

ROBUSTNESS AND EXCHANGE RATE VOLATILITY

EDOUARD DJEUTEM AND KENNETH KASA

ABSTRACT. This paper studies exchange rate volatility within the context of the monetary model of exchange rates. We assume agents regard this model as merely a benchmark, or reference model, and attempt to construct forecasts that are robust to model misspecification. We show that revisions of robust forecasts are more volatile than revisions of nonrobust forecasts, and that empirically plausible concerns for model misspecification can easily explain observed exchange rate volatility.

JEL Classification Numbers: F31, D81

1. INTRODUCTION

Exchange rate volatility remains a mystery. Over the years, many explanations have been offered - bubbles, sunspots, ‘unobserved fundamentals’, noise traders, etcetera. Our paper offers a new explanation. Our explanation is based on a disciplined retreat from the Rational Expectations Hypothesis. The Rational Expectations Hypothesis involves two assumptions: (1) Agents know the correct model of the economy (at least up to a small handful of unknown parameters, which can be learned about using Bayes rule), and (2) Given their knowledge of the model, agents make statistically optimal forecasts. In this paper, we try to retain the idea that agents process information efficiently, while at the same time relaxing what we view as the more controversial assumption, namely, that agents know the correct model up to a finite dimensional parameterization.

Of course, if agents don’t know the model, and do not have conventional finite-dimensional priors about it, the obvious question becomes - How are they supposed to forecast the future? Our answer is to suppose that agents possess a simple benchmark model of the economy, containing a few key macroeconomic variables. We further suppose that agents are aware of their own ignorance, and respond to it strategically by constructing forecasts from the benchmark model that are robust to a wide spectrum of potential misspecifications. We show that revisions of robust forecasts are quite sensitive to new information, and in the case of exchange rates, can easily account for observed exchange rate volatility.

Our paper is closely related to prior work by Hansen and Sargent (2008), Kasa (2001), and Lewis and Whiteman (2008). Hansen and Sargent have pioneered the application of robust control methods in economics. This literature formalizes the idea of a robust policy or forecast by viewing agents as solving dynamic zero sum games, in which a so-called ‘evil agent’ attempts to subvert the control or forecasting efforts of the decisionmaker. Hansen and Sargent show that concerns for robustness and model misspecification shed light on a wide variety of asset market phenomena, although they do not focus on exchange rate volatility. Kasa (2001) used frequency domain methods to derive a robust version of the

well known Hansen and Sargent (1980) prediction formula. This formula is a key input to all present value asset pricing models. Lewis and Whiteman (2008) use this formula to study stock market volatility. They show that concerns for model misspecification can explain observed violations of Shiller’s variance bound. They also apply a version of Hansen and Sargent’s detection error probabilities to gauge the empirical plausibility of the agent’s fear of model misspecification. Since robust forecasts are the outcome of a minmax control problem, one needs to make sure that agents are not being excessively pessimistic, by hedging against models that could have been easily rejected on the basis of observed historical time series. Lewis and Whiteman’s results suggest that explaining stock market volatility solely on the basis of a concern for robustness requires an excessive degree of pessimism on the part of market participants. Interestingly, when we modify their detection error calculations slightly, we find that robust forecasts *can* explain observed exchange rate volatility.¹

Since there are already many explanations of exchange rate volatility, a fair question at this point is - Why do we need another one? We claim that our approach enjoys several advantages compared to existing explanations. Although bubbles and sunspots can obviously generate a lot of volatility, these models require an extreme degree of expectation coordination. So far, no one has provided a convincing story for how bubbles or sunspots emerge in the first place. Our approach requires a more modest degree of coordination. Agents must merely agree on a simple benchmark model, and be aware of the fact that this model may be misspecified.² It is also clear that noise traders can generate a lot of volatility. However, as with bubbles and sunspots, there is not yet a convincing story for where these noise traders come from, and why they aren’t driven from the market. An attractive feature of our approach is that, if anything, agents in our model are *smarter* than usual, since they are aware of their own lack of knowledge about the economy.³

Our approach is perhaps most closely related to the ‘unobserved fundamentals’ arguments in West (1987), Engel and West (2004), and Engel, Mark, and West (2007). These papers all point out that volatility tests aren’t very informative unless one is confident

¹This is not the first paper to apply robust control methods to the foreign exchange market. Li and Tornell (2008) show that a particular type of structured uncertainty can explain the forward premium puzzle. However, they do not calculate detection error probabilities. Colacito and Croce (2011) develop a dynamic general equilibrium model with time-varying risk premia, and study its implications for exchange rate volatility. They adopt a ‘dual’ perspective, by focusing on a risk-sensitivity interpretation of robust control. However, they do not focus on Shiller bounds or detection error probabilities, as we do here.

²On bubbles, see *inter alia* Meese (1986) and Evans (1986). On sunspots, see Manuelli and Peck (1990) and King, Weber, and Wallace (1992). It should be noted that there *are* ways to motivate the emergence of sunspots via an adaptive learning process (Woodford (1990)), but then this just changes the question to how agents coordinated on a very particular learning rule. Another possibility is to avoid the coordination problem altogether, by postulating a collection of heterogeneous individuals who follow simple rules of thumb, and whose behavior evolves over time via a selection process (see, e.g., Arifovic (1996)).

³On the role of noise traders in fx markets, see Jeanne and Rose (2002). A more subtle way noise traders can generate volatility is to prevent prices from revealing other traders’ private information. This can produce a hierarchy of higher order beliefs about other traders’ expectations. Kasa, Walker, and Whiteman (2010) show that these higher order belief dynamics can explain observed violations of Shiller bounds in the US stock market.

that the full array of macroeconomic fundamentals are captured by a model.⁴ As a result, they argue that rather than test whether markets are ‘excessively volatile’, it is more informative to simply compute the fraction of observed exchange rate volatility that *can* be accounted for by innovations in observed fundamentals. Our perspective is similar, yet subtly different. In West, Engel-West, and Engel-Mark-West, fundamentals are only unobserved by the outside econometrician. Agents within the (Rational Expectations) model are presumed to observe them. In contrast, in our model it is the agents themselves who suspect there might be missing fundamentals, in the form of unobserved shocks that are correlated both over time and with the observed fundamentals. In fact, however, their benchmark model could be perfectly well specified. (In the words of Hansen and Sargent, their doubts are only ‘in their heads’). It is simply the prudent belief that they *could be* wrong that makes agents aggressively revise forecasts in response to new information.

In contrast to ‘unobserved fundamentals’ explanations, which are obviously untestable, there is a sense in which our model *is* testable. Since specification doubts are only ‘in their heads’, we can ask whether an empirically plausible degree of doubt can rationalize observed exchange rate volatility. That is, we only permit agents to worry about alternative models that could have plausibly generated the observed time series of exchange rates and fundamentals, where plausible is defined as an acceptable detection error probability, in close analogy to a significance level in a traditional hypothesis test. We find that given a sample size in the range of 100-150 quarterly observations, detection error probabilities in the range of 10-20% can explain observed exchange rate volatility.

The remainder of the paper is organized as follows. Section 2 briefly outlines the monetary model of exchange rates. We assume agents regard this model as merely a benchmark, and so construct forecasts that are robust to a diffuse array of unstructured alternatives. Section 3 briefly summarizes the data. We examine quarterly data from 1973:1-2011:3 on six US dollar bilateral exchange rates: the Australian dollar, the Canadian dollar, the Danish kroner, the Japanese yen, the Swiss franc, and the British pound. Section 4 contains the results of a battery of traditional excess volatility tests: Shiller’s original bound applied to linearly detrended data, the bounds of West (1988) and Campbell-Shiller (1987), which are robust to inside information and unit roots, and finally, and a couple of more recent tests proposed by Engel and West (2004) and Engel (2005). Although the results differ somewhat by test and currency, a fairly consistent picture of excess volatility emerges. Section 5 contains the results of our robust volatility bounds. We first apply Kasa’s (2001) robust Hansen-Sargent prediction formula, based on a so-called H^∞ approach to robustness, and show that in this case the model actually predicts exchange rates should be far *more* volatile than observed exchange rate volatility. We then follow Lewis and Whiteman (2008) and solve a frequency domain version of Hansen and Sargent’s evil agent game, which allows us to calibrate the degree of robustness to a given detection error probability. This is accomplished by assigning a penalty parameter to the evil agent’s actions. We find that observed exchange rate volatility can be explained if agents are hedging against models that have a 10-20% chance of being the true data-generating process. Section 6 contains a few concluding remarks.

