Notes on Engel-West, "Exchange Rates and Fundamentals" (JPE, 2005)

<u>What question does the paper address?</u>

The question is whether present-value models of nominal exchange rates are "good" models of the exchange rate.

A present value model for the exchange rate is one in which the exchange rate depends on the expected present discounted value of current and future fundamentals"

$$s_t = a\sum_{j=0}^{\infty} b^j E_t x_{t+j} ,$$

where a is some constant, b is a discount factor ($0 < b < 1$), and x_t represents the economic fundamentals.

For example, we have seen that the Dornbusch model can be written as:

$$s_t = \frac{1}{1+\lambda}\sum_{j=0}^{\infty}\left(\frac{\lambda}{1+\lambda}\right)^j E_t x_{t+j}\text{, where } x_t \equiv m_t - m_t^* + q_t.$$

The Taylor rule model can be written as:

$$s_t = \sum_{j=0}^{\infty}\left(\frac{1}{1+\delta\gamma}\right)^j E_t z_{t+j}\text{, where}$$

$$z_t \equiv \frac{\delta\gamma}{1+\delta\gamma}(p_t - p_t^*) - \frac{\rho}{1+\delta\gamma}E_t(\pi_{t+1} - \pi_{t+1}^*) - \frac{\gamma}{1+\delta\gamma}(v_t - v_t^*).$$

Why is this an interesting question?

There is substantial doubt in the literature that exchange rate models have any empirical validity. This doubt arises principally because the models cannot be used to forecast the change in the exchange rate, $s_{t+1} - s_t$. That is, time *t* information cannot be used to produce forecasts of changes in the exchange rate that are significantly better (have a significantly lower m.s.e.) than the forecast of “no change”: $s_{t+1} - s_t = 0$.

The most important paper in this vein is Meese and Rogoff (JIE, 1983.) They build sticky-price and flexible-price variants of the monetary model. They form “forecasts” of the exchange rate based on the model, but find that the forecast of no change has a lower m.s.e. than the model.

One surprising thing about Meese-Rogoff is that they use actual *ex post* values of fundamentals to make the model “forecasts”, but they still do no better than the forecast of no change.

This is not as bad for the models as it at first sounds. Why?

1. First, under the Meese-Rogoff methodology, the "no change" forecast at least gets to use s_t in its forecast of s_{t+k}. But Meese-Rogoff do not allow the models to have that information.

2. This methodology can always be manipulated to give the model really good or really bad *ex post* information. While it wasn't intentional, Meese-Rogoff did not give the models good *ex post* information.

Here is an example. Suppose we take the case of the flexible-price model that posits that the nominal exchange rate is determined by relative money supplies less relative money demand levels:

$$s_t = m_t - m_t^* + \lambda(i_t - i_t^*) + u_t - u_t^*.$$

Meese and Rogoff use $\hat{s}_{t+k} = m_{t+k} - m_{t+k}^* + \lambda(i_{t+k} - i_{t+k}^*)$ as the forecast of the *k*-period ahead exchange rate. If the model were true, the forecast would miss because of the error terms (perhaps money demand shocks), $u_{t+k} - u_{t+k}^*$.

But there are equally valid ways of writing the "equation" for the exchange rate. These produce different *ex post* forecasts that might have smaller errors.

For example, we could use covered interest parity to write $f_t - s_t = i_t - i_t^*$, where f_t is the log of the one-period forward rate.

Then we have:

$$s_t = m_t - m_t^* + \lambda(f_t - s_t) + u_t - u_t^*.$$

We can rewrite this as:

$$s_t = \frac{1}{1+\lambda}(m_t - m_t^*) + \frac{\lambda}{1+\lambda} f_t + \frac{1}{1+\lambda}(u_t - u_t^*).$$

If our forecast using *ex post* information was

$$\hat{s}_{t+k} = \frac{1}{1+\lambda}(m_{t+k} - m_{t+k}^*) + \frac{\lambda}{1+\lambda} f_{t+k},$$

the forecast error would be only $\frac{1}{1+\lambda}(u_{t+k} - u_{t+k}^*)$. The variance of the forecast error is much smaller than when we use $\hat{s}_{t+k} = m_{t+k} - m_{t+k}^* + \lambda(i_{t+k} - i_{t+k}^*)$.

My point is that it might seem surprising that we cannot do better than the forecast of "no change" even using *ex post* information, but it all depends on what *ex post* information you use.

However, in general the performance of exchange-rate models in genuine out-of-sample forecasting exercises is not good.

Mark (1995) has some success using models to forecast exchange rates at long-horizons.

Mark defines the "fundamental" determining exchange rates as:

$$x_t = m_t - m_t^* - (y_t - y_t^*).$$

He then posits an forecasting equation:

$$E_t(s_{t+k} - s_t) = \beta_k(x_t - s_t).$$

He finds some success forecasting dollar exchange rates for Canada, Switzerland, Germany, and Japan at long horizons (16 quarters.)

However, that paper has been criticized. The paper originally used data through 1991. Other researchers (for example, Faust, Rogers and Wright (JIE, 2003)) found that as the sample was extended into the 1990s, the forecasting power of the models disappeared.

Others questioned whether the bootstrapped significance levels in Mark's paper were correct, since he presupposed cointegration between the exchange rate and the fundamentals.

