Long-Horizon Forecasts of Asset Prices when the Discount Factor is close to Unity[*]

Charles Engel
University of Wisconsin

Jian Wang
Federal Reserve Bank of Dallas

Jason Wu
Federal Reserve Board

September 2009
Revised September 2010

Abstract

Engel and West (EW, 2005) argue that as the discount factor gets closer to one, present-value asset pricing models place greater weight on future fundamentals. Consequently, current fundamentals have very weak forecasting power and asset prices in these models appear to follow approximately a random walk. We connect the Engel-West explanation to the studies of long-horizon regressions. As expected, we find that under EW's assumption that fundamentals are I(1) and observable to the econometrician, long-horizon regressions generally do not have significant forecasting power when the discount factor is large. However, when EW's assumptions are violated in a particular way, our analytical and simulation results show that long-horizon regressions can have substantial power, even when the discount factor is close to one and the power of short-horizon regressions is low. One example for the substantial power improvement at long horizons is the existence of unobservable stationary fundamentals, such as the risk premium, in present-value asset pricing models. Consistent with our model's prediction, we find that the risk premium calculated from survey data can forecast exchange rates at long horizons. These results suggest that the presence of stationary unobservable fundamentals may have played a large role in the power improvement of long-horizon regressions found in empirical studies.

JEL codes: F31, F41, G12, G15

[*] Charles Engel, Department of Economics, 1180 Observatory Drive, University of Wisconsin, Madison, WI 53706. cengel@ssc.wisc.edu. (608) 262-3697. Jian Wang, Research Department, Federal Reserve Bank of Dallas, 2200 N. Pearl Street, Dallas, TX 75201. jian.wang@dal.frb.org. (214) 922-6471. Jason Wu, Board of Governors of the Federal Reserve System, 20th Street and Constitution Avenue, NW, Washington, D.C. 20551. jason.j.wu@frb.gov. (202) 452-2556. A previous version of this paper was circulated under the title "Can Long-horizon Forecasts Beat the Random Walk under the Engel-West Explanation?" The views in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System.

1 Introduction

A long-standing puzzle in international finance is the disconnect between nominal exchange rates and macroeconomic fundamentals. Engel and West (2005) argue that this disconnect is consistent with exchange rates being determined by economic fundamentals. In connection with the Engel-West explanation, we explore the relationship between fundamentals-based exchange rate models, the disconnect puzzle, and the finding that exchange rate models can forecast exchange rates at long horizons.

Meese and Rogoff's (1983a, 1983b) seminal papers find that a simple random walk model can perform as well as various structural and time series exchange rate models based on out-of-sample forecasting accuracy criteria. While subsequent studies find evidence of exchange rate predictability, most results remain fragile or quantitatively moderate.[1] Engel and West (2005) take a new line of attack on this puzzle. They show that if the exchange rate is determined as a forward-looking asset price, the models predict that the exchange rate approximately follows a random walk when the fundamentals are I(1) and the discount factor is large (close to one). Engel and West (2005) also show that a wide range of exchange rate models can be written in the form of a present-value asset pricing model, and discount factors estimated from monthly or quarterly data are indeed close to one. As a result, exchange rate models should not be judged by whether they can predict exchange rates more accurately than the random walk.[2]

In this paper, we examine whether exchange rates can be better distinguished from a random walk process at longer horizons, under the Engel-West explanation (E-W explanation henceforth). The long-horizon predictability of exchange rates has been an active empirical research topic in the last decade. The underlying rationale is: "While short-horizon changes tend to be dominated by noise, this noise is apparently averaged out over time, thus revealing systematic exchange-rate movements that are determined by economic fundamentals."[3] The early claims of success include Mark(1995), Chinn and Meese (1995), and MacDonald and Taylor (1994). However, these findings have been challenged in later studies. Kilian (1999) finds no long-horizon predictability after updating the dataset used by Mark (1995). He also shows that the long-horizon predictability in Mark (1995) could come from the misspecification of the bootstrap method. Berkowitz and Giorgianni (2001) argue that the long-horizon predictability in Mark (1995) may be driven by the assumption that the exchange rate and fundamentals are cointegrated.

Recently, Engel, Mark, and West (2007) show that out-of-sample forecasting power of models can be increased

---

[1]For instance, see Mark (1995), Groen (2000), Mark and Sul (2001), Kilian and Taylor (2003), Faust, Rogers, and Wright (2003), Molodtsova and Papell (2009), and Wang and Wu (2008). See Cheung, Chinn and Pascual (2005) for a recent comprehensive study.

[2]Rossi (2005) gives another example in which exchange rates approximately follow a random walk even if they are determined by economic fundamentals.

[3]See Mark (1995).

in long-horizon forecasts, especially when using panel data. Wang and Wu (2008) find long-horizon predictability of exchange rates in the context of interval forecasting. Campbell (2001) also finds that long-horizon regressions have power advantages when there is a persistent predictable component in the asset return. More recently, Mark and Sul (2006) study the same model of Campbell (2001) and find that long-horizon regressions have an asymptotic power advantage over the short-horizon regressions if the regressor is endogenous.

The focus of this paper is on whether there is some underlying basis for long-horizon predictability of exchange rates and whether this can be reconciled with the E-W explanation. We study the population R-squared of short- and long-horizon regressions in the context of a present-value asset pricing model as in Engel and West (2005). Two situations are considered. First, we assume that economic fundamentals are nonstationary, following Engel and West (2005). If fundamentals are I(1) and changes of underlying economic variables follow a VAR(p) process, the theoretical R-squared of long-horizon regressions converges to zero for all time horizons, when the discount factor approaches one. That is, for a large discount factor, the long-horizon regressions do not have significant power advantage over a random walk. This is consistent with the E-W explanation and the intuition is straightforward. According to the E-W explanation, the change of the exchange rate in present-value models converges (in probability) to a white noise process as the discount factor approaches one, if economic fundamentals are I(1). In this case, the h-period change of the exchange rate is the sum of h white noises, and therefore a white noise as well.

The more interesting case is when Engel and West's (2005) assumptions that fundamentals are I(1) do not hold, so that some economic fundamentals are I(0) rather than I(1). There are two reasons for us to explore this case. First, Engel and West (2004) find that while observable fundamentals can account for a sizable fraction of exchange rate variation, there is still substantial unexplained variation that may be accounted for by unobservable fundamentals, such as the risk premium in Uncovered Interest Parity (UIP). We calculated this risk premium from Consensus Survey data and it appears to be stationary. Second, economic fundamentals in some exchange rate models, such as bilteral output gap and inflation differentials in the Taylor rule model, are theoretically stationary. In data, evidence of stationarity of these variables are mixed, but it is also well known that unit root tests often suffer from lower power. Hence, we cannot rule out the possibility that these fundamentals are stationary.

In contrast to the case with only I(1) fundamentals, the R-squared of short- and long-horizon regressions no longer converges to zero as the discount factor converges to one. Furthermore, we derive reasonably general analytical conditions under which long-horizon regressions have more power than short-horizon regressions. A special case of these conditions is illustrated in a simple example in Engel, Mark, and West (2007), in which

a stationary, but persistent, unobservable fundamental brings substantial power improvements in long-horizon regressions.

To evaluate whether or not these conditions are satisfied in the data, we consider two standard exchange rate models - the Monetary and Taylor Rule models. Both models include a risk premium, which is a stationary unobservable fundamental. For seven foreign currency-U.S. dollar exchange rates, we calculate the risk premium from Consensus Forecasts of exchange rates, assuming the survey data is an appropriate measure of market expectations. The risk premium and other variables were used to form a dataset of economic fundamentals. We then estimate model coefficients and the variance-covariance matrix from the fundamentals. The data and estimated model coefficients are used to simulate the exchange rate for each country. The population R-squared obtained by simulations at long horizons for several countries is substantially higher than the population R-squared of corresponding short-horizon regressions.

Our results suggest that stationary unobservable fundamentals, such as the risk premium, may play an important role in reconciling the E-W explanation and the empirical evidence of long-horizon predictability. Studies related to this paper include Clarida, Sarno, Taylor, and Valente (2003). They find that the term structure of forward premia may contain information about the risk premium that is useful for forecasting exchange rates. Adrian, Etula, and Shin (2009) find that fluctuations in risk premia captured by the variation in the aggregate balance sheets of financial intermediaries are useful in forecasting exchange rates.[4]

While this paper focuses on the exchange rate, our results are derived for a general present-value asset pricing model, hence they can be applied to other asset prices as well. Campbell (2001) studies long-horizon predictability of stock returns. There is indeed a close analogy of his model to the case with stationary fundamentals of this paper and this connection is detailed in the appendix.[5]

The remainder of the paper is organized as follows: Section 2 provides a brief introduction to the Engel-West explanation. In Section 3, we derive the long-horizon regressions from present-value asset pricing models and study the power of long-horizon regressions when the discount factor is close to one. Two cases are discussed in this section: (1) all observable fundamentals are I(1), and (2) some observable and unobservable fundamentals are I(0). Section 4 presents our simulation results. Section 5 summarizes major findings and concludes.

---

[4]See Chen and Tsang (2009) for a recent study along this line of research.
[5]See Appendix A.1 for details.

2 Asset Pricing Model and E-W Explanation

Engel and West (2005) show that a wide range of exchange rate models can be written as a special case of linear present-value asset pricing models, which have a general form of:

$$s_t = (1-b)\sum_{j=0}^{\infty} b^j E_t \left(f_{1,t+j} + u_{1,t+j}\right) + b\sum_{j=0}^{\infty} b^j E_t \left(f_{2,t+j} + u_{2,t+j}\right), \tag{1}$$

where s_t is the asset price, $0 < b < 1$ is the discount factor, $\{f_{1,t}, f_{2,t}\}$ are observable fundamentals and $\{u_{1,t}, u_{2,t}\}$ are unobservable fundamentals. The fundamentals are functions of a vector of economic variables, $\{\mathbf{X}_t\}$. For instance, in the Monetary model in Mark (1995), s_t is the logarithm of the exchange rate and $f_{1,t} = (m_t - m_t^*) - (y_t - y_t^*)$, where $\{m_t, y_t\}$ and $\{m_t^*, y_t^*\}$ are money and output for home and foreign countries, respectively. To avoid confusion, we call the elements in $\mathbf{X}_t$, for instance m_t and y_t, economic variables and $f_{i,t}$ and $u_{i,t}$ fundamentals.

Some well-known applications of the present-value model include Campbell and Shiller (1987, 1988) and West (1988). In equation (1), the asset price equals the sum of the expected present value of future fundamentals. Engel and West (2005) demonstrate analytically that if the discount factor b is close to unity, and either (1) $f_{1,t} + u_{1,t} \sim I(1)$ and $f_{2,t} + u_{2,t} = 0$, or (2) $f_{2,t} + u_{2,t} \sim I(1)$ with $f_{1,t} + u_{1,t}$ unrestricted, then the exchange rate s_t approximately follows a random walk.

To develop intuition, consider a simple example where $u_{1,t} = f_{2,t} = u_{2,t} = 0$. Equation (1) reduces to:

$$s_t = (1-b)\sum_{j=0}^{\infty} b^j E_t(f_{1,t+j}). \tag{2}$$

If the fundamental $f_{1,t}$ follows a random walk, it is straightforward from the above equation that $s_t = f_{1,t}$, which means that the exchange rate is also a random walk. However, first differences of the fundamentals are typically serially correlated in the data. For simplicity, we assume that the first difference of $f_{1,t}$ follows an AR(1) process:

$$\Delta f_{1,t+1} = \phi \Delta f_{1,t} + \varepsilon_{t+1}, \tag{3}$$

where $|\phi| < 1$. From equations (2) and (3), we obtain:

$$\Delta s_{t+1} = \frac{\phi(1-b)}{1-b\phi} \Delta f_{1,t} + \frac{1}{1-b\phi}\varepsilon_{t+1}. \tag{4}$$

In equation (4), the first difference of the exchange rate is a weighted average of the first difference of the

fundamental at time t and "news" arriving at time $t+1$. Δs_{t+1} is serially correlated because $\Delta f_{1,t}$ is serially correlated. However, when b approaches one, the coefficient in front of $\Delta f_{1,t}$ converges to zero. As a result, Δs_{t+1} converges to $\frac{1}{1-\phi}\varepsilon_{t+1}$, which is simply white noise. The intuition for this result is that more weight is given to the new information rather than the changes of current fundamentals when the discount factor is large. Engel and West (2005) show that the discount factor estimated from data is indeed very close to one. They argue that the exchange rate may be determined by fundamentals, but cannot be distinguished from a random walk.

3   Long-Horizon Regressions under E-W Explanation

In this section, we drive long-horizon regressions from the linear present-value models studied in Engel and West (2005). We investigate conditions under which the E-W explanation is consistent with the empirical findings that economic models can forecast exchange rates at long horizons.

Long-horizon regressions have been widely used to test the efficiency of asset markets.[6] A strand of literature has attempted to test the long-horizon predictability of exchange rates, including Mark (1995), Chinn and Meese (1995), Kilian (1999) and Berkowitz and Giorgianni (2001). Let $s_{t+h} - s_t$ be the h-period change of the exchange rate and z_t be the deviation of the exchange rate from its long-run equilibrium level. Typically, in-sample or out-of-sample forecasts of $s_{t+h} - s_t$ from the model:

$$s_{t+h} - s_t = c_h + \beta_h z_t + \upsilon_{t+h} \tag{5}$$

are compared with in-sample or out-of-sample forecasts from the random walk model (equation (5) with the restriction $\beta_h = 0$).[7] If the R-squared of equation (5) increases, or the out-of-sample relative Mean Square Prediction Error (MSPE) decreases with time horizon h, it is taken as evidence that the exchange rate converges to its long-run value over time and therefore is predictable at long horizons.

We derive long-horizon regressions from the following linear present-value model:

$$s_t = (1-b) \sum_{j=0}^{\infty} b^j E_t f_{1,t+j} + b \sum_{j=0}^{\infty} b^j E_t (f_{2,t+j} + u_{2,t+j}). \tag{6}$$

Notice that compared to equation (1), u_{1t} was omitted in the model to simplify the set up. Two notable exchange rate models in the literature fall under the setup in equation (6): the Monetary model and the Taylor

---

[6]See, for instance, Fama and French (1988) and Campbell and Shiller (1988).
[7]One can also restrict $c_h = 0$ to obtain the driftless random walk model.

