\hat{\mu}_t^2-\log\left[\frac{1}{|\widetilde{\Sigma}|^{\frac{1}{2}}}\frac{1}{\left|\widetilde{\Sigma}^{-1}-2\widetilde{A}^T\widetilde{A}\right|^{\frac{1}{2}}}\exp\left(\widetilde{D^*}^T\widetilde{\Omega}^T\widetilde{D^*}\right)\right]\\
=&-2\log\delta+\phi\left(\rho^2-1\right)\hat{\mu}_t+\frac{1}{2}\eta\left(1-\rho^4\right)\hat{\mu}_t^2-2\widetilde{D^*}^T\widetilde{\Omega}^T\widetilde{D^*}+\frac{1}{2}\log\left(1-\eta\sigma^2-\eta\sigma^2\rho^2\right)
\end{aligned}
\tag{B.42}
$$

$$
\begin{aligned}
y^*_{t,1}=&-\log\mathbb{E}_t\exp\left[\log\delta-\phi\hat{\mu}_{t+1}+\frac{1}{2}\eta\hat{\mu}_{t+1}^2+\phi\hat{\mu}_t-\frac{1}{2}\eta\hat{\mu}_t^2\right]\\
=&-\log\delta-\phi\hat{\mu}_t+\frac{1}{2}\eta\hat{\mu}_t^2-\log\mathbb{E}_t\exp\left[\frac{1}{2}\eta\hat{\mu}_{t+1}^2-\phi\hat{\mu}_{t+1}\right]\\
=&-\log\delta+\phi(\rho-1)\hat{\mu}_t+\frac{1}{2}\eta(1-\rho^2)\hat{\mu}_t^2-\log\mathbb{E}_t\exp\left[(\eta\rho\hat{\mu}_t-\phi)\varepsilon_{t+1}+\frac{1}{2}\eta\varepsilon_{t+1}^2\right]\\
=&-\log\delta+\phi(\rho-1)\hat{\mu}_t+\frac{1}{2}\eta(1-\rho^2)\hat{\mu}_t^2-\frac{(\eta\rho\hat{\mu}_t-\phi)^2\sigma^2}{2(1-\eta\sigma^2)}+\frac{1}{2}\log(1-\eta\sigma^2)
\end{aligned}
\tag{B.43}
$$

$$
\begin{aligned}
&y^*_{t+1,1}\\
=&-\log\mathbb{E}_{t+1}\exp\left[\log\delta-\phi\hat{\mu}_{t+2}+\frac{1}{2}\eta\hat{\mu}_{t+2}^2+\phi\hat{\mu}_{t+1}-\frac{1}{2}\eta\hat{\mu}_{t+1}^2\right]\\
=&-\log\delta-\phi\hat{\mu}_{t+1}+\frac{1}{2}\eta\hat{\mu}_{t+1}^2-\log\mathbb{E}_{t+1}\exp\left[\frac{1}{2}\eta\hat{\mu}_{t+2}^2-\phi\hat{\mu}_{t+2}\right]\\
=&-\log\delta+\phi(\rho-1)\hat{\mu}_{t+1}+\frac{1}{2}\eta(1-\rho^2)\hat{\mu}_{t+1}^2-\log\mathbb{E}_{t+1}\exp\left[(\eta\rho\hat{\mu}_{t+1}-\phi)\varepsilon_{t+2}+\frac{1}{2}\eta\varepsilon_{t+2}^2\right]\\
=&-\log\delta+\phi(\rho-1)\hat{\mu}_{t+1}+\frac{1}{2}\eta(1-\rho^2)\hat{\mu}_{t+1}^2-\frac{(\eta\rho\hat{\mu}_{t+1}-\phi)^2\sigma^2}{2(1-\eta\sigma^2)}+\frac{1}{2}\log(1-\eta\sigma^2)
\end{aligned}
\tag{B.44}
$$

$$
\begin{aligned}
y^*_{t,1}-2y^*_{t,2}=&2\widetilde{D^*}^T\widetilde{\Omega}^T\widetilde{D^*}+\log\delta+\phi\rho(1-\rho)\hat{\mu}_t-\frac{1}{2}\eta(\rho^2-\rho^4)\hat{\mu}_t^2-\frac{(\eta\rho\hat{\mu}_t-\phi)^2\sigma^2}{2(1-\eta\sigma^2)}\\
&-\frac{1}{2}\log\left(1-\eta\sigma^2-\eta\sigma^2\rho^2\right)+\frac{1}{2}\log(1-\eta\sigma^2)
\end{aligned}
\tag{B.45}
$$

$$
\begin{aligned}
&\mathbb{E}_t rx^*_{t+1,2} \\
=&\mathbb{E}_t(y^*_{t+1,1}) + y^*_{t,1} - 2y^*_{t,2} \\
=&-\frac{(\eta\rho\hat{\mu}_t - \phi)^2\sigma^2 + (\eta\rho^2\hat{\mu}_t - \phi)^2\sigma^2 + \eta^2\rho^2\sigma^4}{2(1-\eta\sigma^2)} + \log(1-\eta\sigma^2) + \frac{1}{2}\eta(1-\rho^2)\sigma^2 \\
&+ 2\widetilde{D^*}^T\Omega^T\widetilde{D^*} - \frac{1}{2}\log\left(1-\eta\sigma^2-\sigma^2\rho^2\right) \\
=&-\frac{\eta^2\rho^2\sigma^2(1+\rho^2)}{2(1-\eta\sigma^2)}\hat{\mu}_t^2 + \frac{2\eta\rho\sigma^2\phi(\rho+1)}{2(1-\eta\sigma^2)}\hat{\mu}_t + 2\widetilde{D^*}^T\Omega^T\widetilde{D^*} + \text{a constant term}
\end{aligned} \tag{B.46}
$$