⁴Remember, there is an important difference between unobserved fundamentals and unobserved information about observed fundamentals. The latter can easily be accommodated using the methods of Campbell and Shiller (1987) or West (1988).

2. THE MONETARY MODEL OF EXCHANGE RATES

The monetary model has been a workhorse model in open-economy macroeconomics. It is a linear, partial equilibrium model, which combines Purchasing Power Parity (PPP), Uncovered Interest Parity (UIP), and reduced-form money demand equations to derive a simple first-order expectational difference equation for the exchange rate. It presumes monetary policy and other fundamentals are exogenous. Of course, there is evidence against each of these underlying ingredients. An outside econometrician would have reasons to doubt the specification of the model. Unlike previous variance bounds tests using this model, we assume the agents within the model share these specification doubts.

Since the model is well known, we shall not go into details. (See, e.g., Mark (2001) for a detailed exposition). Combining PPP, UIP, and identical log-linear money demands yields the following exchange rate equation:

$$s_t = (1 - \beta)f_t + \beta E_t s_{t+1} \quad (2.1)$$

where s_t is the log of the spot exchange rate, defined as the price of foreign currency. The variable f_t represents the underlying macroeconomic fundamentals. In the monetary model, it is just

$$f_t = (m_t - m_t^*) - \lambda(y_t - y_t^*)$$

where m_t is the log of the money supply, y_t is the log of output, and asterisks denote foreign variables. In what follows, we assume $\lambda = 1$, where λ is the income elasticity of money demand. The key feature of equation (2.1) is that it views the exchange rate as an *asset price*. Its current value is a convex combination of current fundamentals, f_t , and expectations of next period's value. In traditional applications employing the Rational Expectations Hypothesis, E_t is defined to be the mathematical expectations operator. We relax this assumption. Perhaps not surprisingly, β turns out to be an important parameter, as it governs the weight placed on expectations in determining today's value. In the monetary model, this parameter is given by, $\beta = \alpha/(1 + \alpha)$, where α is the interest rate semi-elasticity of money demand. Following Engel and West (2005), we assume $.95 < \beta < .99$.

By imposing the no bubbles condition, $\lim_{j \rightarrow \infty} E_t \beta^j s_{t+j} = 0$, and iterating eq. (2.1) forward, we obtain the following present value model for the exchange rate

$$s_t = (1 - \beta) E_t \sum_{j=0}^{\infty} \beta^j f_{t+j} \quad (2.2)$$

which expresses the current exchange rate as the expected present discounted value of future fundamentals. If f_t is covariance stationary with Wold representation $f_t = A(L)\varepsilon_t$, then application of the Hansen-Sargent prediction formula yields the following closed-form expression for the exchange rate,

$$s_t = (1 - \beta) \left[\frac{LA(L) - \beta A(\beta)}{L - \beta} \right] \varepsilon_t \quad (2.3)$$

We shall have occasion to refer back to this in Section 5, when discussing robust forecasts. In practice, the assumption that f_t is covariance stationary is questionable. Instead, evidence suggests that for all six countries f_t contains a unit root. In this case, we need

to reformulate equation (2.2) in terms of stationary variables. Following Campbell and Shiller (1987), we can do this by defining the ‘spread’, $\phi_t = s_t - f_t$, and expressing it as a function of expected future *changes* in fundamentals, $\Delta f_{t+1} = f_{t+1} - f_t$ ⁵

$$s_t - f_t = E_t \sum_{j=1}^{\infty} \beta^j \Delta f_{t+j} \tag{2.4}$$

Applying the Hansen-Sargent formula to this expression yields the following closed form expression for the spread,

$$s_t - f_t = \beta \left[\frac{A(L) - A(\beta)}{L - \beta} \right] \varepsilon_t \tag{2.5}$$

where we now assume that $\Delta f_t = A(L)\varepsilon_t$.

3. THE DATA

We study six US dollar exchange rates: the Australian dollar, the Canadian dollar, the Danish kroner, the Japanese yen, the Swiss franc, and the British pound. The data are taken from the IFS, and are quarterly end-of-period observations, expressed as the dollar price of foreign currency. Money supply and income data are from the OECD Main Economic Indicators. Money supply data are seasonally adjusted M1 (except for Britain, which is M4). Income data are seasonally adjusted real GDP, expressed in national currency units. All data are expressed in natural logs.

Table 1 contains summary statistics for Δs_t and Δf_t .

TABLE 1
SUMMARY STATISTICS: QUARTERLY DATA (1973:1-2011:3)

	Aus	Can	Den	Jap	Swz	UK
$\underline{\Delta s}$						
Mean	-.0024	-.0002	.0007	.0080	.0083	-.0030
StDev	.0547	.0311	.0576	.0590	.0649	.0529
ρ_1	.067	.178	.067	.060	-.008	.151
$\underline{\Delta f}$						
Mean	-.0091	-.0071	-.0105	-.0035	-.0053	-.0136
StDev	.0307	.0197	.0277	.0273	.0455	.0236
ρ_1	.197	.367	.277	.394	.092	.461

Notes: (1) USA is reference currency, with s = price of fx in US dollars.
 (2) Both s and f are in natural logs, with $f = m - m^* - (y - y^*)$.
 (3) ρ_1 is the first-order autocorrelation coefficient.

There are three noteworthy features. First, it is apparent that exchange rates are close to random walks, while at the same time changes in fundamentals are predictable. This apparent contradiction was explained by Engel and West (2005). They show that if the discount factor is close to one, as it is with quarterly data, then exchange rates *should* be close to random walks. Second, it is apparent that the mean of Δs_t does not always match up well with the mean of Δf_t . This simply reflects the fact that some of these

⁵This suggests that $s_t - f_t$ should be stationary, i.e., s_t and f_t are cointegrated. As we show in the next section, the evidence here is more mixed.

currencies have experienced long-run *real* appreciations or depreciations. This constitutes prima facie evidence against the monetary model, independent of its volatility implications. Our empirical work accounts for these missing trends by either detrending the data, or by including trends in posited cointegrating relationships. This gives the model the best chance possible to explain exchange rate volatility. Third, and most importantly for the purposes of our paper, note that the standard deviations of Δs_t are around twice as large as the standard deviations of Δf_t . Given the mild persistence in Δf_t , this is a major problem when it comes to explaining observed exchange rate volatility.

As we discuss in more detail in the following section, variance bounds tests require assumptions about the nature of trends in the data. The original tests of Shiller (1981) presumed stationarity around a deterministic trend. Subsequent work showed that this causes a bias toward rejection if in fact the data contain unit roots, which then led to the development of tests that are robust to the presence of unit roots (Campbell and Shiller (1987), West (1988)). Table 2 contains the usual tests for unit roots and cointegration.

TABLE 2
UNIT ROOT AND COINTEGRATION TESTS

	Aus	Can	Den	Jap	Swz	UK
<u>Dickey-Fuller</u>						
s	-1.43	-1.16	-1.96	-2.08	-2.56	-2.70
f	-1.03	-2.09	-1.57	-1.24	-2.27	-1.43
$(s - f)$	-2.05	-1.91	-2.44	-2.37	-3.52	-3.25
<u>Engle-Granger</u>						
(s, f)	-1.44	-1.60	-2.49	-4.16	-3.62	-3.30
<u>Johansen (Trace)</u>						
(s, f)	4.73	11.18	11.33	20.12	23.51	19.21

Notes: (1) DF and EG regressions include constant, trend, and four lags.

(2) (5%, 10%) critical values for $DF = (-3.44, -3.14)$.

(3) (5%, 10%) critical values for $EG = (-3.84, -3.54)$.

(4) Johansen based on VAR(4) with deterministic trend. 5% critical value = 18.40.

Clearly, there is little evidence against the unit root null for both the exchange rate and fundamentals. This casts doubt on the applicability of the original Shiller bound. One can also see that the cointegration evidence is more mixed. There is some evidence in favor of cointegration for Japan, Switzerland, and the UK (especially when using the Johansen test), but little or no evidence for Australia, Canada, or Denmark. Later, when implementing tests based on the unit root specification, we simply assume the implied cointegration restrictions hold.