Rule model.[8] In the Monetary model,

$$s_t = (1-b)\sum_{j=0}^{\infty} b^j E_t \left(m_{t+j} - m_{t+j}^* - \gamma(y_{t+j} - y_{t+j}^*) + q_{t+1} - (v_{t+j} - v_{t+j}^*) \right) - b \sum_{j=1}^{\infty} b^j E_t \rho_{t+j}, \qquad (7)$$

where m_t and y_t are logarithms of the domestic money stock and output, respectively. The superscript $*$ denotes the foreign country. $q_t \equiv s_t + p_t^* - p_t$ is the real exchange rate, v_t is a money demand shock, and $\rho_t \equiv E_t s_{t+1} - s_t - (i_t - i_t^*)$ is the deviation from uncovered interest rate parity. We interpret this deviation as a risk premium in currency markets. In this model $f_{1,t} = m_t - m_t^* - \gamma(y_t - y_t^*) + q_t - (v_t - v_t^*)$, $f_{2,t} = 0$, and $u_{2,t} = -\rho_t$.

In the Taylor Rule model,

$$s_t = (1-b)\sum_{j=0}^{\infty} b^j E_t(p_{t+j} - p_{t+j}^*) - b \sum_{j=0}^{\infty} b^j E_t \left(r_g(y_{t+j}^g - y_{t+j}^{*g}) + r_\pi(\pi_{t+j} - \pi_{t+j}^*) + v_{t+j} - v_{t+j}^* + \rho_t \right), \qquad (8)$$

where p_t, y_t^g and π_t are domestic aggregate price, output gap and the inflation rate, respectively, and v_t is a monetary policy shock. In this case $f_{1,t} = p_t - p_t^*$, and $f_{2,t} + u_{2,t} = -\left(r_g(y_t^g - y_t^{*g}) + r_\pi(\pi_t - \pi_t^*) + v_t - v_t^* + \rho_t \right)$.

As shown in the above examples, the fundamentals in equation (6) are linear functions of economic variables. Collect the economic variables in an $n \times 1$ vector $(\mathbf{X}_t)$ and assume that the first difference of $\mathbf{X}_t$ follows a stationary VAR(p) process:

$$\varepsilon_t = \Delta \mathbf{X}_t - \Phi_1 \Delta \mathbf{X}_{t-1} - \ldots - \Phi_p \Delta \mathbf{X}_{t-p} \equiv (I_n - \Phi(L))\Delta \mathbf{X}_t$$

$$E_t \varepsilon_{t+j} = 0, \forall j > 0 \qquad (9)$$

$$E(\varepsilon_t \varepsilon_t') = \Omega.$$

Although $\mathbf{X}_t$ seemingly contains only I(1) economic variables, our setup also includes cases with stationary economic variables. In these cases, $\mathbf{X}_t$ contains the levels of I(1) variables and the summation of I(0) variables from negative infinity to time t.

Let $\alpha_1, \alpha_{21}, \alpha_{22}$ be $n \times 1$ coefficient vectors. Throughout the paper, we maintain the assumption that:

$$f_{1,t} = \alpha_1' \mathbf{X}_t \sim I(1) \qquad (10)$$

$$u_{2,t} = \alpha_{22}' \Delta \mathbf{X}_t \sim I(0).$$

---

[8]See Appendix A.2 for details of these two models.

The assumption that $f_{1,t} \sim I(1)$ is reasonable and supported by empirical evidence of exchange rate models studied in the literature. The assumption that $u_{2,t} \sim I(0)$ is maintained to guarantee that long-horizon regressions are not spurious.[9] $u_{2,t}$ is a measure of risk premium in the Monetary and Taylor Rule models. Our survey data suggests that $u_{2,t}$ is stationary. More details will be given later. For $f_{2,t}$, two cases are considered:

Case 1.

$$f_{2,t} = u_{2,t} = 0 \text{ or } f_{2,t} = \alpha'_{21}\mathbf{X}_t \sim I(1) \text{ with } \alpha_{21} \neq 0 \text{ and } \alpha_{22} \text{ unrestricted;} \tag{11}$$

Case 2.

$$f_{2,t} = \alpha'_{21}\Delta\mathbf{X}_t \sim I(0) \text{ with } f_{2,t} + u_{2,t} \neq 0. \tag{12}$$

In Case 1, all observable fundamentals (either $f_{1,t}$ or $\{f_{1,t}, f_{2,t}\}$) are I(1). Also, $f_{2,t} + u_{2,t}$ is either I(1) or zero, so Engel and West's (2005) assumptions are satisfied. Case 2, however, implies that $f_{2,t} + u_{2,t} \sim I(0) \neq 0$, which violates Engel and West's assumptions. This case is motivated by exchange rate models studied in the literature and has high empirical relevance. In the Monetary model of equation (7), $f_{2,t}$ equals zero, but the risk premium ($u_{2,t}$) is non-zero. In the Taylor Rule model of equation (8), in addition to the risk premium $u_{2,t}$, $f_{2,t}$ is also stationary in theory.

We derive long-horizon regressions and discuss the properties of the population R-squared of these regressions, under both Cases 1 and 2. In both cases, the regressor or equilibrium error is defined as $z_t = s_t - f_{1,t} - \frac{b}{1-b}f_{2,t}$. This definition of z_t is consistent with most empirical studies on long-horizon regressions, including Mark (1995), Chinn and Meese (1995), Engel, Mark and West (2007), and Molodtsova and Papell (2009). Furthermore, long-horizon regressions are often based on an OLS regression with z_t as the only regressor. Therefore, the population R-squared of interest is $\{Cov\,(s_{t+h} - s_t, z_t)\}^2 / Var(z_t)Var(s_{t+h} - s_t)$. To summarize, for both Cases 1 and 2,

$$
\begin{aligned}
z_t &\equiv s_t - f_{1,t} - \frac{b}{1-b}f_{2,t} \\
R^2(h) &\equiv \frac{\{Cov\,(s_{t+h} - s_t, z_t)\}^2}{Var(z_t)Var(s_{t+h} - s_t)}.
\end{aligned}
\tag{13}
$$

[9]By definition $u_{2,t}$ is unobservable, hence not included as a regressor in long-horizon regressions. If this "left-out" variable is non-stationary, long-horizon regressions will be spurious.

Finally, we use following definitions:

$$
\underset{np\times np}{F} \equiv
\begin{bmatrix}
\Phi_1 & \Phi_2 & \dots & \Phi_{p-1} & \Phi_p \\
I_n & 0 & \dots & 0 & 0 \\
0 & I_n & \dots & 0 & 0 \\
\vdots & \vdots & \dots & \vdots & \vdots \\
0 & 0 & \dots & I_n & 0
\end{bmatrix}, \quad
\underset{np\times 1}{\Delta\mathbf{Y}_t} \equiv
\begin{bmatrix}
\Delta\mathbf{X}_t \\
\Delta\mathbf{X}_{t-1} \\
\vdots \\
\Delta\mathbf{X}_{t-p+1}
\end{bmatrix}, \quad
\underset{np\times 1}{\mathbf{v}_t} \equiv
\begin{bmatrix}
\varepsilon_t \\
0 \\
\vdots \\
0
\end{bmatrix}, \quad
\underset{np\times n}{\iota} \equiv
\begin{bmatrix}
I_n \\
0 \\
\vdots \\
0
\end{bmatrix}.
$$

Notice that $\Delta\mathbf{Y}_t = F\Delta\mathbf{Y}_{t-1}+\mathbf{v}_t$. Define $F(b) \equiv bF(I_{np}-bF)^{-1}$ for $b \in (0,1]$. $F(b)$ exists because all eigenvalues of F lie inside the unit circle by the assumption that $\Delta\mathbf{X}_t$ is a stationary process. Also, let $\mathbf{Q} \equiv E(\mathbf{v}_t\mathbf{v}'_t)$.

3.1   Long-Horizon Regression under Case 1

Proposition 1. *Consider the case specified in (11) and fix $b < 1$. Under the setup in (6), (9), (10) and (13),*

(a) For ρ_h and υ_{t+h} defined in the appendix,

$$
s_{t+h} - s_t = \rho_h z_t - \alpha'_{22}\iota'\left(\sum_{k=0}^{h-1} F^k\right)\Delta\mathbf{Y}_t + \upsilon_{t+h} \tag{14}
$$

is a valid regression where $(z_t, \Delta\mathbf{Y}_t)$ and υ_{t+h} do not correlate. Also,

$$
z_t = \beta'(b)\Delta\mathbf{Y}_t
$$

where $\beta(b) \equiv b\iota\alpha_{22} + F'(b)\iota\alpha(b)$ and $\alpha(b) \equiv \alpha_1 + \frac{b}{1-b}\alpha_{21} + b\alpha_{22}$. Hence, z_t is stationary.

(b) For $np \times 1$ vectors $A_{h,b}$ and $\{B_{k,b}\}_{k=1}^{h-1}$ defined in the appendix,

$$
\frac{1 - R^2(h)}{R^2(h)} = \frac{\sum_{k=0}^{h-1} B'_{k,b}\mathbf{Q}B_{k,b}}{A'_{h,b}E(\Delta\mathbf{Y}_t\Delta\mathbf{Y}'_t)A_{h,b}},
$$

As $b \to 1$, $A_{h,b} = O(1)$ while $B_{h,b} = O(b/(1-b))$ for all $h < \infty$.

Proof. See appendix. $\square$

Proposition 1(a) derives long-horizon regressions based on the present-value model and VAR(p) process of economic variables. Note that in the absence of unobservable economic variables ($\alpha_{22} = 0$), the long-horizon regression in equation (14) reduces to the one that is typically used in empirical studies, for instance, equation

(5). That is, Proposition 1(a) shows that the use of long-horizon regressions in exchange rate forecasting can be rationalized by our present-value model in which all economic variables are observable to the econometrician. We will show shortly that this result also holds in Case 2. Our result generalizes Campbell and Shiller (1987) and Nason and Rogers (2008), which derive the short-horizon regression (i.e. $h = 1$) from a present-value model with only observable fundamentals.

Proposition 1(b) implies that $R^2(h) \to 0$ for all time horizons h as $b \to 1$. That is, the predictive power of long-horizon regressions, as measured by the population R-squared, becomes negligible for all time horizons if the discount factor is close to one. As a result, economic models have no significant power advantage over the random walk in forecasting exchange rates even at long horizons. This extends the Engel-West Theorem to the cases where $h > 1$.

Example 1. To illustrate the results in Proposition 1, consider a simple example in which $f_{2,t} = u_{2,t} = 0$, $\alpha_1 = 1$, and

$$n = 1 \quad , \quad p = 1$$

$$\mathbf{X}_t = X_{1,t} \quad , \quad \varepsilon_t = \varepsilon_{1,t}$$

$$\Phi_1 = \phi_1 \quad , \quad \Omega = \sigma_1^2.$$

This example falls under Case 1, with $f_{1,t} = X_{1,t}, z_t = s_t - f_{1,t}$. Using Proposition 1:

$$\Delta s_{t+1} = \frac{1-b}{b} z_t + \frac{1}{1 - b\phi_1} \varepsilon_{1,t+1}. \tag{15}$$

Comparing equations (15) and (14), $\rho_1 = \frac{1-b}{b}$ and $v_{t+1} = \frac{1}{1-b\phi_1}\varepsilon_{1,t+1}$. After some algebra, we have:

$$R^2(1) = \frac{\rho_1^2 Var(z_t)}{\rho_1^2 Var(z_t) + Var(v_{t+1})}, \text{ where}$$

$$Var(z_t) = \left(\frac{b\phi_1}{1 - b\phi_1}\right)^2 \frac{\sigma_1^2}{1 - \phi_1^2}, \text{ and } Var(v_{t+1}) = \frac{\sigma_1^2}{(1 - b\phi_1^2)}. \tag{16}$$

When the discount factor $b \to 1$, $\rho_1 \to 0$ while $Var(z_t)$ and $Var(v_{t+1})$ converge to constants. Hence, $R^2(1) \to 0$. For instance, if $b = 0.95$, ρ_1 is approximately equal to 0.05. Assuming $\phi_1 = 0.5, \sigma_1^2 = 1$ and using equation (16), we have $R^2(1) \approx 0.001$. ■

Example 2. Consider the long-horizon regressions for the setup in Example 1. Using Proposition 1 and the

fact that $\rho_1 = \frac{1-b}{b}$, we obtain:

$$\rho_h = \rho_1 \frac{1 - \phi_1^h}{1 - \phi_1} \equiv \rho_1 \mu_h. \tag{17}$$

This special result was also derived in Berkowitz and Giorgianni (2001). The multiplier $\mu_h \equiv \frac{1-\phi_1^h}{1-\phi_1} > 1$ increases monotonically in h and ϕ_1 for $\phi_1 \in (0, 1)$. As a result, ρ_h is an increasing function of the time horizon h. However, for a large discount factor, say $b = 0.95$, the level and increases in ρ_h are small unless ϕ_1 is unrealistically large. Even if $\phi_1 = 0.8$, a level difficult to justify given the data, $\lim_{h\to\infty} \mu_h = 5$ and $\lim_{h\to\infty} \rho_h = 0.26$.

Since $u_{2t} = 0$, Proposition 1 implies that:

$$R^2(h) = \frac{\rho_h^2 Var(z_t)}{\rho_h^2 Var(z_t) + Var(v_{t+h})}. \tag{18}$$

Using Proposition 1, it is straightforward, albeit tedious, to show that:

$$v_{t+h} = \frac{\phi(1-b)}{1-b\phi} \left\{ \varepsilon_{t+1} \left(\frac{1}{\phi(1-b)} + \frac{1 - \phi^{h-1}}{1 - \phi} \right) + \varepsilon_{t+2} \left(\frac{1}{\phi(1-b)} + \frac{1 - \phi^{h-2}}{1 - \phi} \right) \right.$$
$$\left. + ... + \varepsilon_{t+h-1} \left(\frac{1}{\phi(1-b)} + 1 \right) + \frac{1}{\phi(1-b)} \varepsilon_{t+h} \right\}. \tag{19}$$

Using equation (19), $Var(v_{t+h})$ can be written as:

$$Var(v_{t+h}) = \frac{\phi_1^2(1-b)^2\sigma_1^2}{(1-b\phi_1)^2} \left\{ \sum_{j=1}^{h-1} \left(\frac{1}{\phi_1(1-b)} + \frac{1 - \phi_1^j}{1 - \phi_1} \right)^2 + \frac{1}{\phi_1^2(1-b)^2} \right\}. \tag{20}$$

Substituting the expression for $Var(z_t)$ in equation (16) and $Var(v_{t+h})$ in equation (20) into equation (18), we can calculate the population R-squared $R^2(h)$.

