Online Appendix for "Ambiguity, Nominal Bond Yields, and Real Bond Yields"

By GUIHAI ZHAO*

This Appendix provides model solution in Section I, details of expectations hypothesis and predictability of bond returns in Section II, and robustness in Section III.

I. *Model solution*

A. *Forcing process*

Under the worst-case measure, the economic dynamics follow

$$z_{t+1} = \phi_a a_t + \mu_z + x_{z,t} + \sigma^z \varepsilon_{t+1}$$
$$x_{z,t+1} = \rho_x x_{z,t} + \sigma^x \varepsilon_{t+1}$$
$$a_{t+1} = \mu_a + a_t + \sigma^a \varepsilon^a_{t+1}$$

where $z_{t+1} = (\Delta g_{t+1}, \pi_{t+1})^T$, $x_{z,t+1} = (x_{c,t+1},\, x_{\pi,t+1})^T$, $a_{t+1} = (a_{c,t+1},\, a_{\pi,t+1})^T$, $\mu_z = (\mu_c,\, \mu_\pi)^T$, $\mu_a = (\mu^a_c,\, \mu^a_\pi)^T$, $\rho_x = \begin{pmatrix} \rho_c & 0 \\ 0 & \rho_\pi \end{pmatrix}$, $\phi_a = \begin{pmatrix} \phi^a_c & 0 \\ 0 & \phi^a_\pi \end{pmatrix}$, $\sigma^z = \begin{pmatrix} \sigma_c & 0 \\ 0 & \sigma_\pi \end{pmatrix}$, $\sigma^x = \begin{pmatrix} \sigma^x_c & \sigma^x_{c\pi} \\ 0 & \sigma^x_\pi \end{pmatrix}$, $\sigma^a = \begin{pmatrix} \sigma_{ac} & \sigma^{ac}_a \\ 0 & \sigma^{a\pi}_a \end{pmatrix}$, $\varepsilon_{t+1} = (\varepsilon_{c,t+1}, \varepsilon_{\pi,t+1})^T$, and $\varepsilon^a_{t+1} = (\varepsilon_{ac,t+1}, \varepsilon_{a,t+1})^T$. The shocks $\varepsilon_{c,t+1}$, $\varepsilon_{\pi,t+1}$, $\varepsilon_{d,t+1}$, $\varepsilon_{ac,t+1}$, and $\varepsilon_{a,t+1} \sim i.i.d.$ $N(0,1)$. ϕ^a_c and ϕ^a_c represent the equilibrium choice of the upper or lower bound, equal to -1 or $+1$.

B. *Stochastic discount factor*

Given the CRRA utility, the nominal stochastic discount factor in log can be written as

$$m^{\$}_{t,t+1} = log\beta - \gamma \Delta g_{t+1} - \pi_{c,t+1} = log\beta - v' z_{t+1}$$

where $v' = (\gamma,\, 1)$. For the real stochastic discount factor, replace v' with $v' = (\gamma,\, 0)$.

* Zhao: Financial Markets Department, Bank of Canada, 234 Wellington Street, Ottawa, ON K1A 0G9 (email: GZhao@bankofcanada.ca).

C. Bond yields

The time-t price of a zero-coupon bond that pays one unit of consumption n periods from now is denoted $P_t^{(n)}$, and it satisfies the recursion

$$P_t^{(n)} = E_{p_t^o}[M_{t,t+1}^{\$} P_{t+1}^{(n-1)}]$$

with the initial condition that $P_t^{(0)} = 1$ and $E_{p_t^o}$ is the expectation operator for the worst-case measure. Given the linear Gaussian framework, I assume that $p_t^{(n)} = log(P_t^{(n)})$ is a linear function of a_t and x_t:

$$p_t^{(n)} = -A^{(n)} - B^{(n)} x_t - C^{(n)} a_t.$$

When we substitute $p_t^{(n)}$ and $p_{t+1}^{(n-1)}$ in the Euler equation, the solution coefficients in the pricing equation can be solved with $B^{(n)} = B^{(n-1)}\rho_x + v' = v'\left(\sum_{i=o}^{n-1}(\rho_x)^i\right)$, $C^{(n)} = C^{(n-1)} + v'\phi_a = v'\phi_a n$, and

$$A^{(n)} = \begin{array}{c} A^{(n-1)} - log\beta + v'\mu_z + C^{(n-1)}\mu_a - B^{(n-1)}\sigma^x \sigma^{z'} v \\ -0.5 * \left(v'\sigma^z\sigma^{z'}v + B^{(n-1)}\sigma^x\sigma^{x'}B^{(n-1)'} + C^{(n-1)}\sigma^a\sigma^{a'}C^{(n-1)'}\right). \end{array}$$