$$
\begin{aligned}
y^*_{t,2} - y^*_{t,1} = &\left[\frac{\eta}{2}\left(\rho^2 - \frac{\rho^4}{2} - \frac{1}{2}\right) + \frac{\eta^2\rho^2\sigma^2}{2(1-\eta\sigma^2)}\right]\hat{\mu}_t^2 + \\
&\left[\frac{1}{2}(\phi\rho^2 - 2\phi\rho + \phi) - \frac{\eta\rho\phi\sigma^2}{1-\eta\sigma^2}\right]\hat{\mu}_t - \widetilde{D^*}^T\widetilde{\Omega}^T\widetilde{D^*} + \text{a constant term}
\end{aligned} \tag{B.47}
$$

where

$$
\begin{aligned}
\widetilde{D^*}^T\widetilde{\Omega}^T\widetilde{D^*} =&\frac{\phi\rho - \eta\rho^3\hat{\mu}_t}{2}\left[\frac{(\phi\rho - \eta\rho^3\hat{\mu}_t)(\eta\sigma^2-1)\sigma^2}{2(\eta\rho^2\sigma^2+\eta\sigma^2-1)} - \frac{\eta\rho\sigma^4(\phi - \eta\rho^2\hat{\mu}_t)}{2(\eta\rho^2\sigma^2+\eta\sigma^2-1)}\right] \\
&- \frac{\phi - \eta\rho^2\hat{\mu}_t}{2}\left[\frac{(\phi - \eta\rho^2\hat{\mu}_t)(\sigma^2 - \eta\rho^2\sigma^4)}{2(\eta\rho^2\sigma^2+\eta\sigma^2-1)} + \frac{\eta\rho\sigma^4(\phi\rho - \eta\rho^3\hat{\mu}_t)}{2(\eta\rho^2\sigma^2+\eta\sigma^2-1)}\right] \\
=&-\frac{\eta^2\rho^4\sigma^2(\rho^2+1)}{4(\eta\rho^2\sigma^2+\eta\sigma^2-1)}\hat{\mu}_t^2 + \frac{\eta\rho^2\sigma^2\phi(\rho^2+1)}{2(\eta\rho^2\sigma^2+\eta\sigma^2-1)}\hat{\mu}_t + \text{a constant term}
\end{aligned} \tag{B.48}
$$

Dropping the constant parts for each equation,

$$
\mathbb{E}_t\widehat{rx}^*_{t+1,2} = \alpha\hat{\mu}_t^2 + \psi^*\hat{\mu}_t \tag{B.49}
$$

$$
\hat{y}^*_{t,2} - \hat{y}^*_{t,1} = \xi\hat{\mu}_t^2 + \kappa^*\hat{\mu}_t \tag{B.50}
$$

where

$$
\begin{aligned}
\alpha =& -\left[\frac{\eta^2\rho^2\sigma^2(1+\rho^2)}{2(1-\eta\sigma^2)} + \frac{\eta^2\rho^4\sigma^2(\rho^2+1)}{2(\eta\rho^2\sigma^2+\eta\sigma^2-1)}\right] \\
\psi^* =& \frac{2\eta\rho\sigma^2\phi(\rho+1)}{2(1-\eta\sigma^2)} + \frac{\eta\rho^2\sigma^2\phi(\rho^2+1)}{\eta\rho^2\sigma^2+\eta\sigma^2-1} \\
\xi =& \frac{\eta}{2}\left(\rho^2 - \frac{\rho^4}{2} - \frac{1}{2}\right) + \frac{\eta^2\rho^2\sigma^2}{2(1-\eta\sigma^2)} + \frac{\eta^2\rho^4\sigma^2(\rho^2+1)}{4(\eta\rho^2\sigma^2+\eta\sigma^2-1)} \\
\kappa^* =& \frac{1}{2}\left[(\phi\rho^2 - 2\phi\rho + \phi) - \frac{2\eta\rho\phi\sigma^2}{1-\eta\sigma^2}\right] - \frac{1}{2}\left[\frac{\eta\rho^2\sigma^2\phi(\rho^2+1)}{\eta\rho^2\sigma^2+\eta\sigma^2-1}\right]
\end{aligned}
$$

$$\begin{aligned} Cov(\mathbb{E}_t\widehat{rx}^*_{t+1,1}, y^*_{t,2} - y^*_{t,1}) =& \mathbb{E}[(\alpha\hat{\mu}_t^2 + \psi^*\hat{\mu}_t)(\xi\hat{\mu}_t^2 + \kappa^*\hat{\mu}_t)] - \mathbb{E}(\alpha\hat{\mu}_t^2 + \psi^*\hat{\mu}_t)\mathbb{E}(\xi\hat{\mu}_t^2 + \kappa^*\hat{\mu}_t) \\ =& \frac{2\alpha\xi\sigma^4}{(1-\rho^2)^2} + \frac{\psi^*\kappa^*\sigma^2}{1-\rho^2} \end{aligned} \tag{B.51}$$

$$Var(\hat{y}_{t,2} - \hat{y}_{t,1}) = Var(\xi\hat{\mu}_t^2 + \kappa^*\hat{\mu}_t) = \frac{2\xi^2\sigma^4}{(1-\rho^2)^2} + \frac{\kappa^{*2}\sigma^2}{1-\rho^2} \tag{B.52}$$

$$\begin{aligned} \beta^*_{2,1} = 1 + \frac{Cov(\mathbb{E}_t rx^*_{t+1,1}, y^*_{t,2} - y^*_{t,1})}{Var(y^*_{t,2} - y^*_{t,1})} =& 1 + \frac{Cov(\mathbb{E}_t\widehat{rx^*}_{t+1,1}, \hat{y}^*_{t,2} - \hat{y}^*_{t,1})}{Var(\hat{y}_{t,2} - \hat{y}_{t,1})} \\ =& 1 + \frac{2\alpha\xi\sigma^4 + \psi^*\kappa^*\sigma^2(1-\rho^2)}{2\xi^2\sigma^4 + \kappa^{*2}\sigma^2(1-\rho^2)} \end{aligned}$$

References

ALVAREZ, F., A. ATKESON, AND P. J. KEHOE (2002): "Money, Interest Rates, and Exchange Rates with Endogenously Segmented Markets," *Journal of Political Economy*, 110(1), 73–112.

——— (2009): "Time-Varying Risk, Interest Rates, and Exchange Rates in General Equilibrium," *Review of Economic Studies*, 76(3), 851–878.