4. TRADITIONAL VOLATILITY TESTS

As motivation for our paper, this section briefly presents results from applying traditional volatility tests. We present them roughly in order of their historical development. Given our purposes, we do not dwell on the (many) statistical issues that arise when implementing these tests.⁶

⁶See Gilles and LeRoy (1991) for an excellent summary of these issues.

4.1. **Shiller (1981).** The attraction of Shiller’s original variance bound test is that it is based on such a simple and compelling logic. Shiller noted that since asset prices are the conditional expectation of the present value of future fundamentals, they should be *less* volatile than the ex post realized values of these fundamentals. More formally, if we define $s_t^* = (1 - \beta) \sum_{j=0}^{\infty} \beta^j f_{t+j}$ as the ex post realized path of future fundamentals, then the monetary model is equivalent to the statement that $s_t = E_t s_t^*$. Next, we can always decompose a random variable into the sum of its conditional mean and an orthogonal forecast error,

$$\begin{aligned} s_t^* &= E_t s_t^* + u_t \\ &= s_t + u_t \end{aligned}$$

Since by construction the two terms on the right-hand side are orthogonal, we can take the variance of both sides and get

$$\sigma_{s^*}^2 = \sigma_s^2 + \sigma_u^2 \quad \Rightarrow \quad \sigma_s^2 \leq \sigma_{s^*}^2$$

It’s as simple as that.⁷ Two practical issues arise when implementing the bound. First, prices and fundamentals clearly trend up over time. As a result, neither possesses a well defined variance, so it is meaningless to apply the bounds test to the raw data. Shiller dealt with this by linearly detrending the data. To maintain comparability, we do the same, although later we present results that are robust to the presence of unit roots. Second, future fundamentals are obviously unobserved beyond the end of the sample. Strictly speaking then, we cannot compute $\sigma_{s^*}^2$, and therefore, Shiller’s bound is untestable. One can always argue that agents within any given finite sample are acting on the basis of some as yet unobserved event. Of course, this kind of explanation is the last refuge of a scoundrel, and moreover, we show that it is unnecessary. Shiller handled the finite sample problem by assuming that s_T^* , the end-of-sample forecast for the discounted present value of future fundamentals, was simply given by the sample average. Unfortunately, subsequent researchers were quick to point out that this produces a bias toward rejection. So in this case, we depart from Shiller by using the unbiased procedure recommended by Mankiw, Romer, and Shapiro (1985). This involves iterating on the backward recursion $s_t^* = (1 - \beta)f_t + \beta s_{t+1}^*$, with the boundary condition, $s_T^* = s_T$.

The first row of Table 3 reports results from applying Shiller’s bound. Evidently, rather than being less volatile than ex post fundamentals, exchange rates are actually between 3 and 32 times as volatile as fundamentals.⁸

⁷As emphasized by Kasa, Walker, and Whiteman (2010), matters aren’t quiet so simple in models featuring heterogeneous beliefs. The law of iterated expectations does not apply to the average beliefs operator. They show conventional applications of Shiller’s bound can easily generate false rejections when there are heterogeneous beliefs.

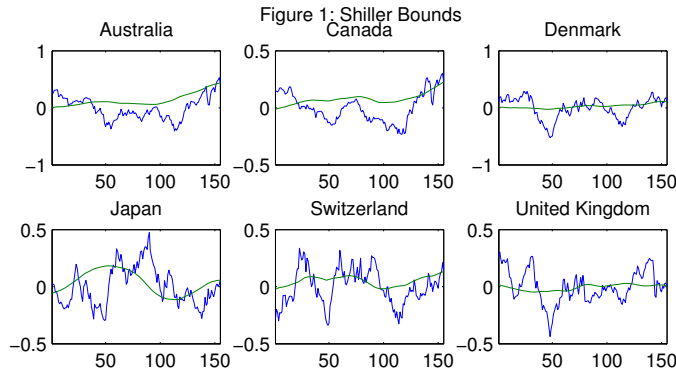
⁸By way of comparison, Shiller (1981) found that U.S. stock prices were 5 times as volatile as fundamentals. In the first application to exchange rates, Huang (1981) found that the pound and deutschemark were between 3 and 10 times too volatile. However, Diba (1987) pointed out that Huang’s results were tainted by a miscalibrated discount factor. With empirically plausible values of β , Huang’s tests showed no signs of excess volatility.

TABLE 3
TRADITIONAL VOLATILITY TESTS

	Aus	Can	Den	Jap	Swz	UK
<u>Shiller</u> $\text{var}(s)/\text{var}(s^*)$	3.13	7.85	25.7	2.93	14.2	32.5
<u>Campbell-Shiller</u> $\text{var}(\phi)/\text{var}(\phi^*)$	3.41	10.2	7.98	211.2	178.2	9.00
<u>West</u> $\text{var}(\varepsilon)/\text{var}(\hat{\varepsilon})$	2.36	1.81	2.33	1.34	1.73	1.77
<u>Engel-West</u> $\text{var}(\Delta s)/\text{var}(\Delta x_H)$	1.83	0.99	1.66	1.30	1.46	1.15
<u>Engel</u> $\text{var}(\Delta s)/\text{var}(\Delta \hat{s})$	5.98	10.8	3.35	7.42	20.7	14.6

Notes: (1) Shiller bound based on detrended data, assuming $s_T^* = s_T$ and $\beta = .98$.
(2) Campbell-Shiller bound based on a VAR(2) for $(\Delta f, \phi)$, assuming $\beta = .98$.
(3) Engel-West bound based on AR(2) for Δf , assuming $\beta = .98 \approx 1$.
(4) Engel bound based on AR(1) for detrended f_t , assuming $\beta = .98$.

Perhaps the most striking result that Shiller presented was a simple time series plot of s_t versus s_t^* . This, more than anything else, is what convinced many readers that stock prices are excessively volatile.⁹ Figure 1 reproduces the Shiller plot for each of our six currencies. These particular plots are based on the assumption $\beta = .98$, but similar plots are obtained for the empirically plausible range, $.95 \leq \beta \leq .99$.



⁹Not everyone was convinced. In particular, Kleidon (1986) warned that these plots could be quite misleading if the underlying data are nonstationary.

As with Shiller, these plots paint a clear picture of ‘excess volatility’. Unfortunately, there are enough statistical caveats and pitfalls associated with these results, that it is worthwhile to consider results from applying some of the more recently proposed bounds tests, starting with the influential work of Campbell and Shiller (1987).

4.2. Campbell and Shiller (1987). The results of Shiller (1981) are sensitive to the presence of unit roots in the data. As we saw in Table 2, there appear to be unit roots in both s_t and f_t (even after the removal of a deterministic trend). Campbell and Shiller (1987) devised a volatility test that is valid when the data are nonstationary. In addition, they devised a clever way of capturing potential information that market participants might have about future fundamentals that is unobserved by outside econometricians. To do this, one simply needs to include current and lagged values of the exchange rate when forecasting future fundamentals. Under the null, the current exchange rate is a sufficient statistic for the present value of future fundamentals. Forecasting future fundamentals with a VAR that includes the exchange rate converts Shiller’s variance bound inequality to a variance *equality*. The model-implied forecast of the present value of fundamentals should be identically equal to the actual exchange rate!

To handle unit roots, Campbell and Shiller (1987) presume that s_t and f_t are cointegrated, and then define the ‘spread’ variable $\phi_t = s_t - f_t$. The model implies that ϕ_t is the expected present value of future values of Δf_t (see eq. (2.4)). If we define the vector $(\Delta f_t, \phi_t)'$ we can then estimate the following VAR,

$$x_t = \Psi(L)x_{t-1} + \epsilon_t$$

By adding lags to the state, this can always be expressed as a VAR(1)

$$\hat{x}_t = \hat{\Psi}\hat{x}_{t-1} + \epsilon_t$$

where $\hat{\Psi}$ is a double companion matrix, and the first element of \hat{x}_t is Δf_t . The model-implied spread, ϕ_t^* , is given by the expected present discounted value of $\{\Delta f_{t+j}\}$, which can be expressed in terms of observables as follows,

$$\phi_t^* = e1' \beta \hat{\Psi} (I - \beta \hat{\Psi})^{-1} \hat{x}_t \quad (4.6)$$

where $e1$ is a selection vector that picks off the first element of \hat{x}_t . The model therefore implies $\text{var}(\phi_t) = \text{var}(\phi_t^*)$.