Nominal bond yields can be calculated as $y_t^{(n)} = -\frac{1}{n}p_t^{(n)} = \frac{A^{(n)}}{n} + \frac{B^{(n)}}{n}x_t + \frac{C^{(n)}}{n}a_t$. The log holding period return from buying an n periods bond at time t and selling it as an $n-1$ periods bond at time $t-1$ is defined as $r_{n,t+1} = p_{t+1}^{(n-1)} - p_t^{(n)}$, and the subjective excess return is $er_{n,t+1} = -Cov_t\left(r_{n,t+1}, m_{t,t+1}^{\$}\right) = -B^{(n-1)}\sigma^x\sigma^{z'}v$. The yield volatility is calculated as

$$\begin{aligned} Var_t\left(y_t^{(n)}\right) &= \left(\frac{B^{(n)}}{n}\sigma^x\right)\left(\frac{B^{(n)}}{n}\sigma^x\right)' \\ &+ \left(\frac{B^{(n)}}{n}\rho_x\sigma^x\right)\left(\frac{B^{(n)}}{n}\rho_x\sigma^x\right)' \\ &+ ... \\ &+ \left(\frac{B^{(n)}}{n}\rho_x^{t-1}\sigma^x\right)\left(\frac{B^{(n)}}{n}\rho_x^{t-1}\sigma^x\right)' \\ &+ t\left(\frac{C^{(n)}}{n}\sigma^a\right)\left(\frac{C^{(n)}}{n}\sigma^a\right)'. \end{aligned}$$

To solve the price and yields for real bonds, we can just replace v' with $v' = (\gamma,\, 0)$.

II. *Expectations hypothesis and predictability of bond returns*

A. *Source of ambiguity - decomposition*

Equation (3) models ambiguity about inflation and real growth in a parsimonious way, which can be further decomposed into two parts. The first part (denoted by $a_{1c,t}$ or $a_{1\pi,t}$) is a random walk with no drift (or a stationary process, for example i.i.d. normal process), which represents agents' ambiguity from transitory shocks in equation (2). The second part (denoted by $a_{2c,t}$ or $a_{2\pi,t}$) is the trend component, which represents agents' ambiguity from the shocks to the expected inflation and growth.

Since the one-step-ahead process contains only the transitory shocks, $a_{1c,t}(= a_{c,t})$ and $a_{1\pi,t}(= a_{\pi,t})$ are the total size of ambiguity for one-quater-ahead inflation and real growth. The specification of $a_{1c,t}$ or $a_{1\pi,t}$ as a random walk with no drift (or a i.i.d. normal process) implies that there is no trend in the realized one-step-ahead ambiguity/dispersion, which is consistent with the data (Figure 1). However, innovations in the two-step-ahead (more than one-step-ahead in general) process consist of both transitory shocks and shocks to expected inflation and growth; hence the total size of ambiguity $a_{c,t+1} = a_{1c,t+1}+a_{2c,t+1}$ and $a_{\pi,t+1} = a_{1\pi,t+1} + a_{2\pi,t+1}$, where $a_{2c,t+1}$ and $a_{2\pi,t+1}$ are the trend components in the total size of ambiguity. For example, one decomposition could be (again, take inflation ambiguity as an example) $a_{1\pi,t+1} = c + \sigma_a \varepsilon_{a,t+1}$ and $a_{2\pi,t+1} = \mu + a_{2\pi,t}$. When investors' perception of ambiguity for long-run inflation is bigger due to long-run uncertainty arising from "inflation scares," shifting endpoints, or stronger disagreement, μ is positive. While in the second subperiod, agents attribute less long-term uncertainty estimated by econometricians to ambiguity, and μ is negative.

At each point in time, due to the trend component in ambiguity from the uncertainty in expected inflation and growth, agents perceive long futures to be more or less ambiguous than short horizons. However, when the time arrives, the second part of the ambiguity containing the trend component does not materialize (expected inflation and growth evolve over time under the true distribution), and the realized one-step-ahead ambiguity does not become bigger or smaller on average. Only when agents evaluate future prospects that are more than one step ahead, does the second part matter.

As shown in Section I, bond prices are solved under the worst-case distribution where the EH roughly holds, and the upward-sloping nominal and real curves are mostly due to the trend component in ambiguity. Whereas the realized yields are calculated using the realized one-step-ahead ambiguity that contains no trends, and this difference makes excess returns on long-term bonds predictable. Note that I focus on the average bond yields in this paper.