Proposition 1 and Example 1 show that $R^2(1) \approx 0$ for a reasonable ϕ_1 and $b \approx 1$. To compare $R^2(h)$ more meaningfully across h, we calculate the ratio of $R^2(h)/R^2(1)$ where $h = 1, ..., 16$. Figure 1 plots the R-squared ratios under different AR(1) coefficients (ϕ_1). When $\Delta f_{1,t}$ is not very persistent ($\phi_1 = 0.4$ for instance), the ratio is less than one for all horizons h and decreases monotonically with h. When $\Delta f_{1,t}$ is more persistent, there is a hump-shaped relation between the R-squared and time horizon: for some small h's, long-horizon regressions have greater power than the short-horizon (i.e. $h = 1$) regressions, but the power improvement for long-horizon regressions is limited. The maximum R-squared in the long-horizon regressions is less than twice that in the short-horizon regression. The power advantage of long-horizon regressions also dies out eventually with the increase of h. Long-horizon regressions generally have less power than short-horizon regressions when $h > 10$.

In order to understand the hump-shaped pattern, Figure 2 plots ρ_h and $Var(v_{t+h})$ across h when $\phi_1 = 0.8$.

As predicted by equation (17), ρ_h increases with the time horizon h, but at a decreasing rate. $Var(\upsilon_{t+h})$, on the other hand, increases at an approximately constant rate. ρ_h increases at a faster rate than the variance of υ_{t+h} at the beginning (for $h \leq 4$ in our example). At these horizons, the numerator of $R^2(h)$ increases faster than the denominator and hence $R^2(h)$ increases in h. However, the growth rate of $var(\upsilon_{t+h})$ eventually catches up and exceeds that of ρ_h as h increases. As a result, $R^2(h)$ starts to decline with h. ■

In summary, in the absence of unobservable fundamentals ($u_{2t} = 0$), long-horizon regressions used in empirical studies are valid regressions in which regressors and error terms are uncorrelated. In the case where observable fundamentals are I(1) (Case 1), the population R-squared of long-horizon regressions ($R^2(h)$) converges to zero as the discount factor b converges to 1, for any horizon h. Hence, long-horizon regressions have negligible power improvements over the random walk for a large discount factor b, consistent with the Engel and West (2005) prediction. An example shows that if the changes of economic fundamentals are persistent, the power of long-horizon regressions, although small, displays a hump shape when compared to the power of the short-horizon regression.

3.2 Long-Horizon Regressions under Case 2

Proposition 2. *Consider the case specified in (12) and a fixed $b < 1$. Under the setup in (6), (9), (10) and (13),*

(a) For a fixed $b < 1$ and θ_h and ν_{t+h} defined in the appendix,

$$s_{t+h} - s_t = \theta_h z_t - \alpha_{22}^{'} \iota^{'} \left(\sum_{k=0}^{h-1} F^k \right) \Delta \mathbf{Y}_t + \nu_{t+h}$$

is a valid regression where $(z_t, \Delta \mathbf{Y}_t)$ and ν_{t+h} do not correlate. Also,

$$z_t = \omega^{'}(b) \Delta Y_t$$

where $\omega(b) \equiv b\iota(\alpha_{22} - \frac{b}{1-b}\alpha_{21}) + F^{'}(b)\iota\eta(b)$ and $\eta(b) \equiv \alpha_1 + b(\alpha_{21} + \alpha_{22})$. Hence, z_t is stationary.

(b) Set $b = 1$. For $np \times 1$ vectors C_h and $\{D_k\}_{k=0}^{h-1}$ defined in the appendix,

$$\frac{1 - R^2(h)}{R^2(h)} = \frac{\sum_{k=0}^{h-1} D_k^{'} \mathbf{Q} D_k}{C_h^{'} E(\Delta \mathbf{Y}_t \Delta \mathbf{Y}_t^{'}) C_h}.$$

Further, if we define:

$$\mathbf{P}_{t+k} \equiv E_t(f_{2,t+k} + u_{2,t+k})$$

$$\Xi_{1,t+k} \equiv f_{2,t+k} + u_{2,t+k} - E_t(f_{2,t+k} + u_{2,t+k})$$

$$\Xi_{2,t+k} \equiv E_{t+k}\left(\sum_{j=0}^{\infty}\Delta f_{1,t+k+j} + f_{2,t+k+j} + u_{2,t+k+j}\right) - E_{t+k-1}\left(\sum_{j=0}^{\infty}\Delta f_{1,t+k+j} + f_{2,t+k+j} + u_{2,t+k+j}\right),$$

then,

$$\frac{1 - R^2(h)}{R^2(h)} = \frac{Var\left(\sum_{k=0}^{h-1}(\Xi_{1,t+k} - \Xi_{2,t+k+1})\right)}{Var(\sum_{k=0}^{h-1}\mathbf{P}_{t+k})}.$$

Proof. See appendix. □

Propositions 2(a) and 2(b) are analogous to Propositions 1(a) and 1(b), except that $R^2(h)$ no longer converges to zero when $b \to 1$. This highlights the fact that the existence of stationary fundamentals (f_{2t} and u_{2t}) opens up the possibility of detecting exchange rate predictability in long-horizon regressions even when the discount factor b is close to one. This is true even when the existence of u_{2t} leads to a misspecification in long-horizon regressions, where only z_t is used as a regressor.

Notice that $\mathbf{P}_{t+k}$ is the time t forecast of $f_{2,t+k} + u_{2,t+k}$, while $\Xi_{1,t+k}$ and $\Xi_{2,t+k}$ are forecast errors, where $\Xi_{1,t+k}$ is the forecast error induced by predicting $f_{2,t+k}+u_{2,t+k}$ at time t, and $\Xi_{2,t+k}$ is the forecast error induced by predicting $\sum_{j=0}^{\infty}\Delta f_{1,t+k+j} + f_{2,t+k+j} + u_{2,t+k+j}$ at time $t + k - 1$ instead of $t + k$. Therefore, Proposition 2(b) says that in order for $R^2(h)$ to increase with h, the following condition needs to be satisfied:

$$\frac{Var\left(\sum_{k=0}^{h}(\Xi_{1,t+k} - \Xi_{2,t+k+1})\right)}{Var\left(\sum_{k=0}^{h-1}(\Xi_{1,t+k} - \Xi_{2,t+k+1})\right)} - 1 \leq \frac{Var(\sum_{k=0}^{h}\mathbf{P}_{t+k})}{Var(\sum_{k=0}^{h-1}\mathbf{P}_{t+k})} - 1. \tag{21}$$

Therefore, equation (21) says that R-squared increases with the forecasting horizon if the percentage increase in the variance of the forecast errors is smaller than the percentage increase in the variance of the forecasts. Intuitively, long-horizon forecasts of exchange rates using fundamentals are useful relative to the random walk when $f_{2t} + u_{2t}$ is sufficiently persistent, while on the other hand errors induced in the forecasting process are reasonably small. A corollary of Proposition 2(b) is that as h increases, $R^2(h)$ converges to zero as the forecast errors accumulate and overwhelm the predictable component.

Although equation (21) gives the condition for long-horizon regressions to have more power than the short-

horizon regressions, the extent of power improvement depends on the structure of the fundamentals. Engel, Mark, and West (2007) gave a simple example in which long-horizon regressions have substantial power improvement. The example can be considered as a special case of Proposition 2. We use a similar example here to develop intuition.

Example 3. Consider the setup in equation (6). Assuming $f_{2,t} = 0$, we have:

$$s_t = (1-b) \sum_{j=0}^{\infty} b^j E_t f_{1,t+j} + b \sum_{j=0}^{\infty} b^j E_t u_{2,t+j}. \tag{22}$$

We further assume $f_{1,t}$ follows a random walk[10] and $u_{2,t}$ follows an AR(1) process:

$$f_{1,t+1} = f_{1,t} + \varepsilon_{t+1}$$

$$u_{2,t+1} = \phi u_{2,t} + v_{t+1},$$

where $|\phi| < 1$. After some algebra, we obtain:

$$s_t = f_{1,t} + \frac{b u_{2,t}}{1 - b\phi}. \tag{23}$$

In equation (23), the exchange rate is determined by a permanent component $f_{1,t}$, and a transitory component $\frac{b u_{2,t}}{1-b\phi}$. Similarly, the h-period change of the exchange rate is also determined by the change in the permanent component and the change in the transitory component:

$$s_{t+h} - s_t = f_{1,t+h} - f_{1,t} + \frac{b}{1 - b\phi} (u_{2,t+h} - u_{2,t})$$

$$= P_{t+h} + T_{t+h}, \tag{24}$$

where $P_{t+h} \equiv f_{1,t+h} - f_{1,t}$ is exchange rate movements due to the change of the permanent component, and $T_{t+h} \equiv \frac{b}{1-b\phi} (u_{2,t+h} - u_{2,t})$ is exchange rate movements due to the change of the transitory component. We can

---

[10]Allowing $f_{1,t}$ to have a stationary component does not change our results. The stationary, or transitory, component has negligible effects on long-horizon predictability if $f_{1,t}$ is I(1) and the discount factor is large.

further elaborate these expressions into:

$$P_{t+h} = \sum_{j=1}^{h} \varepsilon_{t+j}$$

$$T_{t+h} = \frac{b}{1-b\phi} \left[(\phi^h - 1)u_{2,t} + \sum_{j=0}^{h-1} \phi^j v_{t+h-j} \right]. \tag{25}$$

It is helpful to recognize that the right-hand-side of long-horizon regressions is proportional to $u_{2,t}$:

$$z_t = s_t - f_{1t} = \frac{bu_{2,t}}{1-b\phi}. \tag{26}$$

As a result, the term containing $u_{2,t}$ in T_{t+h} can be predicted by z_t. If one regresses T_{t+h} on $u_{2,t}$ (and similarly on z_t, since z_t is proportional to $u_{2,t}$), the population R-squared equals $\frac{1-\phi^h}{2}$, which increases with time horizon h with an upper bound of 0.5. When $u_{2,t}$ is very persistent (ϕ is close to one), T_{t+1} is difficult to predict at the short horizon because the R-squared ($\frac{1-\phi}{2}$) of the short-horizon regression is small when ϕ is close to one. However, it converges to 0.5 as $h \to \infty$. In this case, the transitory part T_{t+h} is helpful for finding long-horizon predictability. ∎

In short, Proposition 2 and Example 3 show that when $f_{2t} + u_{2t}$ is stationary, but persistent, the E-W explanation is consistent with the finding that exchange rate models can forecast exchange rates at long- but not short-horizons.

4   Simulations

Proposition 2 opens up the possibility of finding a non-trivial R-squared in long-horizon regressions even when the discount factor b is close to one. The power of long-horizon regressions depends on the structure of the underlying economic variables. Different parameter values give rise to different degrees of long-horizon predictability. We explore this issue by simulating two standard models in the literature: the Monetary and Taylor Rule models.

In the Monetary model with stationary fundamentals of equation (7), the matrix $\mathbf{\Delta X_t}$ is defined as:

$$\mathbf{\Delta X_t} \equiv \begin{bmatrix} \Delta m_t & \Delta m_t^* & \Delta y_t & \Delta y_t^* & \Delta q_t & \Delta v_t & \Delta v_t^* & \rho_t \end{bmatrix}. \tag{27}$$

The stationary fundamental in this model is $u_{2t} = -\rho_t$, which is typically unobservable to the econometrician.

The risk premium equals:

$$\rho_t = E_t^m s_{t+1} - s_t - (f_t - s_t), \tag{28}$$

where $E_t^m s_{t+1}$ is the market expectation of the exchange rate at time $t+1$, s_t is the spot exchange rate, and f_t is the forward exchange rate.

The risk premium is unobservable because the market expectation of the exchange rate is typically unobservable to the econometrician. For our exercise, we assume that expectations are formed as in the surveys. That is, we calculate the risk premium from survey data (Consensus Forecasts) by assuming that the survey data is a correct measure of market expectations of the exchange rate. 3-month forecasts of the exchange rate are available for 8 countries: Canada, Denmark, Germany, Japan, Norway, Switzerland, UK, and US during 1989Q4-2007Q2.[11] For each country, ρ_t is calculated from equation (28) with $E_t^m s_{t+1}$ replaced by 3-month forecasts of the exchange rate. Figure 3 shows the risk premium that is calculated using the survey data. The risk premium appears to be stationary in the plot. Using the Augmented Dickey-Fuller test, we reject the unit root hypothesis for the risk premium at a 1% significance level for all exchange rates (Table 1).

Most remaining data are obtained from the G10+ dataset provided by Haver Analytics. The money demand shock v_t is recovered from the money demand function:

$$m_t - p_t = \alpha + \gamma y_t + \beta i_t + v_t. \tag{29}$$

The money stock m_t is the seasonally adjusted M2 for all countries except Japan, for which it is M2 plus CDs.[12] p_t is the CPI index and y_t is GDP. The short-term nominal interest rate is measured using 3-month treasure bill rate in Canada and the US, and three-month interbank offer rate for Denmark, Germany, Norway, Switzerland, and the UK. The short-term interest rate in Japan is measured using the 3-month Certificate of Deposit (Gensaki) rate.

We cannot reject the unit root hypothesis for m_t, p_t, y_t, and i_t and the null hypothesis that these variables are not cointegrated at the 5% significance level for most countries. We proceed to take first differences when estimating (29).[13] The OLS regression errors are recovered as Δv_t. Together with other economic variables, we build $\Delta \mathbf{X}_t$ and estimate a VAR(1) process for $\Delta \mathbf{X}_t$.[14] The coefficient matrix and the variance-covariance matrix

---

[11]See Appendix A.4.2 for more details about the data.

[12]Norway's M2 data is from International Financial Statistics and seasonally adjusted using EViews.

[13]Using the Augmented Dickey-Fuller test with a constant and time trend, we fail to reject unit roots for m_t, p_t, y_t, and i_t at various lags for most countries. Using the same test, we fail to reject the null hypothesis that m_t, p_t, y_t, and i_t are not cointegrated at a 1% significance level for all countries.

[14]We restrict the lag length in the VAR to 1 due to the short sample size and the large number of parameters that need to be

of residuals are used to simulate exchange rates with $b = 0.97$ and this process is detailed in Appendix A.4. Long-horizon regressions in equation (5) are estimated with the simulated exchange rate data, where the deviation of the exchange rate from its long-run equilibrium level is defined as $z_t = s_t - [m_t - m_t^* - \gamma(y_t - y_t^*) + q_t - (v_t - v_t^*)]$.

Figure 4(a) shows $R^2(h)$ at various horizons for different countries. Two interesting findings are noted. First, the R-squared is generally small in the short-horizon regression: it is about 0.05 or less in 5 out of 7 countries. Second, in 6 out of 7 countries, $R^2(h)$ displays a hump shape across h. In some countries, the increase of $R^2(h)$ across h is substantial. It rises from about 0.07 to more than 0.3 for Germany and from about 0.11 to 0.28 for Switzerland. Our results confirm the findings in Proposition 2 that the population R-squared does not converge to zero even when the discount factor approaches one, as long as $f_{2,t}$ and $u_{2,t}$ are stationary. The hump shape of $R^2(h)$ supports the notion that long-horizon predictability can exist even when short-horizon predictability is lacking.