The second row of Table 3 reports values of $\text{var}(\phi_t)/\text{var}(\phi_t^*)$ based on a VAR(2) model (including a trend and intercept). We fix $\beta = .98$, although the results are similar for values in the range $.95 \leq \beta \leq .99$. Although Campbell and Shiller’s approach is quite different, the results are quite similar. All ratios substantially exceed unity. Although the Campbell-Shiller method points to relatively less volatility in Denmark and the UK, it actually suggests greater excess volatility for Japan and Switzerland.

4.3. West (1988). West (1988) proposed an alternative test that is also robust to the presence of unit roots. Rather than look directly at the volatility of observed prices and fundamentals, West’s test is based on comparison of two innovation variances. These innovations can be interpreted as one-period holding returns, and so are stationary even when the underlying price and fundamentals processes are nonstationary. Following West (1988), let I_t be the market information set at time- t , and let $H_t \subseteq I_t$ be some subset

that is observable by the outside econometrician. In practice, H_t is often assumed to just contain the history of past (observable) fundamentals. Next, define the following two present value forecasts:

$$x_{tH} = \sum_{j=0}^{\infty} \beta^j E(f_{t+j}|H_t) \quad x_{tI} = \sum_{j=0}^{\infty} \beta^j E(f_{t+j}|I_t)$$

West then derived the following variance bound,

$$E(x_{t+1,H} - E[x_{t+1,H}|H_t])^2 \geq E(x_{t+1,I} - E[x_{t+1,I}|I_t])^2$$

This says that if market participants have more information than H_t , their forecasts should have a smaller innovation variance. Intuitively, when forecasts are based on a coarser information set, there are more things that can produce a surprise. We apply this bound to forecasts of the spread, $\phi_t = s_t - f_t$, with H_t assumed to contain both the history of fundamentals *and* exchange rates. From the Campbell-Shiller logic, the inclusion of s_t in H_t converts West's variance bound inequality to a variance *equality*.

To derive the market's innovation, we can write the present value model recursively as follows

$$\phi_t = \beta E_t(\phi_{t+1} + \Delta f_{t+1})$$

Exploiting the decomposition, $\phi_{t+1} + \Delta f_{t+1} = E_t(\phi_{t+1} + \Delta f_{t+1}) + \varepsilon_{t+1}$, we can then derive the ex post observable expression for the market's innovation

$$\varepsilon_{t+1} = \phi_{t+1} + \Delta f_{t+1} - \frac{1}{\beta} \phi_t$$

Note that ε_{t+1} is just a one-period excess return. We can do the same for the predicted spread, $\hat{\phi}_t$, estimated from the same VAR(2) model used to compute the Campbell-Shiller test. This yields the fitted excess return,

$$\hat{\varepsilon}_{t+1} = \hat{\phi}_{t+1} + \Delta f_{t+1} - \frac{1}{\beta} \hat{\phi}_t$$

Under the null, $\text{var}(\varepsilon) = \text{var}(\hat{\varepsilon})$. The third row of Table 3 contains the results. Interestingly, the West test indicates a much less dramatic degree of excess volatility. Although the point estimates continue to exceed unity, most are in the range 1.5-2.0.

4.4. Engel and West (2004). Engel and West (2004) use results from Engel and West (2005) to derive a variance bound for the limiting case, $\beta \rightarrow 1$. We implement this bound assuming $\beta = .98$. Two things happen as $\beta \rightarrow 1$. First, from West (1988) we know

$$\text{var}(\varepsilon_{t,H}) = \frac{1 - \beta^2}{\beta^2} \text{var}(x_{t,H} - x_{t,I}) + \text{var}(\varepsilon_{t,I})$$

Since $\text{var}(x_{t,H} - x_{t,I})$ is bounded, it is clear that $\text{var}(\varepsilon_{t,H}) \approx \text{var}(\varepsilon_{t,I})$ as $\beta \rightarrow 1$. Second, from Engel and West (2005) we know that s_t converges to a random walk as $\beta \rightarrow 1$. Therefore, $\text{var}(\varepsilon_{t,I}) \approx \text{var}(\Delta s_t)$ as $\beta \rightarrow 1$. Combining, we conclude that $\text{var}(\varepsilon_{t,H}) \approx \text{var}(\Delta s_t)$ under the null as $\beta \rightarrow 1$. To estimate $\text{var}(\varepsilon_{t,H})$ we first estimate a univariate VAR(2) for Δf_t ,

$$\Delta f_t = \hat{\gamma}_0 + \hat{\gamma}_1 \Delta f_{t-1} + \hat{\gamma}_2 \Delta f_{t-2} + u_t$$

We can then estimate $\text{var}(\varepsilon_{t,H})$ as follows

$$\text{var}(\varepsilon_{t,H}) = (1 - \beta\hat{\gamma}_1 - \beta^2\hat{\gamma}_2)^{-2}\text{var}(u)$$

The fourth row of Table 3 reports the ratio $\text{var}(\Delta s_t)/\text{var}(\varepsilon_{t,H})$ for each of our six currencies. Overall, the results are quite similar to the results from the West test, the only significant difference being Canada, which has a point estimate (slightly) below unity. The results are also quite similar to those reported by Engel and West (2004), although as noted earlier, they interpret the results reciprocally, as the share of exchange rate volatility that can be accounted for by observed fundamentals.

4.5. Engel (2005). Engel (2005) derives a variance bound that is closely related to the West (1988) bound. Like the West bound, it is robust to the presence of unit roots. Let \hat{s}_t be the forecast of the present discounted value of future fundamentals based on a subset, H_t , of the market's information. Engel shows that the model, along with the Rational Expectations Hypothesis, implies the following inequality¹⁰

$$\text{var}(s_t - s_{t-1}) \leq \text{var}(\hat{s}_t - \hat{s}_{t-1})$$

For simplicity, we test this bound by assuming that f_t follows an AR(1) around a deterministic trend. That is, letting \tilde{f}_t denote the demeaned and detrended f_t process, we assume $\tilde{f}_t = \rho\tilde{f}_{t-1} + u_t$. In this case, we have

$$\hat{s}_t = \alpha_0 + \alpha_1 \cdot t + \left(\frac{1 - \beta}{1 - \rho\beta} \right) \tilde{f}_t \quad \Rightarrow \quad \Delta\hat{s}_t = \left(\frac{1 - \beta}{1 - \rho\beta} \right) \Delta\tilde{f}_t$$

The bottom row reports values of $\text{var}(\Delta s_t)/\text{var}(\Delta\hat{s}_t)$. Not surprisingly, given our previous results, they all exceed unity by a substantial margin, once again pointing to excess volatility in the foreign exchange market.

5. ROBUST VOLATILITY TESTS

The results reported in the previous section are based on tests that make different assumptions about information, trends, and the underlying data generating process. Despite this, a consistent picture emerges - exchange rates appear to exhibit 'excess volatility'. Of course, we haven't reported standard errors, so it is *possible* these results lack statistical significance (although we doubt it). However, we agree with Shiller (1989). Pinning your hopes on a lack of significance does not really provide much support for the model. It merely says there isn't enough evidence yet to reject it.

Although the previous tests differ along several dimensions, there is one assumption they all share, namely, the Rational Expectations Hypothesis (REH). As noted earlier, this is a joint hypothesis, based on two assumptions: (1) Agents have common knowledge of the correct model, and (2) Agents make statistically optimal forecasts. Of course, many previous researchers have interpreted Shiller's work as evidence against the Rational Expectations Hypothesis (including Shiller himself!). However, there is an important

¹⁰Engel (2005) also derives a second bound. As before, let s_t^* denote the ex post realized present value of future fundamentals. Engel shows that the model implies $\text{var}(s_t^* - s_{t-1}^*) \leq \text{var}(s_t - s_{t-1})$. Note, this is like the Shiller bound, but in first differences. However, note that the direction of the bound is reversed! Although we haven't formally checked it, given our previous results, we suspect this bound would be easily satisfied.

difference between our response and previous responses. Previous responses have focused on the second part of the REH, and so have studied the consequences of various kinds of ad hoc forecasting strategies. In contrast, we try to retain the idea that agents make statistically optimal forecasts, and instead relax the first part of the REH. We further assume that the sort of model uncertainty that agents confront cannot be captured by a conventional (finite-dimensional) Bayesian prior. As a result, we define an optimal forecast in terms of ‘robustness’.¹¹ Of course, it is possible to abandon *both* parts of the REH at the same time. This has been the approach of the adaptive learning literature. For example, Kim (2009) shows that (discounted) Least Squares learning about the parameters of an otherwise known fundamentals process can generate significant exchange rate volatility (although he does not focus on violation of variance bounds inequalities). Markiewicz (2011) assumes agents entertain a set of competing models, each of which is adaptively estimated. She shows that endogenous model switching can generate time-varying exchange rate volatility.