B. *Expectations hypothesis: illustration*

The EH states that the yield for an n periods bond is the average of expected future one-period bond yields. Let $y_t^{(n)} = -\frac{1}{n}p_t^{(n)}$ denote the yield for an n periods bond at time t. The intuition of the EH can be illustrated by the following two-periods example where $2y_t^2 = y_t^1 + E_t\left(y_{t+1}^1\right)$. If the yield curve is upward sloping as in the data, $y_t^1 < y_t^2$, it must be that $y_t^1 < y_t^2 < E_t\left(y_{t+1}^1\right)$, that is, the short rate will rise. However, the realized future short rate does not increase enough in the data, and the EH does not seem to work well. The EH is often formally tested through the following equation:

$$y_{t+1}^{n-1} - y_t^n = \alpha + \beta_n\left(\frac{y_t^n - y_t^1}{n-1}\right) + \epsilon_{t+1}.$$

The EH implies that $\beta_n = 1$. However, in the data, many studies (for example, Campbell and Shiller (1991)) show that $\beta_n < 1$, is often negative, and is decreasing with the horizon n.

C. *Expectations hypothesis: model solution*

To derive implications for the test in Section D.2, let $A \equiv y_t^n - y_t^1$ and $B \equiv (n-1)\left(y_{t+1}^{n-1} - y_t^n\right)$. Since all shocks are Gaussian and orthogonal (in the EH test equation in Section D.2), they can be thought of as the error term. The derivation in this session will ignore all shocks. Given the solution for bond yields, we can solve A and B as

$$\begin{aligned}
A &= \frac{A^{(n)}}{n} - A^{(1)} + \left(\frac{B^{(n)}}{n} - B^{(1)}\right)x_t \\
B &= A + VarCov(n-1) + C^{(n-1)}\left(a_{t+1} - a_t - \mu_a\right) \\
VarCov(n-1) &= 0.5 * \left(B^{(n-1)}\sigma^x\sigma^{x'}B^{(n-1)'} + C^{(n-1)}\sigma^a\sigma^{a'}C^{(n-1)'}\right) \\
&\quad + B^{(n-1)}\sigma^x\sigma^{z'}v.
\end{aligned}$$

So the difference between A and B is $VarCov(n-1) + C^{(n-1)}\left(a_{t+1} - a_t - \mu_a\right)$. $VarCov(n-1)$ is quantitatively very small; the difference is mainly driven by $C^{(n-1)}\left(a_{t+1} - a_t - \mu_a\right)$.

When evaluating future prospects, investors' worst-case beliefs are described by $a_{t+1} = \mu_a + a_t + \sigma^a\varepsilon_{t+1}^a$. Ignoring the shock term $\sigma^a\varepsilon_{t+1}^a$, the difference between A and B now only contains the variance and covariance term $VarCov(n-1)$, which is very small. Thus the EH roughly holds.

However, the realized ambiguity process is described by $a_{t+1} = a_t + \sigma^a\varepsilon_{t+1}^a$, and now the difference is $VarCov(n-1) - \mu_a C^{(n-1)}$. To see intuitively what this difference implies for the EH test coefficient β_n, I first ignore the x_t in A and B, and then calculate A and B for different horizons. For $n = 2$:

$$\begin{aligned} A &= \frac{1}{2}\mu_a C^{(1)} - \frac{1}{2}VarCov(1) \\ B &= -A \\ \beta_2 &\approx -1. \end{aligned}$$

For $n = 3$:

$$\begin{aligned} A &= \mu_a C^{(1)} - \frac{1}{3}\left(VarCov(2) + VarCov(1)\right) \\ B &= -A + \frac{1}{3}VarCov(2) - \frac{2}{3}VarCov(1) \\ \beta_2 &\approx -1. \end{aligned}$$

For $n = 4$:

$$\begin{aligned} A &= \frac{3}{2}\mu_a C^{(1)} - \frac{1}{4}\left(VarCov(3) + VarCov(2) + VarCov(1)\right) \\ B &= -A + \frac{1}{2}\left(VarCov(3) - VarCov(2) - VarCov(1)\right) \\ \beta_2 &\approx -1. \end{aligned}$$

Similarly, I can calculate β_n for $n = 5, 6, 7...$ If we ignore the variance/covariance term and x_t, the coefficient $\beta_n = -1$ for all n. To see the exact value for β_n, we should use simulated values for x_t, and take into account $VarCov(n-1)$, and these terms have a bigger impact on β_n when n is small (this intuition can be confirmed by the regression results reported in Table 4, β_n converges to -1 when n is large).