In a robustness check, we set up the simulation with cross-country differences of economic variables entering $\Delta \mathbf{X}_t$:

$$\Delta \mathbf{X}_t = \left[\begin{array}{ccccc} \Delta(m_t - m_t^*) & \Delta(y_t - y_t^*) & \Delta q_t & \Delta(v_t - v_t^*) & \rho_t \end{array} \right].$$

We estimate a VAR(1) process for $\Delta \mathbf{X}_t$ and simulate exchange rates in the same way as before. Simulation results are reported in Figure 4(b). The shapes of the $R^2(h)$'s are qualitatively similar to those found in the Monetary model, though the $R^2(h)$ for Germany is significantly larger than before.

Next, the Taylor Rule model in equation (8) is simulated. In this example, $f_{1,t} = p_t - p_t^*$ and $f_{2,t} + u_{2,t} = -\left(r_g(y_t^g - y_t^{*g}) + r_\pi(\pi_t - \pi_t^*) + v_t - v_t^* + \rho_t \right)$. Data exhibits strong evidence that p_t and p_t^* are I(1) and ρ_t is I(0). Unit root test results for other variables, however, are mixed. As previously mentioned, it may be difficult to distinguish between an I(1) process and an I(0), but persistent, process. We follow our setup in Section 3.2 and assume that these variables are also I(0).

Following the definition of $\Delta \mathbf{X}_t$ in Section 3.2, we have:

$$\Delta \mathbf{X}_t = \begin{bmatrix} y_t^g & y_t^{g*} & \pi_t & \pi_t^* & v_t & v_t^* & \rho_t \end{bmatrix}. \tag{30}$$

Notice that since p_t and p_t^* are logarithms of prices, Δp_t and Δp_t^* are the same as π_t and π_t^*. So we do not include Δp_t and Δp_t^* in $\Delta \mathbf{X}_t$. GDP gaps y_t^g and y_t^{g*} are quadratically detrended GDP.[15] v_t and v_t^* are residuals

---

estimated.

[15]Using other detrending methods, such as the H-P filter, does not change our results qualitatively.

of regressing the policy rate on the output gap and CPI inflation in each country. We estimate a VAR(1) process for $\Delta \mathbf{X}_t$. The coefficient matrix and variance-covariance matrix of residuals are used to simulate exchange rates. z_t in long-horizon regressions is defined in four different ways, depending on the specification of $f_{2,t}$:

- Case 1: $z_t = s_t - p_t + p_t^* - \frac{b}{1-b}\left(r_g(y_t^g - y_t^{*g}) + r_\pi(\pi_t - \pi_t^*) + v_t - v_t^* + \rho_t\right)$;

- Case 2(a): $z_t = s_t - p_t + p_t^* - \frac{b}{1-b}\left(r_g(y_t^g - y_t^{*g}) + r_\pi(\pi_t - \pi_t^*)\right)$;

- Case 2(b): $z_t = s_t - p_t + p_t^* - \frac{b}{1-b}\left(r_g(y_t^g - y_t^{*g}) + r_\pi(\pi_t - \pi_t^*) + v_t - v_t^*\right)$

- Case 3: $z_t = s_t - p_t + p_t^*$.

These four cases differ in terms of how the stationary fundamental $f_{2,t}$ is defined. For instance, Case 2(a) sets $f_{2,t} = -\left(r_g(y_t^g - y_t^{*g}) + r_\pi(\pi_t - \pi_t^*)\right)$ and $u_{2,t} = -(v_t - v_t^* + \rho_t)$, a case considered by Molodtsova and Papell (2009). In Case 3, $(u_{2,t} = -\left(r_g(y_t^g - y_t^{*g}) + r_\pi(\pi_t - \pi_t^*) + v_t - v_t^* + \rho_t\right))$ is used in studies of long-run Purchasing Power Parity (PPP), for instance, Chinn and Meese (1995).

Figure 5 shows simulation results for each case. We notice several interesting findings in the long-horizon regressions:

- As in the Monetary model, $R^2(h)$ is small (less than 0.1) in most short-horizon regressions.

- In most cases, the long-horizon regressions have much larger $R^2(h)$ than the short-horizon regressions for some time horizon $h > 1$, in particular for Cases 2(b) and 3.

- Case 1 performs the best in all countries at the short horizon. This is intuitive as f_{2t} includes all stationary fundamentals, leaving nothing unobserved by the econometrician.

- At very long horizons, the R-squared remains relative high in the Taylor Rule model compared to that in the Monetary model. This is because the stationary component $f_{2t} + u_{2t}$ is very persistent in the Taylor Rule model. Because this simulation is simply an illustration of the long-horizon predictability of Taylor Rule fundamentals, we shy away from the fact that some fundamentals are borderline non-stationary in the data. Simulation results come close to what we find in the Monetary model if, say, Δy_t^g and Δy_t^{g*} are used in f_{2t} instead of their levels.

5   Conclusion

Engel and West (2005) propose an explanation to Meese and Rogoff's (1983a) finding that exchange rate models cannot forecast exchange rates better than the random walk. With a simple and reasonable modification to their

nonstationarity assumption about economic fundamentals, we find that the E-W explanation is also consistent with the finding that exchange rate models can forecast exchange rates at long horizons. When we allow stationary, and potentially unobservable, fundamentals in the present-value asset pricing models studied in Engel and West (2005), the population R-squared of long-horizon regressions can increase substantially although the R-squared is close to zero in the short-horizon regression. A potential candidate for the stationary fundamental is the risk premium. We calculate the risk premium from survey data and find strong evidence of stationarity. The risk premium, along with other fundamentals, and the model coefficients estimated from the data are used to simulate exchange rates from two standard exchange rate models. The fundamentals can forecast the simulated exchange rates at long horizons, but not at the short one.

Our results suggest that the long-horizon predictability of the exchange rate, which has been found in the literature, may come from some stationary unobservable fundamentals, such as the risk premium. If exchange rate predictability is coming from the risk premium, perhaps panel estimation would perform better since there may be a common component to the risk premium across dollar exchange rates. Groen (2000) and Mark and Sul (2001) find exchange rate predictability by using panel data. Rogoff and Stavrakeva (2008) find forecast improvement after allowing for common cross-country shocks in their panel forecasting specification, although the improvement is not entirely robust to the forecast window.

For simplicity, we only consider a linear model. Incorporating other features, such as nonlinearity, has also been found to be successful in forecasting exchange rates more accurately. For instance, Kilian and Taylor (2003) find strong evidence of predictability at horizons of 2 to 3 years, but not at shorter horizons. Faust, Rogers and Wright (2003) and Molodtsova, Nikolsko-Rzhevskyy and Papell (2008) also find exchange rate predictability using real-time, but not revised, data. Engel and West (2005) and Chen, Rogoff, and Rossi (2008) find the connection between exchange rates and fundamentals in the opposite direction: exchange rates are helpful to forecast fundamentals. Further empirical research along these lines may be fruitful.

In this paper, we do not study the case of a finite sample. In small samples, long-horizon regressions may have serious size distortions when asymptotic critical values are used. For instance, see Mark (1995) and Campbell (2001). It is not clear whether the power advantage of the long-horizon regressions will remain after size is corrected. We also only consider a less general case than Engel and West (2005). $\Delta \mathbf{X_t}$ is assumed to follow a VAR(p) process in this paper while Engel and West (2005) use an ARMA (or equivalently MA(∞)) process. Our setup does not allow $\mathbf{X_t}$ to include stationary or conintegrated variables. The reason to consider a less general setup in this paper is purely technical. We acknowledge that it is desirable to extend our results to a more general setup.

Table 1: Unit Root Tests for the Risk Premium

	Canada	Denmark	Germany	Japan	Norway	Switzerland	UK
ADF test with constant and time trend	-7.552	-7.535	-8.451	-6.982	-8.842	-7.916	-7.368
ADF test with constant only	-8.657	-5.006	-5.599	-5.133	-5.669	-5.321	-7.201

Note:

–Entries are Augmented Dicky-Fuller t-statistics. The unit root hypothesis is rejected at a 99% confidence level in all cases.

–4 lags are included in the above tests. The results do not change qualitatively with the number of lags.

Figure 1:   Asymptotic R-squared with AR(1) Fundamental
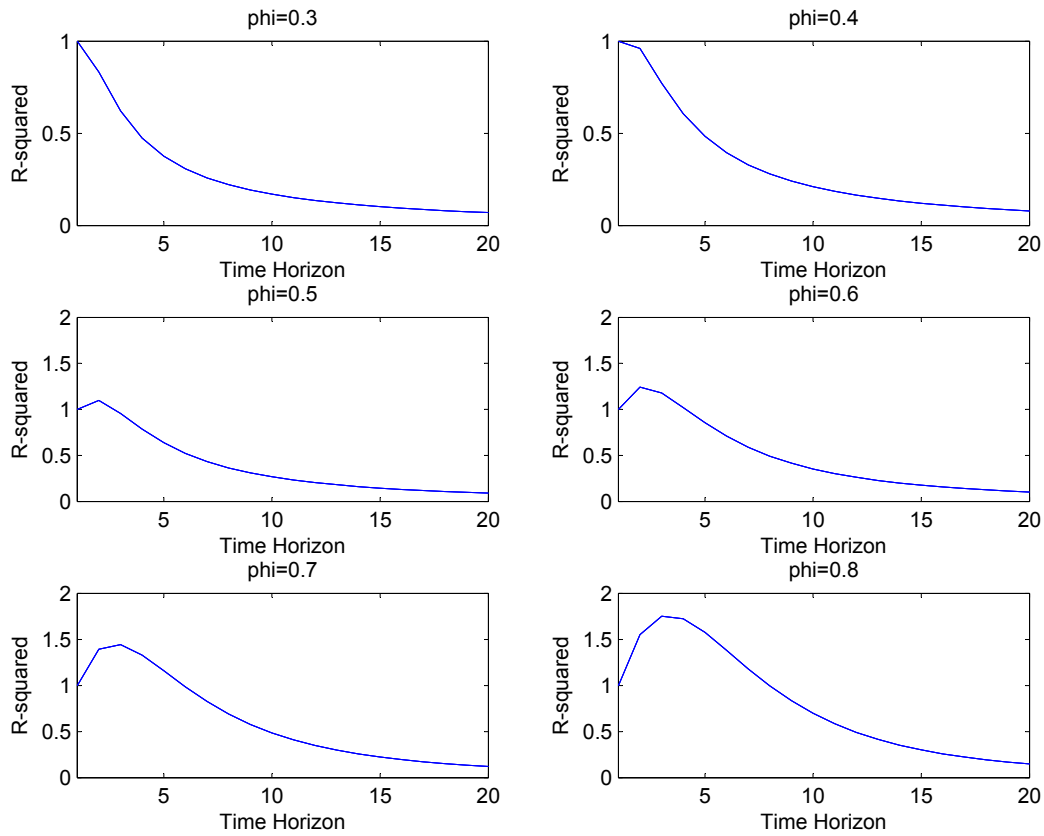
Note:

The R-squared for horizons greater than 1 is normalized by the short-horizon R-squared ($R^2(1)$).

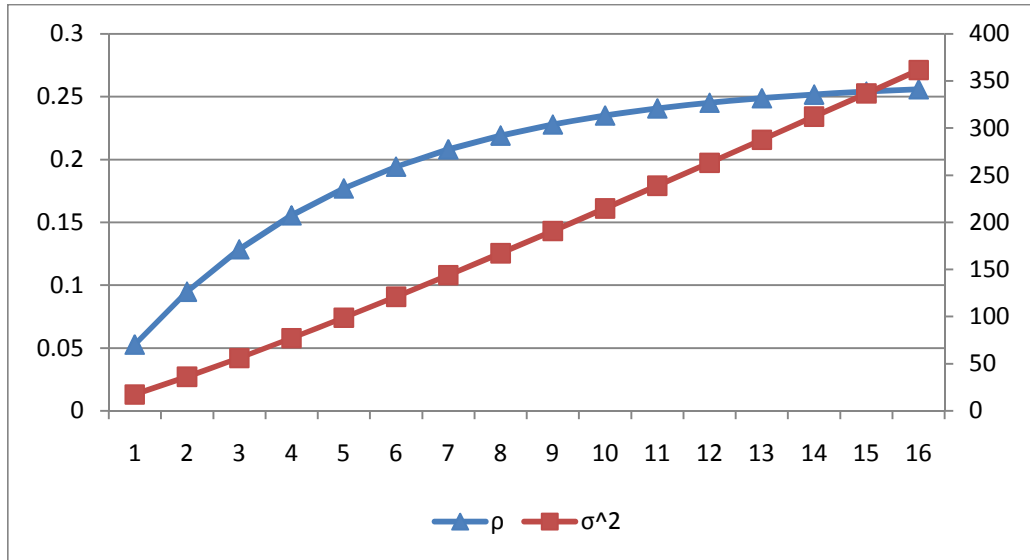Figure 2: ρ_h and $\sigma^2_{\nu_{t+h}}$ at Different Time Horizons