5.1. A Robust Hansen-Sargent Prediction Formula. In contrast to the previous section, where agents *knew* the monetary model was the true data-generating process, here we suppose agents entertain the possibility that they are wrong. Although the monetary model is still regarded as a useful benchmark, agents suspect the model could be misspecified in ways that are difficult to capture with a standard Bayesian prior. To operationalize the idea of a robust forecast, agents employ the device of a hypothetical ‘evil agent’, who picks the benchmark model’s disturbances so as to *maximize* the agent’s mean-squared forecast errors. Since the sequence of error-maximizing disturbances obviously depends on the agent’s forecasts, agents view themselves as being immersed in a dynamic zero-sum game. A robust forecast is a Nash equilibrium of this game.

We provide two solutions to this game. The first makes the agent’s present value forecasts maximally robust, in the sense that the (population) mean-squared error remains totally invariant to a wide spectrum of potential dynamic misspecifications. As stressed by Hansen and Sargent (2008), this may not be an empirically plausible solution, as it may reflect concerns about alternative models that could be easily rejected given the historically generated data. We want our agents to be prudent, not paranoid. It will also turn out to be the case that maximally robust forecasts deliver *too much* exchange rate volatility. Hence, we also construct a solution that limits the freedom of the evil agent to subvert the agent’s forecasts. The more we penalize the actions of evil agent, the closer we get to the conventional minimum mean-squared error forecast. Our empirical strategy is to first select a penalty parameter that replicates observed exchange rate volatility. We then calculate the detection error probability associated with this parameter value. We find that detection error probabilities in the range of 10-20% can explain observed exchange rate volatility.¹²

¹¹It is possible to provide axiomatic foundations that formalize the connection between robustness and optimality. See, e.g., Strzalecki (2011).

¹²Note, our focus on detection error probabilities differentiates our approach from the the ‘rare disasters’ literature. In that literature, agents are hedging against very low probability events. See, e.g., Farhi and Gabaix (2011).

To handle general forms of dynamic misspecification, it is convenient to solve the problem in the frequency domain. As a first step, let's return to the problem of forecasting the present value of fundamentals (eq. (2.2)) under the assumption that the data-generating process is known. In particular, suppose agents know that the Wold representation for fundamentals is $f_t = A(L)\varepsilon_t$. Transformed to the frequency domain our problem becomes (omitting the $(1 - \beta)$ constant for simplicity),

$$\min_{g(z) \in H^2} \frac{1}{2\pi i} \oint \left| \frac{A(z)}{1 - \beta z^{-1}} - g(z) \right| \frac{dz}{z} \quad (5.7)$$

where \oint denotes contour integration around the unit circle, and H^2 denotes the Hardy space of square-integrable analytic functions on the unit disk. Once we find $g(z)$, the time-domain solution for optimal forecast is $s_t = (1 - \beta)g(L)\varepsilon_t$. Restricting the z -transform of $g(L)$ to lie in H^2 guarantees the forecast is 'causal', i.e., based on a square-summable linear combination of current and past values of the underlying shocks, ε_t . The solution of the optimization problem in (5.7) is a classic result in Hilbert space theory (see, e.g., Young (1988), p. 188). It is given by,

$$g(z) = \left[\frac{A(z)}{1 - \beta z^{-1}} \right]_+ \quad (5.8)$$

where $[\cdot]_+$ denotes an 'annihilation operator', meaning 'ignore negative powers of z '. From (5.8) it is a short step to the Hansen-Sargent prediction formula. One simply subtracts off the principal part of the Laurent series expansion of $A(z)$ around the point β (see, e.g., Hansen and Sargent (1980, Appendix A)). This yields

$$g(z) = \frac{A(z) - \beta A(\beta)z^{-1}}{1 - \beta z^{-1}}$$

which is the well know Hansen-Sargent formula. (Note that $g(z) \in H^2$ by construction, since the pole at $z = \beta$ is cancelled by a zero at $z = \beta$).

Now, what if the agent *doesn't* know the true Wold representation of f_t ? In particular, suppose the z -transform of the actual process is $A^a(z) = A^n(z) + \Delta(z)$, where $A^n(z)$ is the agent's original benchmark (or nominal) model, and $\Delta(z)$ is an unknown (one-sided) perturbation function. Applying (5.8) in this case yields the following mean-squared forecast error:

$$\begin{aligned} \mathcal{L}^a &= \mathcal{L}^n + \|\Delta(z)\|_2^2 + \frac{2}{2\pi i} \oint \Delta(z) \left[\frac{A(z)}{1 - \beta z^{-1}} \right]_- \frac{dz}{z} \\ &= \mathcal{L}^n + \|\Delta(z)\|_2^2 + \frac{2}{2\pi i} \oint \Delta(z) \left(\frac{\beta A(\beta)}{z - \beta} \right) \frac{dz}{z} \end{aligned} \quad (5.9)$$

where $[\cdot]_-$ is an annihilation operator that retains only negative powers of z , and \mathcal{L}^a and \mathcal{L}^n denote actual and nominal mean-squared forecast errors. The point to notice is that \mathcal{L}^a could be much greater than \mathcal{L}^n , even when $\|\Delta(z)\|_2^2$ is small, depending on how $\Delta(z)$ interacts with β and $A(z)$. To see this, apply Cauchy's Residue Theorem to (5.9), which yields

$$\mathcal{L}^a = \mathcal{L}^n + \|\Delta(z)\|_2^2 + 2A(\beta)[\Delta(\beta) - \Delta(0)]$$

Notice that the last term is scaled by $A(\beta)$ which, though bounded, could be quite large. It turns out that the key to achieving greater robustness to model misspecification is to switch norms. Rather than evaluate forecast errors in the conventional H^2 sum-of-squares norm, we are now going to evaluate them in the H^∞ supremum norm. In the H^∞ -norm the optimal forecasting problem becomes¹³

$$\min_{g(z) \in H^\infty} \max_{|z|=1} \left| \frac{A(z)}{1 - \beta z^{-1}} - g(z) \right|^2 \quad (5.10)$$

where H^∞ denotes the Hardy space of essentially bounded analytic functions on the unit disk.¹⁴ Problem (5.10) is an example of a wide class of problems known as ‘Nehari’s Approximation Problem’, which involves minimizing the H^∞ distance between a two-sided L^∞ function and a one-sided H^∞ function. For this particular case, Kasa (2001) proves that the solution takes the following form

$$g(z) = \frac{zA(z) - \beta A(\beta)}{z - \beta} + \frac{\beta^2}{1 - \beta^2} A(\beta) \quad (5.11)$$

Notice that the first term is just the conventional Hansen-Sargent formula. The new element here comes from the second term. It shows that a concern for robustness causes the agent to revise his forecasts more aggressively in response to new information. Note that this vigilance is an increasing function of β . As $\beta \rightarrow 1$ the agent becomes more and more concerned about low frequency misspecifications.¹⁵ It is here that the agent is most exposed to the machinations of the evil agent, as even small misjudgments about the persistent component of fundamentals can inflict large losses on the agent. Of course, the agent pays a price for this vigilance, in the form of extra noise introduced at the high end of the frequency spectrum. An agent concerned about robustness is happy to pay this price.

To illustrate the role of robust forecasts in generating exchange rate volatility, we consider the spread, $\phi_t = s_t - f_t$, as a present value forecast of future values of Δf_t , and assume Δf_t is an AR(1), so that $A(L) = 1/(1 - \rho L)$ in the above formulas. Because the forecasts begin at $j = 1$ in this case, the Hansen-Sargent formula changes slightly (see eq. (2.5)). One can readily verify that the second term in the robust forecast remains unaltered.