D. *Expectations hypothesis: formal tests*

To formally assess the EH, I show in Section D.3 that the difference between the left-hand side and right-hand side of the EH test equation in Section D.2 is $\left(y_{t+1}^{n-1} - y_t^n\right) - \frac{y_t^n - y_t^1}{n-1} = v'\phi_a\left((a_{t+1} - a_t) - \mu_a\right) + \frac{VarCov_{n-1}}{n-1}$. Taking advantage of the closed-form solution, I show that the coefficient β_n would be -1 for all n if we ignore $x_{z,t}$ and a variance/covariance term. Because of the low autocorrelation (ρ_x), short-term yields are more sensitive to $x_{z,t}$ and the variance/covariance term, yet β_n for long maturities are mainly driven by the difference above and are close to -1.

To further evaluate the predictability of bond returns, I follow the approach in Cochrane and Piazzesi (2005) by first regressing the average of one-year nominal excess bond returns of two to five years to maturity on one- to five-year forward rates, extracting a single bond factor $\hat{rx}_t$ from this regression, and then forecasting excess bond returns at each maturity n from two to five years, $rx_{t\to t+1}^{n,\$} = const +$

$b_n \hat{rx}_t + error$. They show that the estimate b_n is positive and increasing with horizons.

Based on the closed-form solution, the intuition of EH and the source for the predictability of the realized excess bond returns are provided in Section II.E.

III. *Robustness*

This section provides further checks for the sensitivity of the results in several dimensions.

A. *Ambiguity vs. heterogeneous beliefs vs. volatility*

Apart from ambiguity, forecast dispersion is also widely used as measure for disagreement or uncertainty (volatility) in the literature. For example, Ehling et al. (2018) show that stronger inflation disagreement (or bigger inflation forecast dispersion) increases wealth growth expectations for all types of investors, hence unambiguously raising nominal yields. Meanwhile, higher uncertainty (volatility) typically lowers nominal yields through the precautionary savings channel, or through a flight to quality effect as in Bansal and Shaliastovich (2013). In this model, the impact of inflation forecast dispersion (as a measure for ambiguity) on nominal yields is fundamentally different from disagreement or volatility. Instead of an unambiguous effect, I show that higher inflation ambiguity is associated with higher nominal yields before the late 1990s (upper bound as the worst-case belief), and it is associated with lower nominal yields afterwards (lower bound as the worst-case belief).

To distinguish the different effects, I regress 10-year nominal yields on their lagged values and inflation forecast dispersion for the two subperiods, and results are reported in Table E1. Consistent with the impact of inflation ambiguity, the coefficient on dispersion is positive in the first subperiod and negative in the second subperiod. It could be the case that inflation forecast dispersion contains information for both disagreement and ambiguity. However, the result of the second subperiod suggests that ambiguity has a dominating impact.

B. *Regime shift and learning*

In this paper, I assume that there is an unexpected discrete regime shift at the end of 1999 for the following reasons: (1) the term structure of inflation forecast dispersion has switched from upward sloping to downward sloping after the late 1990s (see Figure 1); (2) Figure E1 shows that the correlation between individual GDP growth and inflation forecasts switched from negative in the first subperiod to positive in the second subperiod; and (3) this is consistent with the literature for regime breaks; for example, Campbell, Pflueger, and Viceira (2014) argue that the first subperiod is the inflation fighting period of Volcker and Greenspan and the second subperiod is the recent period of low inflation and increased central bank transparency.

Table 1—Impact of inflation forecast dispersion on nominal yields

	Constant	Yield_10Y_Lag1	Inflation_Disp_Q4	R^2
Yield_10Y	0.38	0.92	0.09	0.96
(1985-1999)	(0.15)	(0.27)	(0.05)	
Yield_10Y	0.18	0.97	-0.07	0.96
(2000-2017)	(0.08)	(0.01)	(0.05)	

Note: This table presents results from the regression of ten-year nominal Treasury bond yields on their lagged values and four-quarter-ahead inflation forecast dispersion. The data employed in the estimation are on a monthly frequency and cover the period from 1985 to 2017. Robust standard errors are reported in the brackets.

While it is useful to clarify the mechanics by assuming an unanticipated regime switch in the late 1990s, there seems no obvious event in this period that this could be tied to. We may ask what the model would imply if we allowed a more gradual transition between these two regimes. Suppose investors know the probabilities of each regime at time t; then stock and bond prices can be computed as the weighted average of the two solutions in Section I. Given that the probability of regime one (negative correlation between growth and inflation expectation) is very high before the late 1990s and close to zero thereafter (see, for example, the estimation in Song 2017), the mechanism of this paper still works and the model results are quantitatively similar. Because the theoretical framework of learning under ambiguity with a regime switch is not clear yet, I will leave this case for future research.