ALVAREZ, F., L. GUISO, AND F. LIPPI (2012): "Durable Consumption and Asset Management with Transaction and Observation Costs," *American Economic Review*, 102(5), 2272–2300.

ALVAREZ, F., AND U. J. JERMANN (2001): "Quantitative Asset Pricing Implications of Endogenous Solvency Constraints," *Review of Financial Studies*, 14, 1117–1151.

BACKUS, D., S. FORESI, A. MOZUMDAR, AND L. WU (2001): "Predictable changes in yields and forward rates," *Journal of Financial Economics*, 59, 281–311.

BACKUS, D. K., A. W. GREGORY, AND S. E. ZIN (1989): "Risk Premiums in the Term Structure Evidence from Artificial Economics," *Journal of Monetary Economics*, 24, 371–399.

BANSAL, R., AND I. SHALIASTOVICH (2007): "Risk and Return in Bond, Currency and Equity Markets," *Working Paper*.

——— (2012): "A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets," *NBER Working Paper*.

BANSAL, R., AND H. ZHOU (2002): "Term Structure of Interest Rates with Regime Shifts," *The Journal of Finance*, 57, 1997–2043.

BEKAERT, G., AND R. J. HODRICK (2001): "Expectations Hypotheses Tests," *The Journal of Finance*, 56, 1357–1394.

BEKAERT, G., R. J. HODRICK, AND D. A. MARSHALL (1997): "The Implications of First-Order Risk Aversion for Asset Market Risk Premiums," *Journal of Monetary Economics*, 40, 3–39.

——— (2001): "Peso Problem Explanations for Term Structure Anomalies," *Journal of Monetary Economics*, 48, 241–270.

BRANDT, M. W., AND K. Q. WANG (2003): "Time-varying risk aversion and unexpected inflation," *Journal of Monetary Economics*, 50, 1457–1498.

BURASCHI, A., AND A. JILTSOV (2005): "Inflation risk premia and the expectations hypothesis," *Journal of Financial Economics*, 75, 429–490.

CAMPBELL, J. Y., AND R. J. SHILLER (1991): "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *The Review of Economic Studies*, 58 (3), 495–514.

CECCHETTI, S. G., AND G. DEBELLE (2006): "Has the Inflation Process Changed?," *Economic Policy*, 21(46), 311–352.

CHATTERJEE, S., AND D. CORBAE (1992): "Endogenous Market Participation and the General Equilibrium Value of Money," *Journal of Political Economy*, 100, 615–646.

CHERNOV, M., AND P. MUELLER (2012): "The term structure of inflation expectations," *Journal of Financial Economics*, 106, 367–394.

DEN HAAN, W. J. (1995): "The Term Structure of Interest Rates in Real and Monetary Economies," *Journal of Economic Dynamics and Control*, 19, 909–940.

DOMINGUEZ, E., AND A. NOVALES (2000): "Testing the expectations hypothesis in Eurodeposits," *Journal of International Money and Finance*, 19, 713–736.

DONALDSON, J. B., T. JOHNSEN, AND R. MEHRA (1990): "On the Term Structure of Interest Rates," *Journal of Economic Dynamics and Control*, 14, 571–596.

DUFFEE, G. R. (2011): "Information in (and not in) the Term Structure," *The Review of Financial Studies*, 24, 2895–2934.

EVANS, M. D., AND K. K. LEWIS (1994): "Do stationary risk premia explain it all? Evidence from the term structure," *Journal of Monetary Economics*, 33, 285–318.

FAMA, E. F. (1984): "Forward and Spot Exchange Rates," *Journal of Monetary Economics*, 14, 319–338.

FAMA, E. F., AND R. R. BLISS (1987): "The Information in Long-Maturity Forward Rates," *The American Economic Review*, 77, 680–692.

GERLACH, S., AND F. SMETS (1997): "The term structure of Euro-rates: some evidence in support of the expectations hypothesis," *Journal of lnternational Money and Finance*, 16, 305–321.

GUST, C., AND D. LOPEZ-SALIDO (2014): "Monetary policy and the cyclicality of risk," *Journal of Monetary Economics*, 62, 59–75.

HARDOUVELIS, G. A. (1994): "The term structure spread and future changes in long and short rates in the G7 countries: Is there a puzzle?," *Journal of Monetary Economics*, 33, 255–283.

JONGEN, R., W. F. VERSCHOOR, AND C. C. WOLFF (2011): "Time-variation in term premia: International survey-based evidence," *Journal of International Money and Finance*, 30, 605–622.

KEHOE, P. J., AND F. PERRI (2002): "International Business Cycles with endogenous Incomplete Markets," *Econometrica*, 70, 907–928.

KEHOE, T. J., AND D. K. LEVINE (1993): "Debt-Constrained Asset Markets," *Review of Economic Studies*, 60, 865–888.

LONGSTAFF, F. A. (2000): "The term structure of very short-term rates: New evidence for the expectations hypothesis," *Journal of Financial Economics*, 58, 397–415.

LUCAS, R. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 6, 1429–1445.

REIS, R. (2006): "Inattentive consumers," *Journal of Monetary Economics*, 53, 1761–1800.

RUDEBUSCH, G., AND E. SWANSON (2012): "The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks," *American Economic Journal: Macroeconomics*, 4, 105–143.

RUDEBUSCH, G. D., AND T. WU (2008): "A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy," *The Economic Journal*, 118, 906–926.

SEPPALA, J. (2004): "The term structure of real interest rates: theory and evidence from UK index-linked bonds," *Journal of Monetary Economics*, 51, 1509–1549.

SINHA, A. (2009): "Learning and the Yield Curve," *Working Paper of Columbia University*.

WACHTER, J. A. (2006): "A consumption-based model of the term structure of interest rates," *Journal of Financial Economics*, 79(2), 365–399.

WALSH, C. E. (2010): "Monetary Theory and Policy (3rd ed.)," *The MIT Press*.