Table 4 reports ratios of actual to predicted spread variances for both traditional and robust forecasts.

¹³Using the sup norm to attain robustness is a time-honored strategy in both statistics (Huber (1981)) and control theory (Zames (1981)).

¹⁴In principle, we should write ‘inf’ and ‘sup’ in (5.10), but in our case it turns out that the extrema are attained.

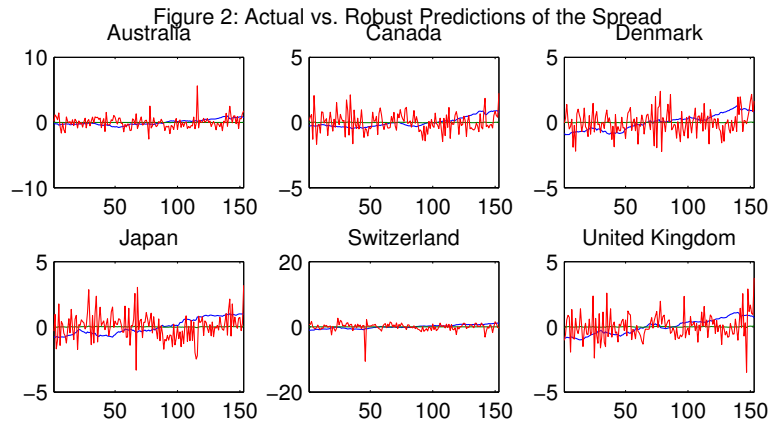
¹⁵This intuition is also the key to the Engel-West (2005) Theorem, which shows that exchange rates converge to random walks as $\beta \rightarrow 1$. This happens because agents become increasingly preoccupied by the permanent component of fundamentals as $\beta \rightarrow 1$.

TABLE 4
ROBUST VOLATILITY TESTS

	Aus	Can	Den	Jap	Swz	UK
<u>NonRobust AR(1) Forecast</u> $\text{var}(\phi)/\text{var}(\hat{\phi})$	3486.	1091.	3482.	1162.	15661.	740.1
<u>Robust AR(1) Forecasts</u> $\text{var}(\phi)/\text{var}(\hat{\phi})$.227	.291	.478	.358	.218	.363
AR coef	.197	.380	.281	.404	.092	.488

Notes: (1) Both bounds based on predictions of $\phi = s - f$ using an AR(1) model for Δf .
(2) When computing present values, it is assumed $\beta = .98$.

Here traditional forecasts look even worse than before, partly because we are failing to account for potential inside information, as we did in the Campbell-Shiller tests, and also partly because an AR(1) benchmark model is probably a bit too simple. Note, however, that this only makes our job harder when it comes to explaining excess volatility. The key results are contained in the middle row of Table 4, which reports ratios of actual to predicted spread variances with robust forecasts. Notice that all are well below one. If anything, robust forecasts generate too much volatility! This is vividly illustrated by Figure 2, which contains plots of the actual spread (blue line) against the nonrobust predicted spread and robust predicted spread (red line). The Shiller plot is inverted!¹⁶



Perhaps not surprisingly in light of eq. (5.11), the results here are rather sensitive to the value of β . For example, the above results assume $\beta = .98$. If instead we set $\beta = .92$, then all variance ratios exceed one and the model continues to generate excess volatility. This is not a serious problem in our view, since with quarterly data $\beta = .98$ is a more plausible value. A potentially more serious issue arises from the possibility that the

¹⁶We use eq. (5.11) to generate the time path of the robust spread. Expressed in terms of observables, this implies

$$\hat{\phi}_t^r = \frac{1}{1 - \beta\rho} \left[\left(\beta\rho + \frac{\beta^2}{1 - \beta^2} \right) \Delta f_t - \frac{\rho\beta^2}{1 - \beta^2} \Delta f_{t-1} \right]$$

agent is being excessively pessimistic here, and is worrying about alternative models that have little likelihood of being the true data-generating process. Fortunately, it is easy to avoid this problem by parameterizing the agent's concern for robustness. We do this by penalizing the actions of the evil agent.

5.2. The Evil Agent Game. The previous results incorporated robustness by evaluating forecast errors in the H^∞ -norm. This delivers the maximal degree of robustness, but may entail unduly pessimistic beliefs. Following Lewis and Whiteman (2008), we now go back to the H^2 -norm, and instead model robustness by penalizing the actions of the hypothetical evil agent. In particular, we assume the evil agent picks shocks subject to a quadratic penalty. As emphasized by Hansen and Sargent (2008), a quadratic penalty is convenient, since in Gaussian environments it can be related to entropy distortions and the Kullback-Leibler Information Criterion, which then opens the door to an interpretation of the agent's worst-case model in terms of detection error probabilities. This allows us to place empirically plausible bounds on the evil agent's actions.

Since our methods presume stationarity, we assume the agent has a benchmark (nominal) model, $\Delta f_t = A^n(L)\varepsilon_t$. In the conventional monetary model, this determines the spread, $\phi_t = s_t - f_t$, as in (2.4). At the same time, he realizes this model may be subject to unstructured perturbations of the form $\Delta(L)\varepsilon_t$, so that the actual model becomes, $\Delta f_t = A^a(L)\varepsilon_t = A^n(L)\varepsilon_t + \Delta(L)\varepsilon_t$. The Evil Agent Game involves the agent selecting a forecast function, $g(L)$, to minimize mean-squared forecast errors, while at the same time the evil agent picks the distortion function, $\Delta(L)$. Hence, both agents must solve a calculus of variations problem. These problems are related to each other, and so we must solve for a Nash equilibrium. Following Lewis and Whiteman (2008), we can express the problem in the frequency domain as follows

$$\min_{g(z)} \max_{A^a(z)} \frac{1}{2\pi i} \oint \left\{ \left| \frac{\beta z^{-1} A^a(z)}{1 - \beta z^{-1}} - g(z) \right|^2 - \theta \left| \frac{A^a(z) - A^n(z)}{1 - \beta z^{-1}} \right|^2 \right\} \frac{dz}{z} \quad (5.12)$$

where for convenience we assume the evil agent picks $A^a(z)$ rather than $\Delta(z)$. The key element here is the parameter θ . It penalizes the actions of the evil agent. By increasing θ we get closer to the conventional minimum mean-squared error forecast. Conversely, the smallest value of θ that is consistent with the concavity of the evil agent's problem delivers the maximally robust H^∞ solution.

The Wiener-Hopf first-order condition for the agent's problem is

$$g(z) - \frac{\beta z^{-1}}{1 - \beta z^{-1}} A^a(z) = \sum_{\infty}^{-1} \quad (5.13)$$

where \sum_{∞}^{-1} denotes an arbitrary function in negative powers of z . The evil agent's Wiener-Hopf equation can be written

$$\frac{(\beta^2 - \theta)A^a(z)}{1 - \beta z} - \beta z \frac{1 - \beta z^{-1}}{1 - \beta z} g(z) + \frac{\theta A^n(z)}{1 - \beta z} = \sum_{\infty}^{-1} \quad (5.14)$$

Applying the annihilation operator to both sides of eq. (5.13), we can then solve for the agent's policy in terms of the policy of the evil agent

$$g(z) = \beta \frac{A^a(z) - A^a(\beta)}{z - \beta}$$

Then, if we substitute this into the evil agent's first-order condition and apply the annihilation operator we get

$$\frac{(\beta^2 - \theta)A^a(z)}{1 - \beta z} - \beta^2 \frac{A^a(z) - A^a(\beta)}{1 - \beta z} + \frac{\theta A^n(z)}{1 - \beta z} = 0$$

which then implies

$$A^a(z) = A^n(z) + \frac{\beta^2}{\theta} A^a(\beta) \quad (5.15)$$

To determine $A^a(\beta)$ we can evaluate the evil agent's first-order condition (5.14) at $z = \beta$, which implies $A^a(\beta) = \frac{\theta A^n(\beta)}{\theta - \beta^2}$. Substituting this back into (5.15) yields

$$A^a(z) = A^n(z) + \frac{\beta^2}{\theta - \beta^2} A^n(\beta) \quad (5.16)$$

This gives the worst-case model associated with any given benchmark model. In game-theoretic terms, it represents the evil agent's reaction function. Notice that as $\theta \rightarrow \infty$ we recover the conventional minimum mean-squared error solution. Conversely, notice the close correspondence to the previous H^∞ solution when $\theta = 1$.