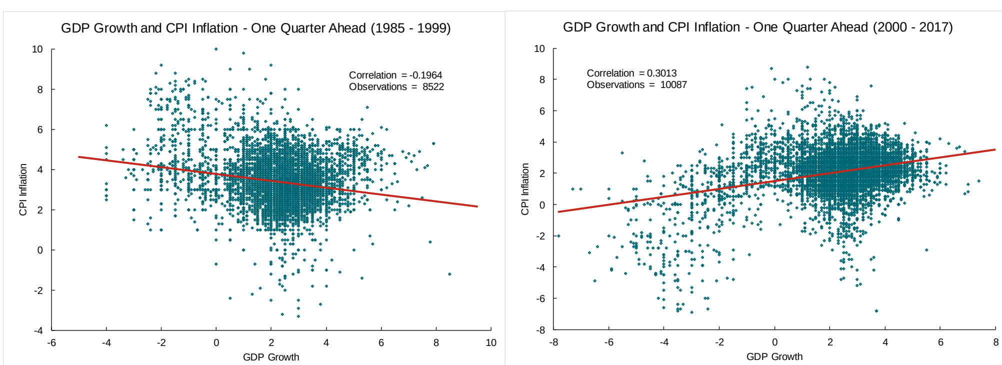

Figure 1. Correlation between individual GDP growth and inflation forecasts

Note: Individual survey data are one-quarter-ahead forecasts from the BCFF and are monthly from 1985 to 2017.

C. Magnitude of ambiguity

Given the specification for the ambiguity process, one natural question is whether the size of the ambiguity is reasonable. I use two different approaches in the literature to provide a sense for the magnitude of the ambiguity.

The first one is the error detection probability suggested by Anderson, Hansen, and Sargent (2003), which quantifies the statistical closeness of two measures by calculating the average error probability in a Bayesian likelihood ratio test of two competing models. Intuitively, measures that are statistically close will be associated with large error probabilities, but measures that are easy to distinguish imply low error probabilities. Formally, let l be the log likelihood function of the worst-case measure relative to the reference measure and P^a be the alternative worst-case measure. Then, the average probability of a model detection error in the corresponding likelihood ratio test is $\epsilon = 0.5 \cdot P(l > 0) + 0.5 \cdot P^a(l < 0)$, where ϵ is just a simple equally weighted average of the probability of rejecting the reference model when it is true ($P(l > 0)$) and the probability of accepting the reference model when the worst-case model is true ($P^a(l < 0)$). In general, a closed-form expression for the detection error probability is not available. The error probability is calculated using simulated data. In this paper, parameters are estimated from data and the error detection probabilities for both output and inflation are at least 5%.

The second approach is by Ilut and Schneider (2014), who argue that (1) the choice of size for ambiguity is guided by the view that agents' ambiguity should be "small enough" relative to the estimated volatility by econometricians, and (2) the extreme forecasts from the implied belief set should perform sufficiently well in forecasting the true data generating process. They show that the maximal size of ambiguity should be smaller than twice the size of the estimated standard error. Yet the size of ambiguity in my model is less than half the size of the estimated standard error, which is small and reasonable according to this standard.

References

Anderson, Evan, Lars Peter Hansen, and Thomas Sargent. 2003. "A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection". *Journal of the European Economic Association* 1 (1): 68–123.

Bansal, Ravi, and Ivan Shaliastovich. 2013. "A Long-run Risks Explanation of Predictability Puzzles in Bond and Currency Markets". *Review of Financial Studies* 26 (1): 1–33.

Campbell, John, Carolin Pflueger, and Luis Viceira. 2014. *Macroeconomic Drivers of Bond and Equity Risks.* NBER Working Paper. National Bureau of Economic Research.

Campbell, John, and Robert Shiller. 1991. "Yield Spreads and Interest Rate Movements: A Bird's Eye View". *The Review of Economic Studies* 58 (3): 495–514.

Cochrane, John, and Monika Piazzesi. 2005. “Bond Risk Premia”. *The American Economic Review* 95 (1): 138–160.

Ehling, Paul, et al. 2018. “Disagreement about Inflation and the Yield Curve”. *Journal of Financial Economics* 127 (3): 459–484.

Ilut, Cosmin, and Martin Schneider. 2014. “Ambiguous Business Cycles”. *The American Economic Review* 104 (8): 2368–2399.

Song, Dongho. 2017. “Bond Market Exposures to Macroeconomic and Monetary Policy Risks”. *The Review of Financial Studies* 30 (8): 2761–2817.