Table 1: Slope coefficient estimates for equation (4.3)

Country	$\beta_{2,1}$	$\beta_{3,1}$	$\beta_{4,1}$	$\beta_{6,1}$	$\beta_{9,1}$	$\beta_{12,1}$	$\beta_{4,3}$	$\beta_{6,3}$	$\beta_{9,3}$	$\beta_{12,3}$	$\beta_{9,6}$	$\beta_{12,6}$
Aus	-0.486	-0.546	-0.761	-1.127	-2.543	-3.536	-0.236	-0.757	-1.625	-2.651	-1.487	-2.419
	(0.308)	(0.461)	(0.517)	(0.610)	(0.832)	(1.048)	(0.234)	(0.354)	(0.518)	(0.660)	(0.340)	(0.429)
Bel	0.017	0.153	–	–	–	–	–	-0.490	-1.242	-2.305	-1.946	-3.169
	(0.296)	(0.399)						(0.503)	(0.744)	(0.813)	(0.731)	(0.651)
Can	0.208	0.788	–	–	–	–	–	0.193	-0.038	-0.266	0.320	0.053
	(0.396)	(0.592)						(0.539)	(0.657)	(0.773)	(0.544)	(0.629)
Fin	0.335	0.698	-3.110	-2.486	-3.181	-4.688	-0.360	0.453	0.665	-0.035	0.198	-0.521
	(0.285)	(0.456)	(0.288)	(1.155)	(1.475)	(1.431)	0.039	(0.429)	(0.632)	(0.658)	(0.474)	(0.564)
Fr	0.995	2.000	2.450	2.589	2.693	2.921	1.642	2.517	2.846	2.777	1.185	1.162
	(0.462)	(0.652)	(0.778)	(0.787)	(0.791)	(0.880)	(0.405)	(0.443)	(0.452)	(0.521)	(0.392)	(0.521)
Ger	0.290	0.693	0.884	1.136	1.102	0.896	0.582	0.929	0.815	0.197	0.232	-0.693
	(0.208)	(0.333)	(0.413)	(0.577)	(0.837)	(1.095)	(0.189)	(0.382)	(0.622)	(0.785)	(0.542)	(0.657)
It	-0.017	0.150	0.172	0.407	0.688	1.202	0.540	0.562	0.802	0.959	0.395	0.404
	(0.296)	(0.433)	(0.495)	(0.567)	(0.666)	(0.747)	(0.204)	(0.276)	(0.333)	(0.410)	(0.267)	(0.353)
Jap	-0.042	-0.100	-0.897	-0.017	-0.435	-1.093	-0.256	0.265	-0.231	-0.841	-0.312	-0.660
	(0.291)	(0.477)	(0.624)	(0.894)	(1.21)	(1.520)	(0.202)	(0.654)	(0.842)	(0.956)	(0.584)	(0.655)
Neth	0.226	0.590	0.810	0.805	0.475	-0.223	0.437	0.627	-0.174	-1.390	-1.248	-2.285
	(0.236)	(0.395)	(0.493)	(0.730)	(1.071)	(1.266)	(0.230)	(0.496)	(0.798)	(0.917)	(0.623)	(0.652)
Nor	-0.907	-0.387	-2.99	-1.389	-5.263	5.847	-0.245	-0.052	-0.181	-1.090	0.036	-0.517
	(0.095)	(0.146)	(0.207)	(3.680)	(0.656)	(4.217)	(0.033)	(0.387)	(0.447)	(0.371)	(0.439)	(0.230)
Nzl	0.524	0.703	0.174	0.315	-0.412	-0.729	0.211	0.484	0.392	0.261	0.731	0.781
	(0.694)	(0.767)	(0.721)	(0.827)	(0.969)	(1.021)	(0.403)	(0.627)	(0.724)	(0.799)	(0.779)	(0.821)
Port	-0.450	-0.384	–	–	–	–	–	-0.549	-0.750	-0.823	-1.771	-2.271
	(0.273)	(0.301)						(0.314)	(0.442)	(0.552)	(0.408)	(0.511)
Spa	0.325	0.598	0.667	0.898	1.318	1.830	0.664	0.564	0.638	0.470	0.141	-0.171
	(0.212)	(0.334)	(0.415)	(0.532)	(0.719)	(0.912)	(0.186)	(0.235)	(0.339)	(0.404)	(0.347)	(0.382)
Swe	-0.152	0.294	-2.633	-0.678	-2.624	-3.754	-0.315	2.208	1.572	1.649	0.655	0.767
	(0.245)	(0.358)	(0.272)	(1.166)	(1.496)	(1.788)	(0.038)	(0.562)	(0.765)	(0.889)	(0.612)	(0.753)
Swit	0.154	0.091	0.247	0.232	-0.386	0.146	0.299	0.574	0.211	-0.023	-1.276	-2.114
	(0.319)	(0.485)	(0.589)	(0.812)	(1.196)	(1.756)	(0.253)	(0.544)	(0.944)	(1.179)	(0.819)	(1.190)
UK	-0.130	0.201	0.610	0.626	0.764	0.493	0.937	1.011	1.081	0.805	0.949	0.756
	(0.241)	(0.319)	(0.409)	(0.464)	(0.590)	(0.735)	(0.244)	(0.309)	(0.435)	(0.510)	(0.389)	(0.448)
US	-0.632	0.371	0.223	0.260	-0.553	-1.562	-0.076	0.066	-0.453	-1.115	-0.501	-0.982
	(0.130)	(0.260)	(0.242)	(0.400)	(0.627)	(0.824)	(0.103)	(0.217)	(0.392)	(0.513)	(0.287)	(0.357)

Estimates of equation (4.3) with Eurocurrency rates for 17 countries. Monthly observations of 1-12 month Euro-rates for Feb 1995 to Jan 2000 are used. Hansen-Hodrick standard errors are below estimated coefficients, in parentheses. For each regression, the sample is the longest possible using data. "–" indicates data is not available here. Data are from Datastream under the fold of "Garban Information Services Eurocurrency Rates". The abbreviations for countries used here are Australia (Aus), Belgium (Bel), Canada (Can), Finland (Fin), France (Fr), Germany (Ger), Italy (It), Japan (Jap), The Netherlands (Neth), Norway (Nor), New Zealand (Nzl), Portugal (Port), Spain (Spa), Sweden (Swe), Switzerland (Swit), United Kingdom (UK), and United States of America (US).