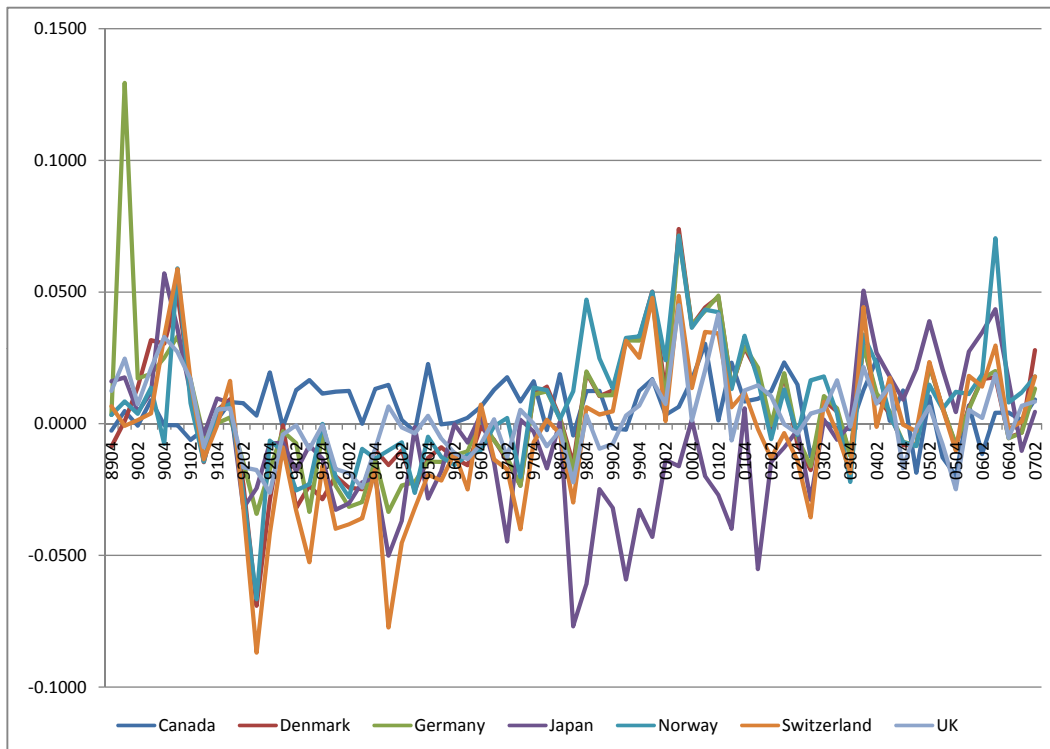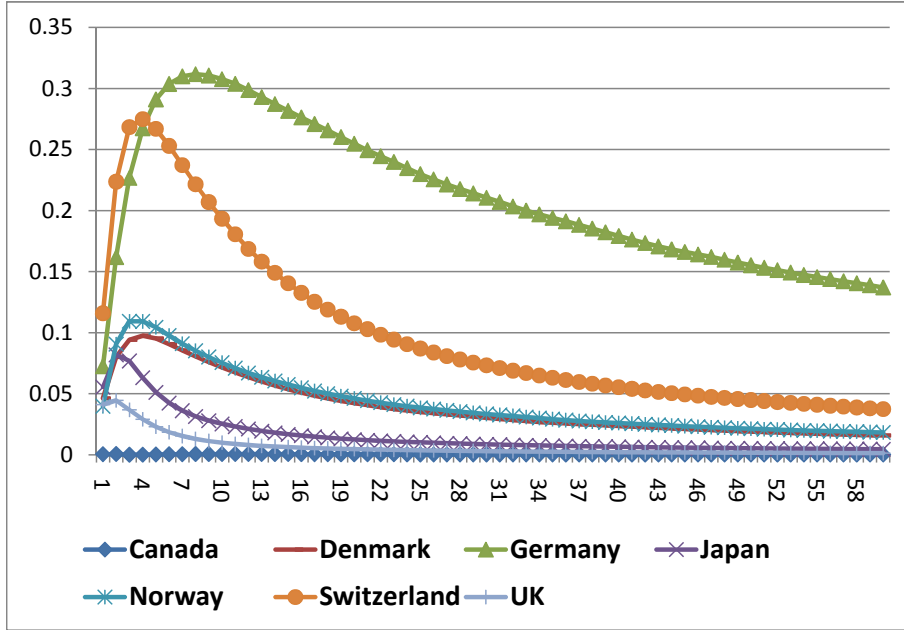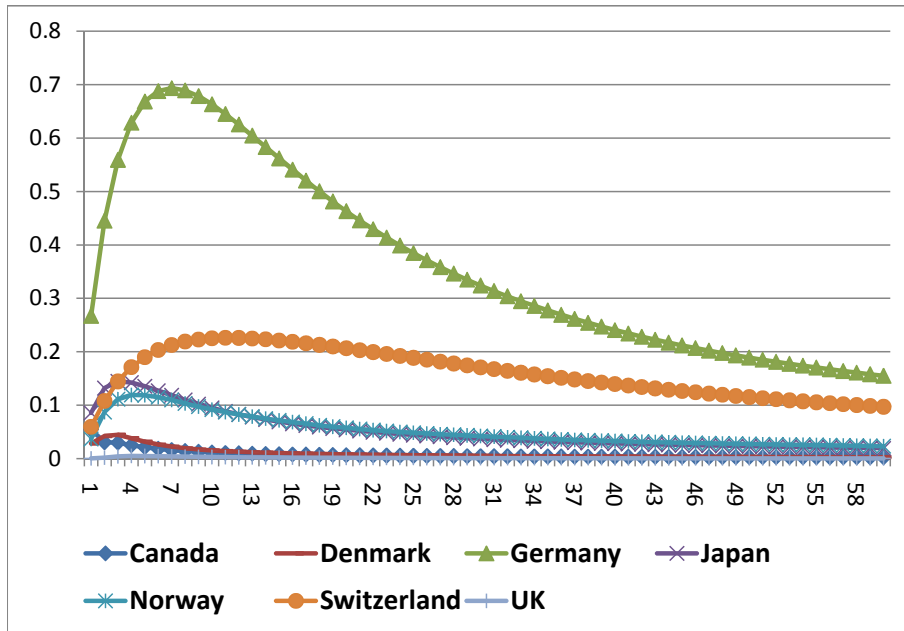Figure 3: Exchange Rate Risk Premium Calculated From Survey Data

Figure 4: Population R-squared with VAR(1) Fundamental: Monetary Model

(a) Monetary Model 1

(b) Monetary Model 2

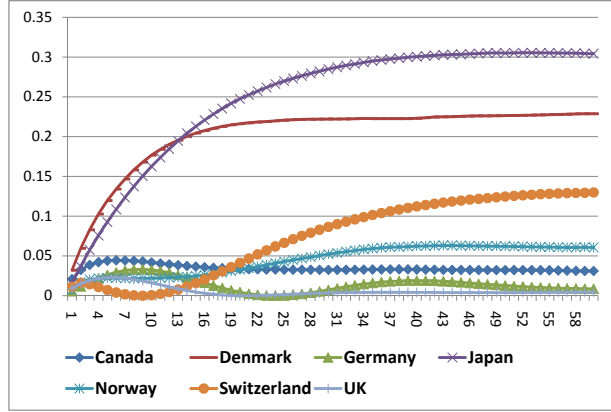–Figure 4(a): $\Delta \mathbf{X_t} \equiv \begin{bmatrix} \Delta m_t & \Delta m_t^* & \Delta y_t & \Delta y_t^* & \Delta q_t & \Delta v_t & \Delta v_t^* & \rho_t \end{bmatrix}$.

–Figure 4(b): $\Delta \mathbf{X_t} = \begin{bmatrix} \Delta(m_t - m_t^*) & \Delta(y_t - y_t^*) & \Delta q_t & \Delta(v_t - v_t^*) & \rho_t \end{bmatrix}$.
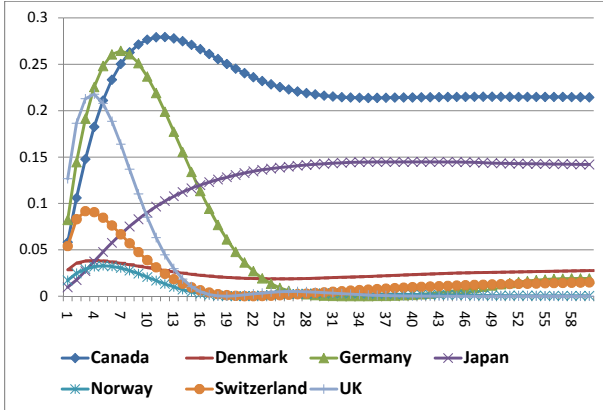
21

Figure 5: Population R-squared with VAR(1) Fundamental: Taylor Rule Model
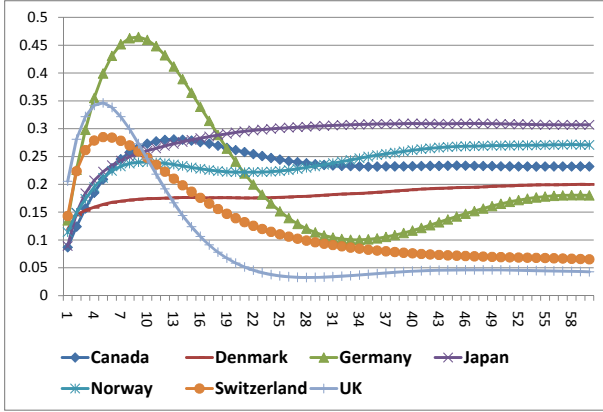
(a) Case 1

(b) Case 2(a)

(c) Case 2(b)
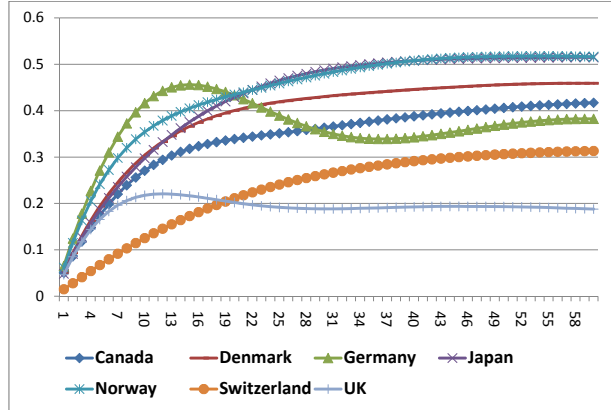
(d) Case 3

–Case 1: $z_t = s_t - p_t + p_t^* - \frac{b}{1-b} \left[r_g(y_t^g - y_t^{*g}) + r_\pi(\pi_t - \pi_t^*) + v_t - v_t^* + \rho_t \right]$

–Case 2(a): $z_t = s_t - p_t + p_t^* - \frac{b}{1-b} \left[r_g(y_t^g - y_t^{*g}) + r_\pi(\pi_t - \pi_t^*) \right]$

–Case 2(b): $z_t = s_t - p_t + p_t^* - \frac{b}{1-b} \left[r_g(y_t^g - y_t^{*g}) + r_\pi(\pi_t - \pi_t^*) + v_t - v_t^* \right]$

–Case 3: $z_t = s_t - p_t + p_t^*$.

References

[1] Adrian, T., E. Etula, and H. S. Shin, 2009. Risk Appetite and Exchange Rates, Federal Reserve Bank of New York Staff Report no. 361.

[2] Berkowitz, J. and L. Giorgianni, 2001. Long-Horizon Exchange Rate Predictability? *The Review of Economics and Statistics* 83(1), 81-91.

[3] Campbell, J., 2001. Why Long Horizon? A Study of Power Against Persistent Alternatives, *Jounal of Empirical Finance* 8, 459-491.

[4] Campbell, J. and R. Shiller, 1987. Cointegration and Tests of Present Value Models, *Journal of Political Economy* 95, 1062-1088.

[5] Campbell, J. and R. Shiller, 1988. The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, *The Review of Financial Studies* 1(3), 195-228.

[6] Chen, Y., K. Rogoff, and B. Rossi, 2008. Can Exchange Rates Forecast Commodity Prices?, NBER Working Paper No.13901.

[7] Chen, Y. and K. Tsang, 2009. What Does the Yield Curve Tell Us About Exchange Rate Predictability?, Working Paper, University of Washington and Virginia Tech.

[8] Cheung, Y., M. D. Chinn, and A. Pascual, 2005. Empirical Exchange Rate Models of the Nineties: Are Any Fit to Survive? *Journal of International Money and Finance* 24, 1150-1175.

[9] Chinn, M. D. and R. A. Meese, 1995. Banking on Currency Forecasts: How Predictable Is Change in Money? *Journal of International Economics* 38, 161-178.

[10] Clarida, R., L. Sarno, M. Taylor, and G. Valente, 2003. The Out-of-sample Success of Term Structure Models as Exchange Rate Predictors: A Step Beyond, *Journal of International Economics* 60, 61-83.

[11] Engel, C., N. Mark, and K. West, 2007. Exchange Rate Models are Not as Bad as You Think, forthcoming in NBER Macroeconomics Annual 2007.

[12] Engel, C., and K. D. West, 2004. Accounting for Exchange-Rate Variability in Present-Value Models When the Discount Factor Is Near 1, *American Economic Review*, Papers and Proceedings 94, May 2004, 119-125.

[13] Engel, C. and K. West, 2005. Exchange Rates and Fundamentals, *Journal of Political Economy* 113, 485-517.

[14] Fama, E. and K. French, 1988. Dividend Yields and Expected Stock Returns, *Journal of Financial Economics* 113, 485-517.

[15] Faust, J., J. Rogers, and J. Wright, 2003. Exchange Rate Forecasting: The Errors We've Really Made, *Journal of International Economics* 60, 35-59.

[16] Groen, J., 2000. The Monetary Exchange Rate Model as a Long-run Phenomenon, *Journal of International Economics* 52, 299-319.

[17] Kilian, L., 1999. Exchange Rates And Monetary Fundamentals: What Do We Learn From Long-Horizon Regressions? *Journal of Applied Econometrics* 14, 491-510.

[18] Kilian, L. and M. Taylor, 2003. Why Is It So Difficult to Beat the Random Walk Forecast of Exchange Rates? *Journal of International Economics* 60, 85-107.

[19] MacDonald, R. and M. P. Taylor, 1994. The Monetary Model of the Exchange Rate: Long-Run Relationships, Short-Run Dynamics, and How to Beat a Random Walk, *Journal of International Money and Finance* 13, 276-290.

[20] Mark, N., 1995. Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability, *American Economic Review* 85, 201-218.

[21] Mark, N. and D. Sul, 2001. Nominal Exchange Rates and Monetary Fundamentals Evidence from a Small Post-Bretton Woods Panel, *Journal of International Economics* 53, 29-52.

[22] Mark, N. and D. Sul, 2006. The Power of Long-Horizon Predictive Regression Tests, Working Paper, University of Notre Dame and University of Auckland.

[23] Meese, R. A., and K. Rogoff, 1983a. Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample? *Journal of International Economics* 14, 3-24.

[24] Meese, R. A., and K. Rogoff, 1983b. The Out of Sample Failure of Empirical Exchange Models, in J. Frenkel, ed., *Exchange Rates and International Macroeconomics* (University of Chicago Press, Chicago).

[25] Molodtsova, T. and D. Papell, 2009. Out-of-sample Exchange Rate Predictability with Taylor Rule Fundamentals, *Journal of International Economics* 77, 167-180.

[26] Molodtsova, T., A. Nikolsko-Rzhevskyy, and D. Papell, 2008. Taylor Rules with Real-Time Data: A Tale of Two Countries and One Exchange Rate, *Journal of Monetary Economics* 55, S63-S79.

[27] Nason, J. and J. Rogers, 2008. Exchange Rates and Fundamentals: A Generalization, FRB-Atlanta Working Paper 2008-16.

[28] Rogoff, K. and V. Stavrakeva, 2008. The Continuing Puzzle of Short-Horizon Exchange Rate Forecasting, Working Paper, Harvard University and the Brookings Institution.