5.3. Detection Error Calibration. The idea here is that the agent believes $A^n(L)\varepsilon_t$ is a useful first-approximation to the actual data-generating process. However, he also recognizes that if he is wrong, he could suffer large losses. To minimize his exposure to these losses, he acts as if $\Delta f_t = A^a(L)\varepsilon_t$ is the data-generating process when formulating his present value forecasts. Reducing θ makes his forecasts more robust, but produces unnecessary noise if in fact the benchmark model is correct. To gauge whether observed exchange rate volatility might simply reflect a reasonable response to model uncertainty, we now ask the following question - Suppose we calibrate θ to match the observed volatility of the spread, $s_t - f_t$. (We know this is possible from the H^∞ results). Given this implied value of θ , how easy would it be for the agent to statistically distinguish the worst-case model, $A^a(L)$, from the benchmark monetary model, $A^n(L)$? If the two would be easy to distinguish, then our explanation doesn't carry much force. However, if the probability of a detection error is reasonably large, then we claim that model uncertainty provides a reasonable explanation of observed exchange rate volatility.

To begin, we need to take a stand on the benchmark model. For simplicity, we assume Δf_t follows an AR(1), so that $A^n(L) = 1/(1 - \rho L)$. (We de-mean the data). Given this, note that eq. (5.16) then implies the worst case model, $A^a(L)$, is an ARMA(1,1), with the same AR root. To facilitate notation, we write this model as follows,

$$\Delta f_t = \frac{1 + \kappa - \rho\kappa L}{1 - \rho L} \varepsilon_t$$

where $\kappa \equiv \frac{\beta^2}{\theta - \beta^2} \frac{1}{1 - \rho\beta}$. Notice that $\kappa \rightarrow 0$ as $\theta \rightarrow \infty$, and the two models become increasingly difficult to distinguish. As always, there are two kinds of inferential errors the agent

could make: (1) He could believe A^n is the true model, when in fact A^a is the true model, or (2) He could believe A^a is the true model, when in fact A^n is the true model. Denote the probability of the first error by $P(A^n|A^a)$, and the probability of the second error by $P(A^a|A^n)$. Following Hansen and Sargent (2008), we treat the two errors symmetrically, and define the detection error probability, \mathcal{E} , to be $\frac{1}{2}[P(A^n|A^a) + P(A^a|A^n)]$. From Taniguchi and Kakizawa (2000) (pgs. 500-503), we have the following frequency domain approximation to this detection error probability,

$$\mathcal{E} = \frac{1}{2} \left\{ \Phi \left[-\sqrt{T} \frac{I(A^n, A^a)}{V(A^n, A^a)} \right] + \Phi \left[-\sqrt{T} \frac{I(A^a, A^n)}{V(A^a, A^n)} \right] \right\} \quad (5.17)$$

where T is the sample size and Φ denotes the Gaussian cdf.¹⁷ Note that detection error probabilities decrease with T . The I functions in (5.17) are the KLIC ‘distances’ between the two models, given by

$$I(A^n, A^a) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left[-\log \frac{|A^n(\omega)|}{|A^a(\omega)|} + \frac{A^n(\omega)}{A^a(\omega)} - 1 \right] d\omega \quad (5.18)$$

$$I(A^a, A^n) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left[-\log \frac{|A^a(\omega)|}{|A^n(\omega)|} + \frac{A^a(\omega)}{A^n(\omega)} - 1 \right] d\omega \quad (5.19)$$

The V functions in (5.17) can be interpreted as standard errors. They are given by the square roots of the following variance functions

$$V^2(A^n, A^a) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left[A^n(\omega) \left(\frac{1}{A^n(\omega)} - \frac{1}{A^a(\omega)} \right) \right]^2 d\omega \quad (5.20)$$

$$V^2(A^a, A^n) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left[A^a(\omega) \left(\frac{1}{A^a(\omega)} - \frac{1}{A^n(\omega)} \right) \right]^2 d\omega \quad (5.21)$$

By substituting in the expressions for A^a and A^n and performing the integrations, we obtain the following expressions for I functions¹⁸

$$\begin{aligned} I(A^n, A^a) &= \frac{1}{2} \left[\log(1 + \kappa) + \frac{1}{(1 + \kappa)^2(1 - \psi^2)} - 1 \right] \\ I(A^a, A^n) &= \frac{1}{2} \left[-\log(1 + \kappa) + (1 + \kappa)^2(1 + \psi^2) - 1 \right] \end{aligned}$$

¹⁷The same formula applies even when the underlying data are non-Gaussian. All that changes is that higher-order cumulants must be added to the variance terms in (5.20)-(5.21). However, for simplicity, we suppose the agent knows the data are Gaussian. Intuitively, we suspect that relaxing this assumption would only strengthen our results.

¹⁸We employed the following trick when evaluating these integrals. First, if $A^n(L) = \frac{1}{1 - \rho L}$, we can then write $A^a = (1 + \kappa) \frac{1 - \psi L}{1 - \rho L}$, where κ is defined above and $\psi = \rho\kappa/(1 + \kappa)$. We then have (omitting inessential constants)

$$\oint -\log \frac{|A^n|}{|A^a|} \frac{dz}{z} = \oint \log(1 + \kappa) \frac{dz}{z} + \oint \log \left| \frac{1 - \psi z}{1 - \rho z} \right| \frac{dz}{z} - \oint \log \left| \frac{1}{1 - \rho z} \right| \frac{dz}{z} = \log(1 + \kappa)$$

where the last equality follows from the well known innovation formula, $\frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(\omega) d\omega = \log(\sigma^2)$, where $f(\omega)$ is the spectral density of a process, and σ^2 is its innovation variance. This cancels the last two terms since they have the same innovation variance. The same trick can be used to evaluate the other log integral.

where $\psi \equiv \rho\kappa/(1 + \kappa)$. As a simple reality check, note that $\theta \rightarrow \infty \Rightarrow \kappa \rightarrow 0 \Rightarrow I \rightarrow 0$, which we know must be the case. Now, doing the same thing for the V functions gives us

$$\begin{aligned}
 V^2(A^n, A^a) &= \frac{1}{2} \frac{1}{(1 + \kappa)^2(1 - \psi^2)} \left[(1 + \kappa)^4 - 2(1 + \kappa)^2 + \frac{1 + \psi^2}{(1 - \psi^2)^2} \right] \\
 V^2(A^a, A^n) &= \frac{1}{2} \left[\frac{1}{(1 + \kappa)^2} - 2(1 + \psi^2) + (1 + 4\psi^2 + \psi^4)(1 + \kappa)^2 \right]
 \end{aligned}$$

We use these formulas as follows. First, we estimate country-specific values of ρ . These are given in the bottom row of Table 4. Then we calibrate θ for each country to match the observed variance of its spread, $s_t - f_t$. The resulting values are reported in the first row of Table 5. Since the earlier H^∞ forecasts generated *too much* volatility, it is not surprising that the implied values of θ all exceed one (but not by much). Finally, as we have done throughout, we set $\beta = .98$. We can then calculate the κ and ψ parameters that appear in the above detection error formulas. The results are contained in the second and third rows of Table 5.

TABLE 5
EVIL AGENT GAME: CALIBRATED DETECTION ERRORS

	Aus	Can	Den	Jap	Swz	UK
θ	1.0442	1.0349	1.0178	1.0273	1.0454	1.0220
Det Error Prob($T = 150$)	.083	.109	.112	.119	.075	.131
Det Error Prob ($T = 100$)	.107	.131	.134	.140	.099	.151

Notes: (1) θ calibrated to match observed variance of $\phi = s - f$.
 (2) When computing present values, it is assumed $\beta = .98$.

We report detection error probabilities for two different sample sizes. The first assumes $T = 150$, which is (approximately) the total number of quarterly observations available in our post-Bretton Woods sample. In this case, the agent could distinguish the two models at approximately the 10% significance level.¹⁹ Although one could argue this entails excessive pessimism, keep in mind that the data are generated in real-time, and so agents did not have access to the full sample when the data were actually being generated. As an informal correction for this, we also report detection error probabilities assuming $T = 100$. Now the detection error probabilities lie between 10-15%.