Table 2: Theoretical Values of $\beta_{n,1}$ under MA(1)-Money Growth

σ	$\beta_{n,1}$	θ								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.0033	$\beta_{2,1}$	1.003	1.007	1.011	1.008	0.999	0.992	0.990	0.989	0.989
	$\beta_{4,1}$	1.002	1.004	1.006	1.008	1.007	1.004	1.001	0.998	0.996
	$\beta_{6,1}$	1.002	1.003	1.005	1.007	1.007	1.005	1.002	1.000	0.998
	$\beta_{9,1}$	1.001	1.003	1.005	1.006	1.006	1.005	1.003	1.001	0.999
0.0115	$\beta_{2,1}$	1.033	1.078	1.110	1.048	0.930	0.869	0.850	0.847	0.849
	$\beta_{4,1}$	1.021	1.046	1.073	1.087	1.069	1.024	0.977	0.944	0.924
	$\beta_{6,1}$	1.019	1.041	1.064	1.079	1.074	1.045	1.006	0.973	0.950
	$\beta_{9,1}$	1.017	1.038	1.059	1.075	1.073	1.052	1.019	0.988	0.965

Table 3: Theoretical Values of $\beta_{2,1}$ under AR(1)-Money Growth

Panel A: Regression slope coefficients for nominal yields, $\beta_{2,1}$

σ	ϕ	ρ								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.0033	10	1.003	1.007	1.013	1.022	1.035	1.054	1.080	1.115	1.156
	20	1.003	1.007	1.014	1.024	1.040	1.067	1.114	1.203	1.307
	30	1.003	1.007	1.014	1.024	1.041	1.069	1.122	1.238	1.515
0.0115	10	1.023	1.051	1.086	1.133	1.202	1.311	1.512	1.838	-1.810
	20	1.033	1.081	1.148	1.239	1.352	1.456	1.399	0.681	-0.313
	30	1.037	1.093	1.179	1.312	1.501	1.674	1.267	-0.146	-0.029
0.0180	10	1.044	1.098	1.170	1.273	1.416	1.472	0.104	-3.283	-1.357
	20	1.085	1.195	1.314	1.365	1.118	0.335	-0.448	-3.369	-1.308
	30	1.106	1.269	1.480	1.524	0.720	-0.369	-0.602	-3.348	-1.278

Panel B: Regression slope coefficients for real yields, $\beta^*_{2,1}$

σ	ϕ	ρ								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.0033	10	1.003	1.006	1.012	1.020	1.031	1.049	1.079	1.128	1.264
	20	1.003	1.007	1.013	1.022	1.037	1.061	1.104	1.188	1.409
	30	1.003	1.007	1.013	1.023	1.038	1.065	1.113	1.215	1.529
0.0115	10	1.021	1.047	1.082	1.132	1.208	1.340	1.607	2.207	-3.018
	20	1.032	1.077	1.141	1.230	1.351	1.497	1.576	0.889	-1.239
	30	1.035	1.089	1.172	1.300	1.489	1.711	1.524	-0.134	-0.711
0.0180	10	1.041	1.095	1.173	1.297	1.503	1.747	0.026	-2.915	-1.358
	20	1.081	1.188	1.314	1.407	1.264	0.470	-0.621	-2.755	-1.305
	30	1.103	1.260	1.472	1.578	0.899	-0.355	-0.735	-2.489	-1.272

Table 4: Average Values of Simulated β under i.i.d-Money Growth

$n =$	2	3	4	5	6	7	8	9	10
$\bar{\beta}_{n,1} =$	1.009	1.006	1.006	1.005	1.005	1.005	1.005	1.005	1.005
$\bar{\beta}_{n,2} =$	–	0.996	0.998	0.998	0.998	0.998	0.998	0.998	0.998

where $\eta = 1000$, $\phi = 20$, $\sigma = 0.0115$, $\bar{\mu} = 1.05$, $\delta = 0.95$.

Table 5: Average Values of Simulated β under MA(1)-Money Growth

$n=$	2	3	4	5	6	7	8	9	10
$\bar{\beta}_{n,1}=$	0.857	0.897	0.924	0.939	0.949	0.955	0.960	0.963	0.966
$\bar{\beta}_{n,2}=$	–	1.037	1.018	1.015	1.014	1.013	1.012	1.012	1.012

where $\eta = 1000$, $\phi = 20$, $\sigma = 0.0115$, $\bar{\mu} = 1.05$, $\delta = 0.95$, $\theta = 0.9$.

Table 6: Average Values of Simulated $\beta_{n,m}$ under AR(1)-Money Growth

Panel A: Regression slope coefficients for nominal yields, $\bar{\beta}_{n,m}$

σ	ρ	m	n								
			2	3	4	5	6	7	8	9	10
0.0080	0.6	1	1.458	1.375	1.326	1.294	1.272	1.256	1.244	1.235	1.228
		2	–	1.301	1.240	1.204	1.181	1.165	1.154	1.146	1.140
	0.9	1	1.088	1.102	1.123	1.148	1.175	1.204	1.234	1.263	1.292
		2	–	1.164	1.174	1.188	1.206	1.226	1.248	1.270	1.293
0.0115	0.6	1	1.565	1.487	1.439	1.407	1.385	1.369	1.357	1.348	1.340
		2	–	1.444	1.369	1.322	1.291	1.270	1.255	1.243	1.234
	0.9	1	0.214	0.160	0.065	-0.036	-0.129	-0.207	-0.266	-0.308	-0.335
		2	–	0.471	0.327	0.191	0.067	-0.032	-0.115	-0.178	-0.224

Panel B: Regression slope coefficients for real yields, $\bar{\beta}^*{}_{n,m}$

σ	ρ	m	n								
			2	3	4	5	6	7	8	9	10
0.0080	0.6	1	1.440	1.345	1.326	1.292	1.259	1.238	1.223	1.212	1.204
		2	–	1.261	1.203	1.171	1.151	1.138	1.129	1.122	1.117
	0.9	1	1.550	1.690	1.800	1.884	1.948	1.995	2.029	2.053	2.070
		2	–	1.817	1.896	1.953	1.993	2.020	2.036	2.046	2.051
0.0115	0.6	1	1.621	1.531	1.473	1.435	1.409	1.389	1.375	1.363	1.354
		2	–	1.449	1.367	1.316	1.283	1.261	1.246	1.234	1.225
	0.9	1	-1.073	-1.095	-1.116	-1.132	-1.140	-1.138	-1.126	-1.103	-1.072
		2	–	-0.733	-0.787	-0.834	-0.872	-0.898	-0.912	-0.913	-0.903

where $\eta = 1000$, $\phi = 20$, $\bar{\mu} = 1.05$, $\delta = 0.95$.