[29] Rossi, B., 2005. Testing Long-horizon Predictive Ability with High Persistence and the Meese-Rogoff Puzzle, *International Economic Review* 46(1), 61-92.

[30] Wang, J. and J. J. Wu, 2008. The Taylor Rule and Interval Forecast for Exchange Rates, Globalization and Monetary Policy Institute Working Paper No. 22, International Finance Discussion Paper No. 963.

[31] West, K. 1988. Dividend Innovations and Stock Price Volatility, *Econometrica* 56, 37-61.

APPENDIX

A.1 Connection with Campbell (2001)

In this section, we draw the analogy between Campbell's (2001) long-horizon exercise and Engel, Mark, and West's (EMW, 2007) long-horizon exercise that we use in Section 3.2.

A useful starting place is to notice the close analogy of the Campbell-Shiller log linearization of the stock price model to the EMW log linearization of the exchange rate model. Take equation (2) of Campbell and Shiller (1988):

$$r_{t+1} = k + (1 - \rho)d_t + \rho p_{t+1} - p_t, \tag{A.1.1}$$

where d_t is the log of dividends, and p_t is the log of the stock price. This is exactly equation (2'), but we have rearranged and canceled terms on the right-hand side. We have also changed the notation on the left-hand side to the one used in Campbell (2001).

In Campbell (2001), x_t is defined by $\lambda x_t \equiv E_t r_{t+1}$. So from equation (A.1.1) we have:

$$\lambda x_t = k + (1 - \rho)d_t + \rho E_t p_{t+1} - p_t. \tag{A.1.2}$$

We can rewrite this equation (dropping the intercept for convenience) as:

$$p_t = (1 - \rho)d_t + \rho E_t p_{t+1} - \lambda x_t. \tag{A.1.3}$$

Now compare to Example 3 in Section 3.2:

$$s_t = (1 - b)f_{1,t} + b E_t s_{t+1} + b u_{2,t}. \tag{A.1.4}$$

Here is the mapping of notations from equation (A.1.3) of Campbell to equation (A.1.4) of EMW:

Table 2: Mapping of Campbell to EMW

Campbell (2001)	p_t	d_t	$-\lambda x_t$	ρ
EMW (2007)	s_t	$f_{1,t}$	$b u_{2,t}$	b

Campbell (2001) assumes that x_t is stationary and follows an AR(1) process:

$$x_{t+1} = \phi x_t + u_{t+1}. \tag{A.1.5}$$

Campbell (2001) closes his model by modeling r_{t+1}:

$$r_{t+1} - E_t r_{t+1} = v_{t+1}, \tag{A.1.6}$$

where v_{t+1} is the expectation error and possibly correlated with u_{t+1}. Comparing the above equation with (A.1.1), we have:

$$p_{t+1} - E_t p_{t+1} = v_{t+1}/\rho. \tag{A.1.7}$$

EMW (2007) models $u_{2,t}$ the same way as Campbell (2001) models x_t

$$u_{2,t+1} = \phi u_{2,t} + u_{t+1}. \tag{A.1.8}$$

They call $u_{2,t}$ the risk premium, which is unobservable to the econometrician. This interpretation is consistent with the Monetary model. $f_{1,t}$ is assumed to be a random walk:

$$f_{1,t+1} = f_{1,t} + \varepsilon_{t+1}. \tag{A.1.9}$$

Substituting equations (A.1.8) and (A.1.9) into (A.1.4), we have:

$$s_{t+1} - E_t s_{t+1} = \varepsilon_{t+1} + \frac{b}{1 - b\phi} u_{t+1}. \tag{A.1.10}$$

So we have that $\varepsilon_{t+1} + \frac{b}{1-b\phi} u_{t+1}$ is analogous to v_{t+1}/ρ in equation (A.1.7). Note that because EMW assume ε_{t+1} and u_{t+1} are uncorrelated, then EMW's $\varepsilon_{t+1} + \frac{b}{1-b\phi} u_{t+1}$ (analogous to Campbell's v_{t+1}) is necessarily correlated with EMW's u_{t+1} (analogous to Campbell's u_{t+1}.) Mark and Sul (2006) find that long-horizon regressions have power advantage over the short-horizon regression when the regressor is endogenous (u_{t+1} and v_{t+1} are correlated.) In this example, we show that the endogeneity exists under EMW's setup.

It is useful to recognize that the long-horizon regression that Campbell simulates does not require a model of the stochastic process of d_t (which is $f_{1,t}$ in EMW.) All he needs to do is model (ex-post returns) the ex-ante risk

premium, λx_t. The only assumption Campbell actually makes is on the stochastic process of the risk premium, and it is the same as in EMW. Campbell (2001) implicitly assumes that the change in d_t is stationary. So, Campbell's simulations are consistent with any I(1) or I(0) data generating process for d_t. As we will now note, EMW's long-horizon regression does require a model of their $f_{1,t}$ (which is analogous to Campbells d_t), because their simulations require that we be able to solve the model for s_t, or at least $s_{t+1} - s_t$.

Next, we turn to the long-horizon regressions, which appear to be different. Campbell's long-horizon regression regresses $r_{t+1} + ... + r_{t+k}$ on x_t, while EMW regress $s_{t+k} - s_t$ on $f_{1,t} - s_t$. While these seem to be completely different regressions, under the assumptions made about the stochastic processes, they are in fact very similar. First, note that under the assumptions of EMW, we have:

$$f_{1,t} - s_t = -\frac{b}{1 - b\phi}u_{2,t}.$$

So, EMW's r.h.s. variable in the long-horizon regression, $f_{1,t} - s_t$, is just proportional to their risk premium, $-bu_{2,t}$. Likewise, Campbell's r.h.s. variable, x_t, is just proportional to his risk premium, λx_t. If EMW had run the same regression for exchange rates as Campbell does for stock prices, then, their r.h.s. variable would be $(1 - b\phi)(f_{1,t} - s_t)$, which is analogous to Campbell's r.h.s. variable, x_t, because Campbell normalizes λ to one.

What is different is the dependent variable. From equation (A.1.1), we have:

$$r_{t+1} + ... + r_{t+k} = (1 - \rho)d_t + \rho p_{t+1} - p_t + ... + (1 - \rho)d_{t+k-1} + \rho p_{t+k} - p_{t+k-1}$$

$$= p_{t+k} - p_t + (\rho - 1)\left[p_{t+k} - d_{t+k-1} + ... + p_{t+1} - d_t\right]. \tag{A.1.11}$$

The l.h.s. variable in EMW would be, by analogy to equation (A.1.11) above

$$s_{t+k} - s_t + (b - 1)\left[s_{t+k} - f_{1,t+k-1} + ... + s_{t+1} - f_{1,t}\right]. \tag{A.1.12}$$

It is straightforward that:

$$s_{t+k} - s_t + (b - 1)\left[s_{t+k} - f_{1,t+k-1} + ... + s_{t+1} - f_{1,t}\right]$$

$$= s_{t+k} - s_t + (b - 1)\left[s_{t+k} - s_t - f_{1,t+k-1} + ... + s_{t+1} - f_{1,t} + s_t\right]$$

$$= b(s_{t+k} - s_t) + (b - 1)\sum_{j=0}^{k-1}(s_{t+j} - f_{1,t+j}). \tag{A.1.13}$$

In short, EMW's long-horizon regression regresses $s_{t+k} - s_t$ on $f_{1,t} - s_t$, but if they had followed the Campbell

methodology, they would regress $b(s_{t+k} - s_t) + (b-1)\sum_{j=0}^{k-1}(s_{t+j} - f_{1,t+j})$ on $(1-b\phi)(f_{1,t} - s_t)$. Note that the two methodologies are very similar for b close to one.

A.2 Exchange Rate Models

A.2.1 Monetary Model

Assume the money market clearing condition in the home country is:

$$m_t = p_t + \gamma y_t - \alpha i_t + v_t,$$

where m_t is the log of the money supply, p_t is the log of the aggregate price, i_t is the nominal interest rate, y_t is the log of output, and v_t is a money demand shock. A symmetric condition holds in the foreign country and we use an asterisk in subscript to denote variables in the foreign country. Subtracting foreign money market clearing condition from the home, we have:

$$i_t - i_t^* = \frac{1}{\alpha}\left[-(m_t - m_t^*) + (p_t - p_t^*) + \gamma(y_t - y_t^*) + (v_t - v_t^*)\right]. \tag{A.2.1}$$

The nominal exchange rate is equal to its purchasing power value plus the real exchange rate:

$$s_t = p_t - p_t^* + q_t. \tag{A.2.2}$$

The uncovered interest rate parity condition in financial market takes the form

$$E_t s_{t+1} - s_t = i_t - i_t^* + \rho_t, \tag{A.2.3}$$

where ρ_t is the uncovered interest rate parity shock. Substituting equations (A.2.1) and (A.2.2) into (A.2.3), we have:

$$s_t = (1-b)\left[m_t - m_t^* - \gamma(y_t - y_t^*) + q_t - (v_t - v_t^*)\right] - b\rho_t + bE_t s_{t+1}, \tag{A.2.4}$$

29

where $b = \alpha/(1+\alpha)$. Solving s_t recursively and applying the "no-bubbles" condition, we have:

$$s_t = E_t \left\{ (1-b) \sum_{j=0}^{\infty} b^j \left[m_{t+j} - m_{t+1}^* - \gamma(y_{t+j} - y_{t+1}^*) + q_{t+1} - (v_{t+j} - v_{t+j}^*) \right] - b \sum_{j=1}^{\infty} b^j \rho_{t+j} \right\}. \qquad (A.2.5)$$

In the standard Monetary model, such as Mark (1999), purchasing power parity ($q_t = 0$) and uncovered interest rate parity hold ($\rho_t = 0$). Furthermore, it is assumed that the money demand shock is zero ($v_t = v_t^* = 0$) and $\gamma = 1$. Equation (A.2.5) then reduces to:

$$s_t = E_t \left\{ (1-b) \sum_{j=0}^{\infty} b^j \left(m_{t+j} - m_{t+j}^* - (y_{t+j} - y_{t+j}^*) \right) \right\}.$$

If we release the above assumptions and allow both observable and unobservable fundamentals in the model, $f_{1t} = m_t - m_t^* - \gamma(y_t - y_t^*) + q_t - (v_t - v_t^*)$, $u_{1t} = 0$, $f_{2t} = 0$, and $u_{2t} = -\rho_t$ in this example. Assume the risk premium ρ_t is stationary. By our definition of the error correction term z_t:

$$z_t \equiv s_t - f_{1t} = s_t - (m_t - m_t^* - \gamma(y_t - y_t^*) + q_t - (v_t - v_t^*)) \qquad (A.2.6)$$

$$= p_t - p_t^* - (m_t - m_t^*) + \gamma(y_t - y_t^*) + v_t - v_t^*. \qquad (A.2.7)$$

In the second part of the above equation, we replace $s_t - q_t$ with $p_t - p_t^*$. If we assume purchasing power parity ($q_t = 0$), the error correction term becomes:

$$z_t \equiv s_t - f_{1t} = s_t - (m_t - m_t^* - \gamma(y_t - y_t^*) - (v_t - v_t^*)). \qquad (A.2.8)$$

A.2.2 Taylor Rule Model

We follow Engel and West (2005) to assume that both countries follow the Taylor Rule and the foreign country targets the exchange rate in its Taylor Rule. The interest rate differential is:

$$i_t - i_t^* = r_s(s_t - \bar{s}_t^*) + r_g(y_t^g - y_t^{*g}) + r_\pi(\pi_t - \pi_t^*) + v_t - v_t^*, \qquad (A.2.9)$$

where $\bar{s}_t^*$ is the targeted exchange rate. Assume that monetary authorities target the PPP level of the exchange rate: $\bar{s}_t^* = p_t - p_t^*$. Substituting this condition and the interest rate differential of the UIP condition, we have:

$$s_t = (1-b)(p_t - p_t^*) - b\left[r_g(y_t^g - y_t^{*g}) + r_\pi(\pi_t - \pi_t^*) + v_t - v_t^*\right] - b\rho_t + bE_t s_{t+1}, \tag{A.2.10}$$

where $b = \frac{1}{1+r_s}$. Comparing to our models with both observable and unobservable fundamentals, in this example, $f_{1t} = p_t - p_t^*$, $u_{1t} = 0$, and $f_{2t} + u_{2t} = -\left[r_g(y_t^g - y_t^{*g}) + r_\pi(\pi_t - \pi_t^*) + v_t - v_t^* + \rho_t\right]$. Under the condition that f_{1t} and f_{2t} are I(1) and ρ_t is I(0), s_t, f_{1t}, and f_{2t} are cointegrated with the conintegrating vector of $\begin{pmatrix} 1 & -1 & -\frac{b}{1-b} \end{pmatrix}$. The error correction term is $z_t \equiv s_t - f_{1t} - \frac{b}{1-b}f_{2t}$.

A.3   Proof of Propositions

Proof of 1(a). First note that using equation(6) and the definition of $\alpha(b)$, we have:

$$z_t = bu_{2t} + \alpha'(b)\sum_{j=1}^{\infty} b^j E_t \Delta \mathbf{X}_{t+j}. \tag{A.3.1}$$

Notice that:

$$\Delta \mathbf{Y}_{t+j} = F^j \Delta \mathbf{Y}_t + \xi_{t+j} \tag{A.3.2}$$

$$\xi_{t+j} \equiv \sum_{i=1}^{j} F^{j-i}\mathbf{v}_{t+i}.$$

Hence, $E_t \Delta \mathbf{X}_{t+j} = \iota' F^j \Delta \mathbf{Y}_t$, and

$$z_t = b\alpha_{22}' \iota' \Delta \mathbf{Y}_t + \alpha'(b)\iota'\left(\sum_{j=1}^{\infty} b^j F^j\right)\Delta \mathbf{Y}_t$$

$$= \beta'(b)\Delta \mathbf{Y}_t.$$

For $b < 1$, z_t is stationary because $\Delta \mathbf{Y}_t$ is stationary.