6. CONCLUDING REMARKS

This paper has proposed a solution to the excess volatility puzzle in foreign exchange markets. Our solution is based on a disciplined retreat from the Rational Expectations Hypothesis. We abandon the assumption that agents know the correct model of the

¹⁹It is a little misleading to describe these results in the language of hypothesis testing. The agent is not conducting a traditional hypothesis test, since both models are treated symmetrically. It is more accurate to think of the agent as conducting a *pair* of hypothesis tests with two different nulls, or even better, as a Bayesian who is selecting between models with a uniform prior. The detection error probability is then expected loss under a 0-1 loss function.

economy, while retaining a revised notion of statistically optimal forecasts. We show that an empirically plausible concern for robustness can explain observed exchange rate volatility, even in a relatively simple environment like the constant discount rate/flexible-price monetary model.

Of course, there are many competing explanations already out there, so why is ours better? We think our approach represents a nice compromise between the two usual routes taken toward explaining exchange rate volatility. One obvious way to generate volatility is to assume the existence of bubbles, herds, or sunspots. Although these models retain the idea that agents make rational (self-fulfilling) forecasts, in our opinion they rely on an implausible degree of expectation coordination. Moreover, they are often not robust to minor changes in market structure or information. At the other end of the spectrum, many so-called ‘behavioral’ explanations have the virtue of not relying on strong coordination assumptions, but only resolve the puzzle by introducing rather drastic departures from conventional notions of optimality.

As noted at the outset, our paper is closely related to Lewis and Whiteman (2008). They argue that robustness can explain observed US stock market volatility. However, they also find that if detection errors are based only on the agent’s ability to discriminate between alternative models for the economy’s exogenous dividend process, then implausibly small detection error probabilities are required. If instead detection errors are based on the agent’s ability to discriminate between bivariate models of dividends *and* prices, then stock market volatility can be accounted for with reasonable detection errors. This is not at all surprising, since robustness delivers a substantially improved fit for prices. Interestingly, we find that even if detection errors are only based on the exogenous fundamentals process, exchange rate volatility can be accounted for with reasonable detection error probabilities. Still, one could argue that they are a bit on the low side, so it might be a useful extension to apply the bivariate Lewis-Whiteman approach to computing detection error probabilities. We conjecture that this would only strengthen our results.

REFERENCES

- ARIFOVIC, J. (1996): "The Behavior of the Exchange Rate in the Genetic Algorithm and Experimental Economies," *Journal of Political Economy*, 104, 510–41.
- CAMPBELL, J. Y., AND R. J. SHILLER (1987): "Cointegration and Tests of Present Value Models," *Journal of Political Economy*, 95, 1062–1088.
- COLACITO, R., AND M. M. CROCE (2011): "International Asset Pricing with Risk-Sensitive Agents," mimeo, University of North Carolina.
- DIBA, B. T. (1987): "A Critique of Variance Bounds Tests for Monetary Exchange Rate Models," *Journal of Money, Credit and Banking*, 19, 104–111.
- ENGEL, C. (2005): "Some New Variance Bounds for Asset Prices," *Journal of Money, Credit and Banking*, 37, 949–955.
- ENGEL, C., N. MARK, AND K. D. WEST (2007): "Exchange Rate Models Are Not As Bad As You Think," *NBER Macroeconomics Annual*, 22, 381–441.
- 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*, 94, 119–125.
- (2005): "Exchange Rates and Fundamentals," *Journal of Political Economy*, 113(3), 485–517.
- EVANS, G. W. (1986): "A Test for Speculative Bubbles in the Sterling-Dollar Exchange Rate: 1981–84," *American Economic Review*, 76, 621–36.
- FARHI, E., AND X. GABAIX (2011): "Rare Disasters and Exchange Rates," mimeo, Harvard University.
- GILLES, C., AND S. F. LEROY (1991): "Econometric Aspects of the Variance-Bounds Tests: A Survey," *The Review of Financial Studies*, 4, 753–791.
- HANSEN, L. P., AND T. J. SARGENT (1980): "Formulating and Estimating Dynamic Linear Rational Expectations Models," *Journal of Economic Dynamics and Control*, 2, 7–46.
- HANSEN, L. P., AND T. J. SARGENT (2008): *Robustness*. Princeton University Press, Princeton.
- HUANG, R. D. (1981): "The Monetary Approach to Exchange Rate in an Efficient Foreign Exchange Market: Tests Based on Volatility," *The Journal of Finance*, 36, 31–41.
- HUBER, P. J. (1981): *Robust Statistics*. Wiley, New York.
- JEANNE, O., AND A. ROSE (2002): "Noise Trading and Exchange Rate Regimes," *Quarterly Journal of Economics*, 117, 537–69.
- KASA, K. (2001): "A Robust Hansen-Sargent Prediction Formula," *Economics Letters*, 71, 43–48.
- KASA, K., T. B. WALKER, AND C. H. WHITEMAN (2010): "Heterogeneous Beliefs and Tests of Present Value Models," mimeo, University of Iowa.
- KIM, Y. S. (2009): "Exchange Rates and Fundamentals Under Adaptive Learning," *Journal of Economic Dynamics and Control*, 33, 843–63.
- KING, R., W. WEBER, AND N. WALLACE (1992): "Nonfundamental Uncertainty and Exchange Rates," *Journal of International Economics*, 32, 83–108.
- KLEIDON, A. W. (1986): "Variance Bounds Tests and Stock Price Valuation," *Journal of Political Economy*, 94, 953–1001.
- LEWIS, K. F., AND C. H. WHITEMAN (2008): "Robustifying Shiller: Do Stock Prices Move Enough to be Justified by Subsequent Changes in Dividends?," mimeo, University of Iowa.
- LI, M., AND A. TORNELL (2008): "Exchange Rates Under Robustness: An Account of the Forward Premium Puzzle," mimeo, UCLA.
- MANKIW, N. G., D. ROMER, AND M. D. SHAPIRO (1985): "An Unbiased Reexamination of Stock Market Volatility," *Journal of Finance*, 40, 677–87.
- MANUELLI, R. E., AND J. PECK (1990): "Exchange Rate Volatility in an Equilibrium Asset Pricing Model," *International Economic Review*, 31, 559–74.
- MARK, N. C. (2001): *International Macroeconomics and Finance: Theory and Econometric Methods*. Blackwell, Oxford.
- MARKIEWICZ, A. (2011): "Model Uncertainty and Exchange Rate Volatility," mimeo, Erasmus School of Economics.
- MEESE, R. A. (1986): "Testing for Bubbles in Exchange Markets: A Case of Sparkling Rates?," *Journal of Political Economy*, 94, 345–73.

- SHILLER, R. J. (1981): "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?," *American Economic Review*, 71(3), 421–436.
- (1989): *Market Volatility*. MIT Press.
- STRZALECKI, T. (2011): "Axiomatic Foundations of Multiplier Preferences," *Econometrica*, 79, 47–73.
- TANIGUCHI, M., AND Y. KAKIZAWA (2000): *Asymptotic Theory of Statistical Inference for Time Series*. Springer, New York.
- WEST, K. D. (1987): "A Standard Monetary Model and the Variability of the Deutschemark-Dollar Exchange Rate," *Journal of International Economics*, 23, 57–76.
- (1988): "Dividend Innovations and Stock Price Volatility," *Econometrica*, 56, 37–61.
- WOODFORD, M. (1990): "Learning to Believe in Sunspots," *Econometrica*, 58, 277–307.
- YOUNG, N. (1988): *An Introduction to Hilbert Space*. Cambridge University Press, Cambridge.
- ZAMES, G. (1981): "Feedback and Optimal Sensitivity: Model Reference Transformation, Multiplicative Seminorms, and Approximate Inverses," *IEEE Transactions on Automatic Control*, 26, 301–20.

Edouard Djeutem
Department of Economics
Simon Fraser University
8888 University Drive
Burnaby, BC, V5A 1S6 CANADA
email: etsagued@sfu.ca

Kenneth Kasa
Department of Economics
Simon Fraser University
8888 University Drive
Burnaby, BC, V5A 1S6 CANADA
email: kkasa@sfu.ca