Table 7: Values of Simulated β under AR(4)-Money Growth

Panel A: Regression slope coefficients for nominal yields, $\bar{\beta}_{n,m}$

$n =$	2	3	4	5	6	7	8	9	10
$\beta_{n,1} =$	-0.492	-0.589	-0.188	-0.009	-0.458	-0.584	-0.419	-0.382	-0.415
$\beta_{n,2} =$	–	-0.368	-0.434	-0.277	-0.324	-0.445	-0.481	-0.495	-0.339

Panel B: Regression slope coefficients for real yields, $\bar{\beta^*}_{n,m}$

$n =$	2	3	4	5	6	7	8	9	10
$\beta_{n,1} =$	-0.960	-0.667	-0.208	0.103	-0.174	-0.140	-0.010	-0.003	0.072
$\beta_{n,2} =$	–	-1.381	-0.960	-0.182	-0.064	-0.336	-0.206	-0.019	-0.004

Parameter value: $\eta = 1000$, $\phi = 10$, $\sigma = 0.008$, $\bar{\mu} = 1.05$, $\delta = 0.95$, $\rho_1 = 0.3$, $\rho_2 = 0.2$, $\rho_3 = 0.2$, $\rho_4 = 0.2$.

Table 8: Values of Simulated β under AR(12)-Money Growth

Panel A: Regression slope coefficients for nominal yields, $\bar{\beta}_{n,m}$

$n =$	2	3	4	5	6	7	8	9	10
$\beta_{n,1} =$	-0.853	-0.525	-0.091	0.043	-0.085	0.005	0.079	-0.110	-0.699
$\beta_{n,2} =$	–	-1.287	-0.779	-0.168	-0.120	-0.182	-0.090	-0.613	-0.181

Panel B: Regression slope coefficients for real yields, $\bar{\beta^*}_{n,m}$

$n =$	2	3	4	5	6	7	8	9	10
$\beta_{n,1} =$	-0.854	-0.499	-0.156	-0.030	-0.150	-0.100	-0.054	-0.038	0.004
$\beta_{n,2} =$	–	-1.362	-0.793	-0.243	-0.203	-0.294	-0.178	-0.122	-0.051

Parameter value: $\eta = 1000$, $\phi = 10$, $\sigma = 0.008$, $\bar{\mu} = 1.05$, $\delta = 0.95$, $\rho_1 = 0.1$, $\rho_2 = 0.08$, $\rho_3 = 0.08$, $\rho_4 = 0.05$, $\rho_5 = 0.05$, $\rho_6 = 0.05$, $\rho_7 = 0.05$, $\rho_8 = 0.05$, $\rho_9 = 0.05$, $\rho_{10} = 0.05$, $\rho_{11} = 0.05$, $\rho_{12} = 0.05$,

Table 9: Values of Inflation Parameter

Country	$\hat{\sigma}$	$\hat{\rho}$	Prob. of $\hat{\rho} = 0$
Bel	0.0027	0.072	0.440
Can	0.0019	0.188	0.072
Fin	0.0029	0.133	0.153
Fr	0.0021	0.127	0.171
Ger	0.0031	0.141	0.130
It	0.0021	0.452	0.000
Jap	0.0042	0.209	0.024
Neth	0.0037	0.180	0.054
Nor	0.0030	0.027	0.769
Port	0.0048	0.406	0.000
Spa	0.0032	0.101	0.276
Swe	0.0058	0.256	0.005
Swit	0.0024	0.432	0.000
UK	0.0051	0.160	0.085
US	0.0018	0.264	0.004

Table 10: Results of Regression for Eq. (6.1)

(n, m)	$b_{n,m}$	$d_{n,m}$	prob. of $b_{n,m}$=0	prob. of $d_{n,m}$=0	prob. of F-stat.	R^2
$(2, 1)$	-1.615	-122.269	0.112	0.161	0.143	0.322
$(3, 1)$	-2.743	-235.523	0.035	0.037	0.028	0.512
$(4, 1)$	-1.524	-583.460	0.762	0.211	0.428	0.191
$(6, 1)$	-0.562	-253.776	0.894	0.499	0.777	0.061
$(9, 1)$	-2.467	-531.937	0.657	0.292	0.553	0.137
$(12, 1)$	-0.628	-675.853	0.924	0.319	0.562	0.134
$(6, 3)$	-1.675	70.987	0.469	0.722	0.686	0.073
$(9, 3)$	-2.062	-19.992	0.445	0.931	0.736	0.059
$(12, 3)$	-0.620	19.928	0.853	0.945	0.978	0.005
$(9, 6)$	-3.641	-69.560	0.132	0.726	0.301	0.213
$(12, 6)$	-3.487	-22.527	0.282	0.934	0.543	0.115