Now consider the one-step-ahead regression. Manipulations of equation (6) yield:

$$\Delta s_{t+1} = \Delta f_{1t} + \frac{b}{1-b}\Delta f_{2t} + \alpha'(b)\sum_{j=0}^{\infty} b^j\left(E_{t+1}\Delta \mathbf{X}_{t+1+j} - E_t \Delta \mathbf{X}_{t+j}\right).$$

31

Equation (A9) of Engel and West (2005) shows that:

$$\sum_{j=0}^{\infty} b^j \left(E_{t+1}\Delta\mathbf{X}_{t+1+j} - E_t\Delta\mathbf{X}_{t+1+j} \right) = (I_n - \Phi(b))^{-1}\varepsilon_{t+1}.$$

Substituting this and (A.3.1) into the expression for Δs_{t+1}, we have:

$$
\begin{aligned}
\Delta s_{t+1} &= \alpha^{'}(b)\Delta\mathbf{X}_t - bu_{2t} + \alpha^{'}(b)\left\{ \sum_{j=0}^{\infty} b^j \left(E_t\Delta\mathbf{X}_{t+1+j} - E_t\Delta\mathbf{X}_{t+j} \right) + (I_n - \Phi(b))^{-1}\varepsilon_{t+1} \right\} \\
&= \frac{1-b}{b}\alpha^{'}(b)\sum_{j=1}^{\infty} b^j E_t\Delta\mathbf{X}_{t+j} - bu_{2t} + \alpha^{'}(b)(I_n - \Phi(b))^{-1}\varepsilon_{t+1} \\
&= \frac{1-b}{b}z_t - u_{2t} + \alpha^{'}(b)(I_n - \Phi(b))^{-1}\varepsilon_{t+1}.
\end{aligned}
$$

Define the following matrices:

$$\underset{n\times nh}{\Gamma^{'}_h(b)} \equiv \left[(I_n - \Phi(b))^{-1}, ..., (I_n - \Phi(b))^{-1} \right], \qquad \underset{nh\times 1}{\varepsilon_{t,h}} \equiv \left[\varepsilon^{'}_{t+1}, ..., \varepsilon^{'}_{t+h} \right]^{'}.$$

Then, by iterating the one-step-ahead regression h times, we have:

$$s_{t+h} - s_t = \frac{1-b}{b}\sum_{k=0}^{h-1} z_{t+k} - \sum_{k=0}^{h-1} u_{2t+k} + \alpha^{'}(b)\Gamma^{'}_h(b)\varepsilon_{t,h}. \tag{A.3.3}$$

Using (A.3.2) and the fact that for $b < 1$, $\beta(b)\beta^{'}(b)$ is non-singular, we obtain:

$$
\begin{aligned}
z_{t+k} &= \beta^{'}(b)\Delta\mathbf{Y}_{t+k} \\
&= \beta^{'}(b)\left(F^k\Delta\mathbf{Y}_t + \xi_{t+k} \right) \\
&= \beta^{'}(b)F^k(\beta(b)\beta^{'}(b))^{-1}\beta(b)z_t + \beta^{'}(b)\xi_{t+k}.
\end{aligned}
$$

After some straightforward, but tedious, calculations,

$$\sum_{k=0}^{h-1} \xi_{t+k} = F^{'}_h\varepsilon_{t,h}, \qquad \underset{np\times nh}{F^{'}_h} \equiv \left[\sum_{j=0}^{h-2} F^j\iota \quad \sum_{j=0}^{h-3} F^j\iota \quad ... \quad \sum_{j=0}^{1} F^j\iota \quad I_{np}\iota \quad 0 \right].$$

Plugging these expressions into (A.3.3), we have:

$$
\begin{aligned}
s_{t+h} - s_t &= \frac{1-b}{b}\beta'(b)\left(\sum_{k=0}^{h-1}F^k\right)\left(\beta(b)\beta'(b)\right)^{-1}\beta'(b)z_t + \frac{1-b}{b}\beta'(b)\sum_{k=0}^{h-1}\xi_{t+k} - \alpha'_{22}\iota'\sum_{k=0}^{h-1}\left(F^k\Delta\mathbf{Y}_t + \xi_{t+k}\right) + \beta'(b)\Gamma'_h(b)\varepsilon_{t,h} \\
&\equiv \rho_h z_t - \alpha'_{22}\iota'\left(\sum_{k=0}^{h-1}F^k\right)\Delta\mathbf{Y}_t + \upsilon_{t+h},
\end{aligned}
$$

where

$$
\rho_h \equiv \frac{1-b}{b}\beta'(b)\left(\sum_{k=0}^{h-1}F^k\right)\left(\beta(b)\beta'(b)\right)^{-1}\beta'(b)
$$

$$
\upsilon_{t+h} \equiv \left\{\left(\frac{1-b}{b}\beta'(b) - \alpha'_{22}\iota'\right)F'_h + \alpha'(b)\Gamma'_h(b)\right\}\varepsilon_{t,h}.
$$

∎

Proof of 1(b).

Because $E\Delta\mathbf{X}_t = 0$ and $E(s_{t+h} - s_t) = Ez_t = 0$, we have $R^2(h) = \frac{\{E(s_{t+h}-s_t)z_t\}^2}{Ez_t^2 E(s_{t+h}-s_t)^2}$. First, consider the numerator. Using $z_t = \beta'(b)\Delta\mathbf{Y}_t$ and re-arranging, we have:

$$
s_{t+h} - s_t = \left\{\frac{1-b}{b}\beta'(b) - \alpha'_{22}\iota'\right\}\left(\sum_{k=0}^{h-1}F^k\right)\Delta\mathbf{Y}_t + \upsilon_{t+h}.
$$

Then,

$$
\begin{aligned}
\{E(s_{t+h} - s_t)z_t\}^2 &= \left\{\frac{1-b}{b}\beta'(b) - \alpha'_{22}\iota'\right\}\left(\sum_{k=0}^{h-1}F^k\right)E(\Delta\mathbf{Y}_t\Delta\mathbf{Y}'_t)\beta(b)\beta'(b)E(\Delta\mathbf{Y}_t\Delta\mathbf{Y}'_t)\left(\sum_{k=0}^{h-1}F'^k\right)\left\{\frac{1-b}{b}\beta(b) - \iota\alpha_{22}\right\} \\
&= Ez_t^2\left\{\frac{1-b}{b}\beta'(b) - \alpha'_{22}\iota'\right\}\left(\sum_{k=0}^{h-1}F^k\right)E(\Delta\mathbf{Y}_t\Delta\mathbf{Y}'_t)\left(\sum_{k=0}^{h-1}F'^k\right)\left\{\frac{1-b}{b}\beta(b) - \iota\alpha_{22}\right\},
\end{aligned}
$$

where the last equality is due to the fact that $E^2z_t = \beta'(b)E(\Delta\mathbf{Y}_t\Delta\mathbf{Y}'_t)\beta(b)$ and $(\beta(b)\beta'(b))^{-1}$ exists for $b < 1$.

Now define:

$$
A_{h,b} \equiv \left(\sum_{k=0}^{h-1}F'^k\right)\left\{\frac{1-b}{b}\beta(b) - \iota\alpha_{22}\right\}. \tag{A.3.4}
$$

33

Then, we have:

$$
\begin{aligned}
\{E(s_{t+h} - s_t)z_t\}^2 &= E z_t^2 A'_{h,b} E(\Delta \mathbf{Y}_t \Delta \mathbf{Y}'_t) A_{h,b} \\
A_{h,b} &\overset{b \to 1}{\to} \left(\sum_{k=0}^{h-1} F'^k \right) \left\{ F'(b) \iota \alpha_{21} - \iota \alpha_{22} \right\} \\
&= O(1).
\end{aligned}
$$

Next, consider the denominator:

$$
\begin{aligned}
E(s_{t+h} - s_t)^2 &= \left\{ \frac{1-b}{b} \beta'(b) - \alpha'_{22} \iota' \right\} \left(\sum_{k=0}^{h-1} F^k \right) E(\Delta \mathbf{Y}_t \Delta \mathbf{Y}'_t) \left(\sum_{k=0}^{h-1} F'^k \right) \left\{ \frac{1-b}{b} \beta(b) - \iota \alpha_{22} \right\} + E v_{t+h}^2 \\
&= A'_{h,b} E(\Delta \mathbf{Y}_t \Delta \mathbf{Y}'_t) A_{h,b} + E v_{t+h}^2 \\
E v_{t+h}^2 &= \left\{ \left(\frac{1-b}{b} \beta'(b) - \alpha'_{22} \iota' \right) F'_h + \alpha'(b) \Gamma'_h(b) \right\} \Omega \otimes I_{nh} \left\{ F_h \left(\frac{1-b}{b} \beta(b) - \iota \alpha_{22} \right) + \Gamma_h(b) \alpha(b) \right\},
\end{aligned}
$$

where $\otimes$ denotes the Kronecker product. Using the definitions of F_h and $\Gamma_h(b)$, and the fact that $\mathbf{Q} = \iota \Omega \iota'$ and $\iota' \iota = I_n$, we have:

$$
\begin{aligned}
E v_{t+h}^2 = \sum_{k=0}^{h-1} &\left\{ \left(\frac{1-b}{b} \beta'(b) - \alpha'_{22} \iota' \right) \left(\sum_{j=0}^{k-1} F^j \right) + \alpha'(b)(I_n - \Phi(b))^{-1} \iota' \right\} \mathbf{Q} \\
&\left\{ \left(\sum_{j=0}^{k-1} F'^j \right) \left(\frac{1-b}{b} \beta(b) - \iota \alpha_{22} \right) + \iota (I_n - \Phi'(b))^{-1} \alpha(b) \right\}.
\end{aligned}
$$

Define:

$$
B_{k,b} \equiv \left\{ \left(\sum_{j=0}^{k-1} F'^j \right) \left(\frac{1-b}{b} \beta(b) - \iota \alpha_{22} \right) + \iota (I_n - \Phi'(b))^{-1} \alpha(b) \right\}. \tag{A.3.5}
$$

Then, we have:

$$
E(s_{t+h} - s_t)^2 = A'_{h,b} E(\Delta \mathbf{Y}_t \Delta \mathbf{Y}'_t) A_{h,b} + \sum_{k=0}^{h-1} B'_{k,b} \mathbf{Q} B_{k,b},
$$

and

$$
\frac{1-b}{b} B_{k,b} \overset{b \to 1}{\to} \iota (I_n - \Phi'(1))^{-1} \alpha_{21}
$$

$$
= O(1).
$$

Putting the numerator and denominator together, we have:

$$
\begin{aligned}
\frac{1}{R^2(h)} - 1 &= \frac{E(s_{t+h} - s_t)^2}{\{E(s_{t+h} - s_t)z_t\}^2 / E z_t^2} - 1 \\
&= \frac{\sum_{k=0}^{h-1} B_{k,b}' \mathbf{Q} B_{k,b}}{A_{h,b}' E(\Delta \mathbf{Y}_t \Delta \mathbf{Y}_t') A_{h,b}}
\end{aligned}
$$

∎

Proof of 2(a).

The proof of Proposition 2(a) is similar to the proof of Proposition 1(a), but there are some critical differences. First, notice that using the definition of $\eta(b)$, we have:

$$
z_t = \frac{-b^2}{1 - b} f_{2t} + b u_{2t} + \eta'(b) \sum_{j=1}^{\infty} b^j E_t \Delta \mathbf{X}_{t+j}. \tag{A.3.6}
$$

Again using the fact that $E_t \Delta \mathbf{X}_{t+j} = \iota' F^j \Delta \mathbf{Y}_t$, we have:

$$
\begin{aligned}
z_t &= b \left(\alpha_{22}' - \frac{b}{1-b} \alpha_{21}' \right) \iota' \Delta \mathbf{Y}_t + \eta'(b) \iota' \left(\sum_{j=1}^{\infty} b^j F^j \right) \Delta \mathbf{Y}_t \\
&= \omega'(b) \Delta \mathbf{Y}_t.
\end{aligned}
$$

For $b < 1$, z_t is stationary because $\Delta \mathbf{Y}_t$ is stationary.

Using the same mechanics as the derivation for the one-step-ahead regression in Proposition 1(a), it can be shown that:

$$
\Delta s_{t+1} = \frac{1-b}{b} z_t - u_{2t} + \eta'(b)(I_n - \Phi(b))^{-1} \varepsilon_{t+1}.
$$

Notice that the key difference between propositions 1(a) and 2(a) lies in the difference between $\alpha(b), \beta(b)$ and $\eta(b), \omega(b)$. By iterating the one-step-ahead regression ahead h times, we have:

$$
s_{t+h} - s_t = \frac{1-b}{b} \sum_{k=0}^{h-1} z_{t+k} - \sum_{k=0}^{h-1} u_{2t+k} + \eta'(b) \Gamma_h'(b) \varepsilon_{t,h}. \tag{A.3.7}
$$

Again, following the same mechanics as in Proposition 1(a), we have:

$$z_{t+k} = \omega^{'}(b)F^{k}(\omega(b)\omega^{'}(b))^{-1}\omega(b)z_{t} + \omega^{'}(b)\xi_{t+k}.$$

Plugging this into (A.3.7), we have:

$$
\begin{aligned}
s_{t+h} - s_t &= \frac{1-b}{b}\omega^{'}(b)\left(\sum_{k=0}^{h-1}F^{k}\right)\left(\omega(b)\omega^{'}(b)\right)^{-1}\omega^{'}(b)z_t + \frac{1-b}{b}\omega^{'}(b)\sum_{k=0}^{h-1}\xi_{t+k} - \alpha_{22}^{'}\iota^{'}\sum_{k=0}^{h-1}\left(F^{k}\Delta\mathbf{Y}_t + \xi_{t+k}\right) + \eta^{'}(b)\Gamma_{h}^{'}(b)\varepsilon_{t,h}\\
&\equiv \theta_h z_t - \alpha_{22}^{'}\iota^{'}\left(\sum_{k=0}^{h-1}F^{k}\right)\Delta\mathbf{Y}_t + \nu_{t+h},
\end{aligned}
$$

where

$$
\begin{aligned}
\theta_h &\equiv \frac{1-b}{b}\omega^{'}(b)\left(\sum_{k=0}^{h-1}F^{k}\right)\left(\omega(b)\omega^{'}(b)\right)^{-1}\omega^{'}(b)\\
\nu_{t+h} &\equiv \left\{\left(\frac{1-b}{b}\omega^{'}(b) - \alpha_{22}^{'}\iota^{'}\right)F_{h}^{'} + \eta^{'}(b)\Gamma_{h}^{'}(b)\right\}\varepsilon_{t,h}.
\end{aligned}
$$

∎

Proof of 2(b).