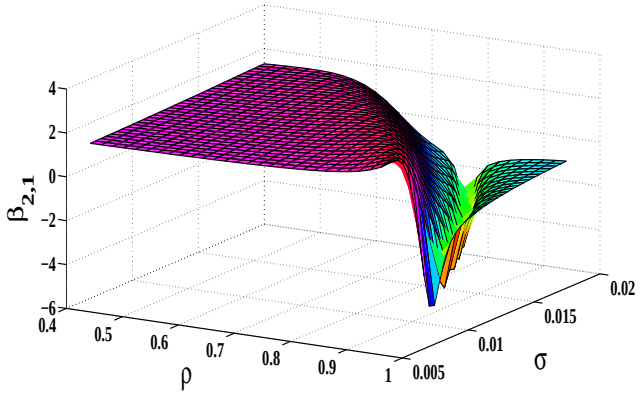

Figure 1: Comovement between $\beta_{2,1}$ and σ and ρ

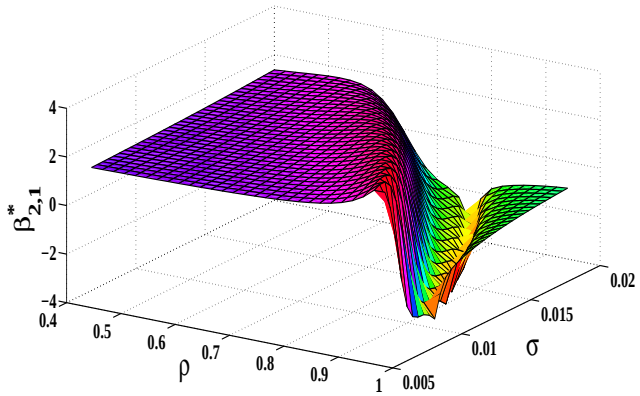

Figure 2: Comovement between $\beta^*_{2,1}$ and σ and ρ

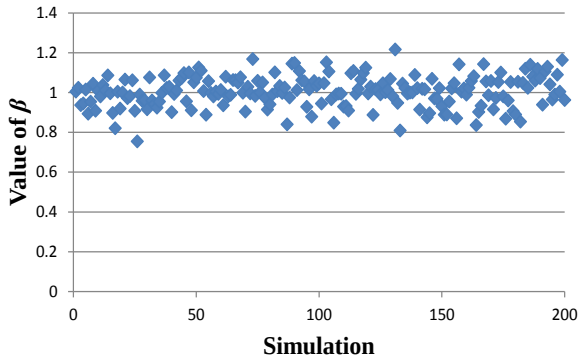

Figure 3: $\beta_{10,1}$ under i.i.d-inflation ($\bar{\beta}_{10,1} = 1.005 \pm 0.077$)

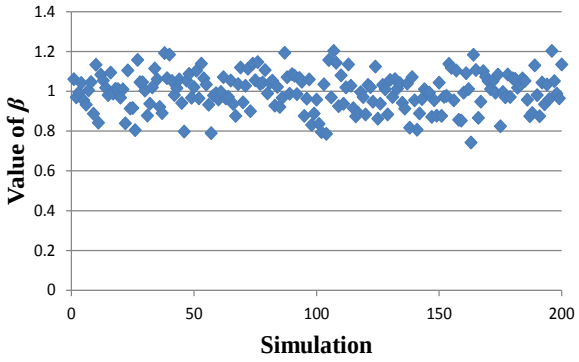

Figure 4: $\beta_{10,2}$ under i.i.d-inflation ($\bar{\beta}_{10,2} = 0.998 \pm 0.093$)

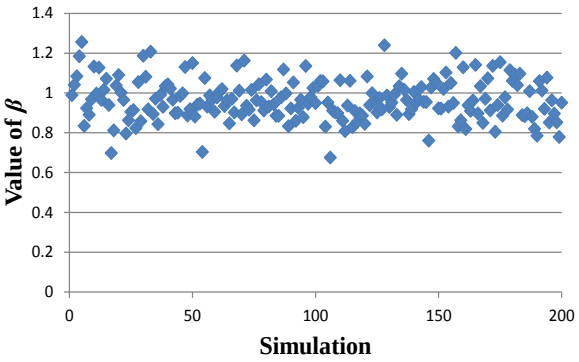

Figure 5: $\beta_{10,1}$ under MA(1)-inflation ($\bar{\beta}_{10,1} = 0.966 \pm 0.102$)

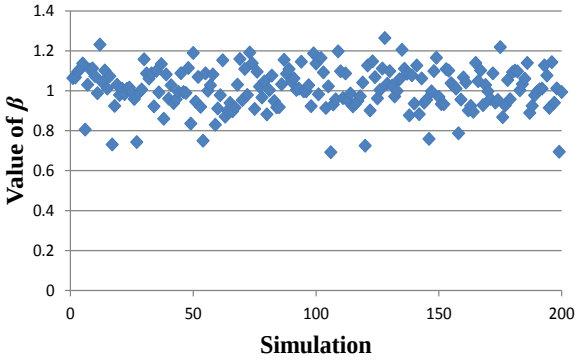

Figure 6: $\beta_{10,2}$ under MA(1)-inflation ($\bar{\beta}_{10,2} = 1.012 \pm 0.103$)

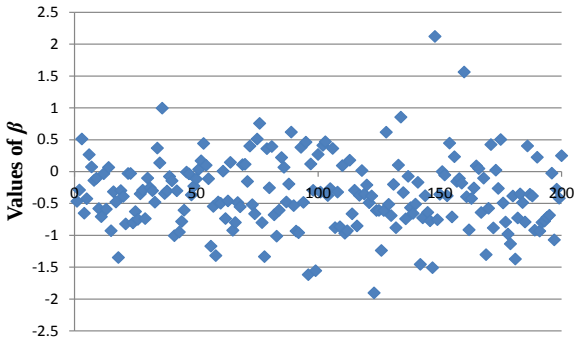

Figure 7: $\beta_{10,1}$ under AR(1)-inflation ($\bar{\beta}_{10,1} = -0.335 \pm 0.550$)

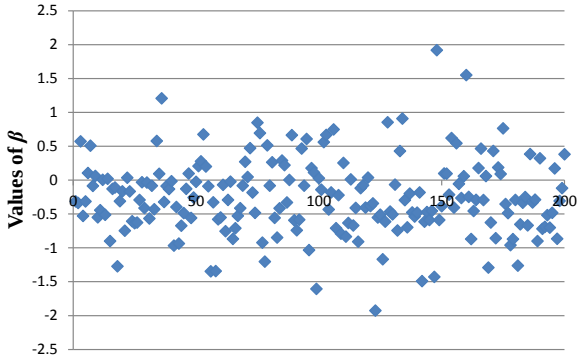

Figure 8: $\beta_{10,2}$ under AR(1)-inflation ($\bar{\beta}_{10,2} = -0.224 \pm 0.555$)