Using the same derivation as in Proposition 1(b), and the fact that $z_t = \omega^{'}(b)\Delta\mathbf{Y}_t$, we have:

$$s_{t+h} - s_t = \left\{\frac{1-b}{b}\omega^{'}(b) - \alpha_{22}^{'}\iota^{'}\right\}\left(\sum_{k=0}^{h-1}F^{k}\right)\Delta\mathbf{Y}_t + \nu_{t+h}.$$

As a result,

$$\{E(s_{t+h} - s_t)z_t\}^2 = Ez_t^2\left\{\frac{1-b}{b}\omega^{'}(b) - \alpha_{22}^{'}\iota^{'}\right\}\left(\sum_{k=0}^{h-1}F^{k}\right)E(\Delta\mathbf{Y}_t\Delta\mathbf{Y}_t^{'})\left(\sum_{k=0}^{h-1}F^{'k}\right)\left\{\frac{1-b}{b}\omega(b) - \iota\alpha_{22}\right\},$$

and

$$E(s_{t+h} - s_t)^2 = \left\{\frac{1-b}{b}\omega^{'}(b) - \alpha_{22}^{'}\iota^{'}\right\}\left(\sum_{k=0}^{h-1} F^k\right) E(\Delta \mathbf{Y}_t \Delta \mathbf{Y}_t^{'})\left(\sum_{k=0}^{h-1} F^{'k}\right)\left\{\frac{1-b}{b}\omega(b) - \iota\alpha_{22}\right\} + E\nu_{t+h}^2$$

$$E\nu_{t+h}^2 = \sum_{k=0}^{h-1}\left\{\left(\frac{1-b}{b}\omega^{'}(b) - \alpha_{22}^{'}\iota^{'}\right)\left(\sum_{j=0}^{k-1} F^j\right) + \eta^{'}(b)(I_n - \Phi(b))^{-1}\iota^{'}\right\}\mathbf{Q}$$

$$\left\{\left(\sum_{j=0}^{k-1} F^{'j}\right)\left(\frac{1-b}{b}\omega(b) - \iota\alpha_{22}\right) + \iota(I_n - \Phi^{'}(b))^{-1}\eta(b)\right\}$$

If $b = 1$, $\frac{1-b}{b}\omega(b)|_{b=1} = -\iota\alpha_{21}$, and $\eta(b)|_{b=1} = \eta(1) = \alpha_1 + \alpha_{21} + \alpha_{22}$. Therefore, if we define:

$$C_h \equiv \left(\sum_{k=0}^{h-1} F^{'k}\right)\iota(\alpha_{21} + \alpha_{22}) \tag{A.3.8}$$

$$D_k \equiv \left(\sum_{j=0}^{k-1} F^{'j}\right)\iota(\alpha_{21} + \alpha_{22}) + \iota(I_n - \Phi^{'}(b))^{-1}\eta(1), \tag{A.3.9}$$

for $b = 1$, we have:

$$\frac{1}{R^2(h)} - 1 = \frac{E(s_{t+h} - s_t)^2}{\{E(s_{t+h} - s_t)z_t\}^2/Ez_t^2} - 1$$

$$= \frac{\sum_{k=0}^{h-1} D_k^{'}\mathbf{Q}D_k}{C_h^{'}E(\Delta \mathbf{Y}_t \Delta \mathbf{Y}_t^{'})C_h}$$

as required. Inspecting the denominator of this expression, we see: that

$$C_h^{'}E(\Delta \mathbf{Y}_t \Delta \mathbf{Y}_t^{'})C_h = E(C_h^{'}\Delta \mathbf{Y}_t)^2$$

$$= Var\left((\alpha_{21}^{'} + \alpha_{22}^{'})\iota^{'}\left(\sum_{k=0}^{h-1} F^k\right)\Delta \mathbf{Y}_t\right)$$

$$= Var\left((\alpha_{21}^{'} + \alpha_{22}^{'})\iota^{'}\sum_{k=0}^{h-1} E_t\Delta \mathbf{Y}_{t+k}\right)$$

$$= Var\left(\sum_{k=0}^{h-1} E_t(f_{2,t+k} + u_{2,t+k})\right)$$

$$= Var\left(\sum_{k=0}^{h-1} \mathbf{P}_{t+k}\right),$$

while the numerator can be expressed as:

$$
\begin{aligned}
\sum_{k=0}^{h-1} D_k^{'} \mathbf{Q} D_k &= Var(\nu_{t+h}) \\
&= Var\left(\left(-(\alpha_{21}^{'} + \alpha_{22}^{'})\iota^{'} F_h^{'} + \eta^{'}(1)\Gamma^{'}(1)\right)\varepsilon_{t,h}\right) \\
&= Var\left(-(\alpha_{21}^{'} + \alpha_{22}^{'})\iota^{'}\sum_{k=0}^{h-1}\xi_{t+k} + \eta^{'}(1)(I_n - \Phi(b))^{-1}\sum_{k=1}^{h}\varepsilon_{t+k}\right).
\end{aligned}
$$

To complete the proof, note that:

$$
\begin{aligned}
(\alpha_{21}^{'} + \alpha_{22}^{'})\iota^{'}\xi_{t+k} &= (\alpha_{21}^{'} + \alpha_{22}^{'})\iota^{'}(\Delta\mathbf{Y}_{t+k} - E_t\Delta\mathbf{Y}_{t+k}) \\
&= (\alpha_{21}^{'} + \alpha_{22}^{'})(\Delta\mathbf{X}_{t+k} - E_t\Delta\mathbf{X}_{t+k}) \\
&\equiv \Xi_{1,t+k}.
\end{aligned}
$$

By equation (A9) of Engel and West (2005) and the fact that $\eta(1) = \alpha_1 + \alpha_{21} + \alpha_{22}$, we have:

$$
\begin{aligned}
\eta^{'}(1)(I_n - \Phi(b))^{-1}\varepsilon_{t+k} &= \eta^{'}(1)\left(E_{t+k}\sum_{j=0}^{\infty}\Delta\mathbf{X}_{t+k+j} - E_{t+k-1}\sum_{j=0}^{\infty}\Delta\mathbf{X}_{t+k+j}\right) \\
&\equiv \Xi_{2,t+k}.
\end{aligned}
$$

Therefore, we have $\sum_{k=0}^{h-1} D_k^{'}\mathbf{Q}D_k = Var\left(\sum_{k=0}^{h-1}(\Xi_{1,t+k} - \Xi_{2,t+k+1})\right)$. ∎

A.4   Details on Simulations

A.4.1   General Setup

As discussed in Section 4, the data vector $\Delta\mathbf{X}_t$ is set up according to different models.[16] For a given lag order p, a VAR(p) process on $\Delta\mathbf{X}_t$ can be estimated. The coefficients estimates, $\widehat{\Phi}(L)$ and error variance-covariance matrix $\widehat{\Omega}$ are obtained. This process is detailed in Appendix A.4.3.

The coefficient vectors α_1, α_{21} and α_{22} may contain coefficient estimates. For instance, in the Taylor Rule model, coefficients on cross-country output gap and inflation differences need to be estimated from the Taylor Rule relationship.

Let $Sim = 1,000,000$ and the superscript $*$ denote simulated variables. $n \times 1$ vectors of Gaussian errors

---

[16]We subtract the sample means from the data so that $\Delta\mathbf{X}_t$ is mean-zero.

ε_t^* are drawn for $t = 1, ..., Sim$. With that, pseudo-true fundamentals $\Delta \mathbf{X}_t^*$ are generated recursively using the equation:

$$(I_n - \widehat{\Phi}(L))\Delta \mathbf{X}_t^* = \widehat{\Omega}^{1/2}\varepsilon_t^*$$

for $t = 1, ..., Sim$ and $\widehat{\Omega}^{1/2}$ is obtained by Cholesky decomposition. To begin the recursion, we assume $\Delta \mathbf{X}_0^*, ..., \Delta \mathbf{X}_{-p}^* = 0$.

Next we generate z_t^*. Because all our simulation exercises falls under Case 2, using Proposition 2, we have:

$$
\begin{aligned}
z_t^* &= \widehat{\omega}'(b)\Delta \mathbf{Y}_t^* \\
\widehat{\omega}(b) &= \left\{ b\left(\alpha_{22}' - \frac{1-b}{b}\alpha_{21}'\right)\iota' + (\alpha_1' + b\alpha_{21}' + b\alpha_{22}')\iota'\widehat{F}(b) \right\}\Delta \mathbf{Y}_t^*,
\end{aligned}
$$

where $\widehat{F}(b)$ is an estimate of $F(b)$ by replacing $\Phi's$ with $\widehat{\Phi}'s$ and $\Delta \mathbf{Y}_t^* = [\Delta \mathbf{X}_t^{*'}, ..., \Delta \mathbf{X}_{t-p+1}^{*'}]'$.

Finally, following the one-horizon regression in the proof of Proposition 1, Δs_{t+1}^* for $t = 0, ..., Sim - 1$ is generated using:

$$\Delta s_{t+1}^* = \frac{1-b}{b}z_t^* - \alpha_{22}'\Delta \mathbf{X}_t^* + (\alpha_1' + b\alpha_{21}' + b\alpha_{22}')(I_n - \widehat{\Phi}(b))^{-1}\widehat{\Omega}^{1/2}\varepsilon_{t+1}^*.$$

Assuming $s_0^* = 0$, $s_t^*, t = 1, ..., Sim$ can be calculated recurisvely. We discard the first 1,000 observations of all generated variables to avoid start-up effects.

Then, using $Sim - 1000$ observations we run the long-horizon regressions (i.e. regressing $s_{t+h}^* - s_t^*$ on z_t^*) to obtain $R^2(h)$. Since Sim is already large, increases in Sim did not affect the results qualitatively.

A.4.2   Data Description

We collected quarterly data (1989Q4 to 2007Q2)[17] for 8 countries: US, Canada, Denmark, Germany, Japan, Norway, Switzerland, and UK. From the Consensus Forecast, we obtain the Uncovered Interest-rate Parity (UIP) risk premium between the US and other countries. Most remaining data are from the G10+ dataset provided by Haver Analytics. Our dataset includes:

- Uncovered Interest-rate Parity (UIP) risk premium

- Money supply (Seasonally adjusted M2 for all countries other than Japan, in which M2+CDs is used.

---

[17]The data for Denmark is restricted to 1990Q1 and 2007Q2 due to missing GDP data.

Norway's M2 data are from International Financial Statistics (IFS) and are not seasonally adjusted. We seasonally adjusted the data using Eviews.)

- GDP (Chained real GDP for all countries. Data for Germany and Japan are calculated from nominal GDP and the GDP deflator, which we obtained from IFS.)

- CPI (In UK, we use HICP. The data for Canada and Japan are from OECD.)

- Short term interest rate (For Canada and US: 3-month treasury bill rate; for Denmark, Germany, Norway, Switzerland, and UK: 3-month interbank offer rate; for Japan: 3-month Certificate of Deposit (Gensaki) Rate. Japan's data are obtained from OECD).

- Short term interest rate targeted by the central bank (for Canada: Overnight Money Market Financing Rate (Effective); for Denmark: National Bank Discount Rate; for Germany: Base Rate; for Japan: Overnight Call Rate (Uncollateralized); for Norway: 3-month Interbank Offer Rate; for Switzerland: 3-month Interbank Offer Rate; for United Kingdom: Base Rate; and for US: Federal Funds (effective) Rate.)

- Exchange rates (All exchange rates are in the form of units of foreign currency per US dollar.)

A.4.3 Estimating VAR(p)

A.4.3.1 Monetary Model

Fundamentals in the Monetary model include (log) money supply (m_t, m_t^*), (log) output (y_t, y_t^*), the (log) real exchange rate (q_t), a money demand shock (v_t, v_t^*), and risk premium (ρ_t).[18] We use the US as the base country and denote its variables with an asterisk. For each country, real money supply ($m_t - p_t$) is regrssed on total income (y_t) and the short-term interest rate (i_t). We cannot reject the unit root hypothesis for m_t, p_t, y_t, and i_t and the null hypothesis that these variables are not cointegrated at a 5% significance level for most countries. Therefore, we proceed to take the first difference for all variables in our regression. The regression residuals are recovered as Δv_t and Δv_t^*. The real exchange rate is calculated from $q_t = s_t + p_t^* - p_t$, where s_t is the (log) nominal exchange rate, which is in the form of units of national currency per US dollar.

Under the Augmented Dickey-Fuller Unit Root test, we generally cannot reject the null hypothesis at a 95% confidence level in all countries for m_t, m_t^*, y_t, y_t^*, q_t, v_t, and v_t^*. Unit roots are strongly rejected for ρ_t in all

---

[18]If we assume PPP holds, the real exchange rate is zero.

countries.[19] Following the definition in the paper, we have:

$$\Delta \mathbf{X_t} \equiv \left[\begin{array}{cccccccc} \Delta m_t & \Delta m_t^* & \Delta y_t & \Delta y_t^* & \Delta q_t & \Delta v_t & \Delta v_t^* & \rho_t \end{array} \right]. \tag{A.4.10}$$

We estimate VAR(p) coefficients for ΔX_t and the covariance matrix of residuals with $p = 1, 2, 3, 4$. The estimated coefficients and covariance matrix are used for simulations.

In an alternative setup, we regress $\Delta[(m_t - p_t) - (m_t^* - p_t^*)]$ on $\Delta(y_t - y_t^*)$ and $(i_t - i_t^*)$. The residuals are recovered as $\Delta(v_t - v_t^*)$. The matrix $\Delta \mathbf{X_t}$ contains the following variables:

$$\Delta \mathbf{X_t} = \left[\begin{array}{cccc} \Delta(m_t - m_t^*) & \Delta(y_t - y_t^*) & \Delta q_t & \Delta(v_t - v_t^*) & \rho_t \end{array} \right]. \tag{A.4.11}$$

A.4.3.2 Taylor Rule Model

The economic variables in the Taylor Rule model include: the (log) aggregate price (p_t), the output gap (y_t^g), CPI inflation (π_t), the monetary policy shock in the Taylor Rule (v_t), and the risk premium in the UIP condition (ρ_t). Corresponding variables in the foreign country are denoted with asterisks. As discussed in the paper, CPI is I(1) and all other variables are I(0). As a result, we define:

$$\boldsymbol{\Delta} \mathbf{X_t} = \begin{bmatrix} \Delta p_t & \Delta p_t^* & y_t^g & y_t^{g*} & \pi_t & \pi_t^* & v_t & v_t^* & \rho_t \end{bmatrix}. \tag{A.4.12}$$

Note that Δp_t (Δp_t^*) and π_t (π_t^*) are collinear. We drop Δp_t and Δp_t^* from $\boldsymbol{\Delta} \mathbf{X_t}$ when estimating VAR(p) coefficients and variance matrixes.

We also tried an alternative setup in which the matrix $\boldsymbol{\Delta} \mathbf{X_t}$ is defined as:

$$\boldsymbol{\Delta} \mathbf{X_t} = \begin{bmatrix} y_t^g - y_t^{g*} & \pi_t - \pi_t^* & v_t - v_t^* & \rho_t \end{bmatrix}. \tag{A.4.13}$$

Note that we drop $\Delta(p_t - p_t^*)$ from $\boldsymbol{\Delta} \mathbf{X_t}$ because it is collinear with $\pi_t - \pi_t^*$.

---

[19]This result is true whether or not we include a time trend in our test, and for various lags (from 4 to 8) in a VAR(p).