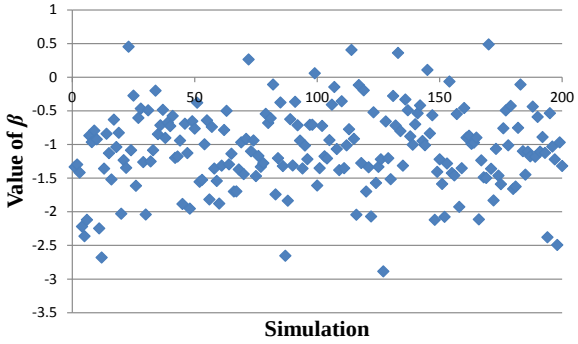

Figure 9: $\beta^*_{10,1}$ under AR(1)-inflation ($\bar{\beta^*}_{10,1} = -1.072 \pm 0.598$)

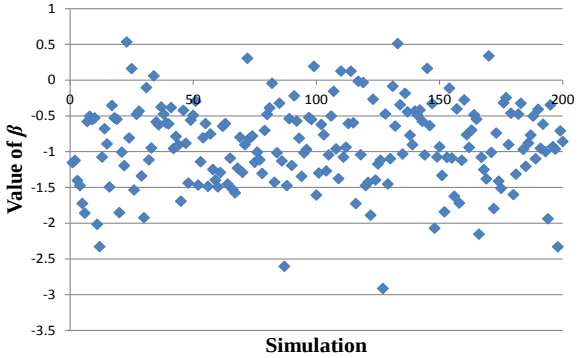

Figure 10: $\beta^*_{10,2}$ under AR(1)-inflation ($\bar{\beta^*}_{10,2} = -0.903 \pm 0.578$)

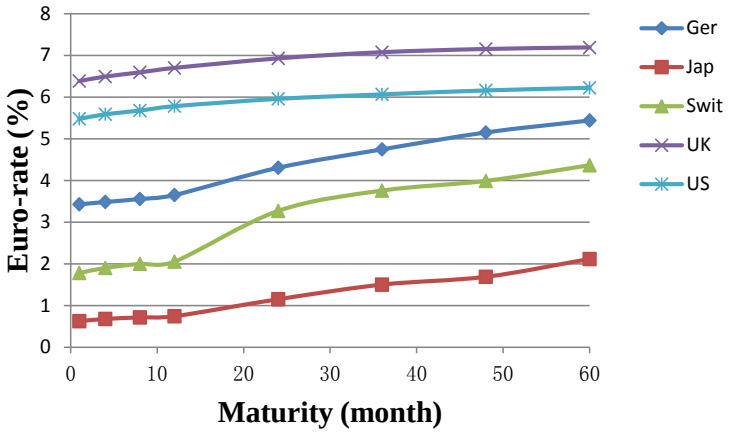

Figure 11: Average Euro Rates

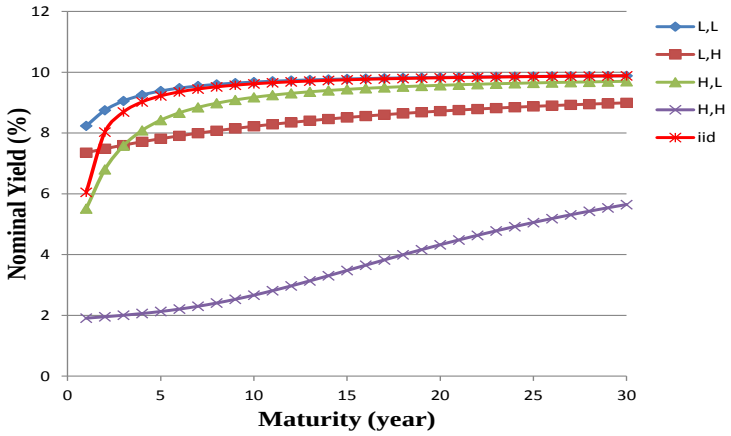

Figure 12: Average Simulated Nominal Yields

L,L: $\sigma = 0.008, \rho = 0.6$; L,H: $\sigma = 0.008, \rho = 0.9$; H,L: $\sigma = 0.0115, \rho = 0.6$;
L,H: $\sigma = 0.0115, \rho = 0.9$; iid: i.i.d-inflation with $\sigma = 0.0115$

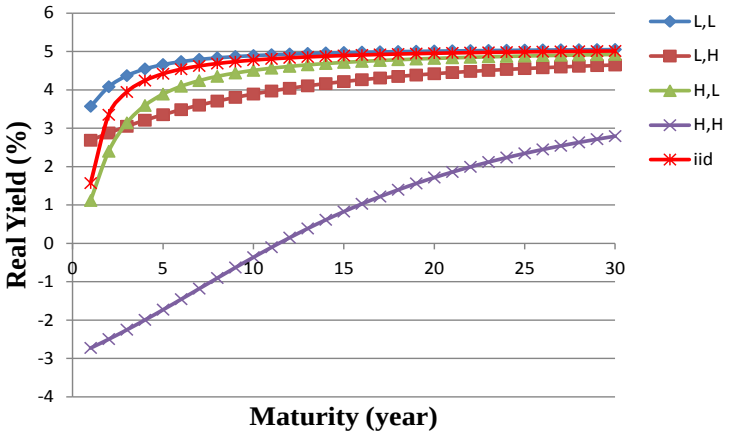

Figure 13: Average Simulated Real Yields

L,L: $\sigma = 0.008, \rho = 0.6$; L,H: $\sigma = 0.008, \rho = 0.9$; H,L: $\sigma = 0.0115, \rho = 0.6$;
L,H: $\sigma = 0.0115, \rho = 0.9$; iid: i.i.d-inflation with $\sigma = 0.0115$