Chinn, Cheung, and Garcia Pascual (JIMF, 2005) find little consistent support for the claim that models have out-of-sample forecasting power.

Mark and Sul (JIE, 2001) and Groen (2000) use panel methods to derive some support for the claim that the models have power in out-of-sample forecasts at long horizons.

However, it is fair to say that there is little agreement that exchange rate models are useful in making forecasts of exchange rate changes out of sample.

How do Engel and West address the question?

The paper has two main points:

1. Many exchange rate models actually have the implication that the exchange rate change should be nearly unforecastable. That is, they imply that the exchange rate is nearly a random walk.

2. How then do we judge exchange rate models? Since exchange rate models are forward looking, then the exchange rate ought to incorporate useful information about the future fundamentals. One test of the models, then, is to test the implication that the exchange rate is useful in forecasting future fundamentals. They find weak evidence in favor of this hypothesis.

Theorem in Engel-West

Suppose the exchange rate is determined by a present-value model:

$$s_t = a\sum_{j=0}^{\infty} b^j E_t x_{t+j} .$$

Theorem states that if x_t is I(0), then as the discount factor gets large, $b \to 1$, the exchange rate approaches a random walk.

Here is the intuition of the theorem.

Do a Beveridge-Nelson decomposition of fundamentals: $x_t = rw_t + tr_t$. Then,

$$\begin{aligned} s_t &= \Sigma_{j=0}^{\infty} b^j E_t rw_{t+j} + \Sigma_{j=0}^{\infty} b^j E_t tr_{t+j} \\ &= rw_t /(1-b) + \Sigma_{j=0}^{\infty} b^j E_t tr_{t+j} \end{aligned}$$

$$s_t - s_{t-1} = \underbrace{(rw_t - rw_{t-1})/(1-b)} + \underbrace{\left(\Sigma_{j=0}^{\infty} b^j E_t tr_{t+j} - \Sigma_{j=0}^{\infty} b^j E_{t-1} tr_{t-1+j}\right)}$$

$$b \to 1 \qquad VAR \to \infty \qquad\qquad VAR \to \text{constant}$$

When the discount factor is close to one, markets put a high weight on the future. In the distant future, the influence of the transitory component diminishes. Uncertainty about the future values of fundamentals in the future primarily arises from uncertainty about the permanent component. As the discount factor goes to one, more and more weight is put on the future, and the influence of the permanent component of fundamentals is larger and larger. In the limit, the fraction of the variance of the exchange rate explained by the transitory component goes to zero.

Is this theory useful in practice? How close does *b* have to be to one?

Population Auto- and Cross-correlations of Δs_t

$$s_t = \sum_{j=0}^{\infty} b^j E_t x_{t+j};\ x_t \sim AR(2) \text{ with roots } \eta \text{ and } \varphi$$

b	η	φ		Δs_{t-1}	Δs_{t-2}	Δs_{t-3}		Δx_{t-1}	Δx_{t-2}	Δx_{t-3}
0.5	1	0.5		0.27	0.14	0.07		0.28	0.14	0.07
		0.8		0.52	0.42	0.34		0.56	0.44	0.36
0.9	1	0.5		0.05	0.03	0.01		0.06	0.03	0.01
		0.8		0.09	0.07	0.06		0.13	0.11	0.09
0.95	1	0.5		0.03	0.01	0.01		0.03	0.01	0.01
		0.8		0.04	0.04	0.03		0.07	0.05	0.04
0.90	0.90	0.5		0.04	-0.01	-0.03		0.02	-0.03	-0.05
0.90	0.95	0.5		0.05	0.01	-0.01		0.04	-0.00	-0.02
0.95	0.95	0.5		0.02	-0.00	-0.01		0.01	-0.02	-0.03
0.95	0.99	0.5		0.02	0.01	0.00		0.03	0.01	-0.00

How do we validate the models? Engel and West suggest that the exchange rate should have forecasting power for future fundamentals.

Indeed, there is a well-known methodology for testing present-value models derived by Campbell and Shiller (JPE, 1987). They exploit the fact that if the true model is $s_t = a\sum_{j=0}^{\infty} b^j E_t x_{t+j}$, then s_t should contain all the information relevant for forecasting the sum $a\sum_{j=0}^{\infty} b^j x_{t+j}$. That is, at time *t*, only s_t should be used in forecasting $a\sum_{j=0}^{\infty} b^j x_{t+j}$. They show how we can test the restriction that the "best" forecast of $a\sum_{j=0}^{\infty} b^j x_{t+j}$ is given by s_t.

Engel and West argue that the Campbell-Shiller methodology is too stringent because it applies to models where we can measure exactly the fundamental, x_t. In exchange rate models, some fundamentals are "unobserved" or mismeasured.

So Engel-West propose testing the weaker restriction, which is just that the exchange rate contains useful information in forecasting future values of x_t.

They test whether the exchange rate "Granger causes" x_t.

They find weak evidence in favor of this hypothesis. The exchange rate seems to be useful in forecasting some fundamentals for some countries, but the success is not uniform.

They conclude that we should not judge models by whether the model is useful in forecasting the exchange rate. But the model is supported by evidence that the exchange rate forecasts the future fundamentals.

Critique of Engel-West

1. The most obvious problem is the one just mentioned – that the exchange rate doesn't do all that well in forecasting future fundamentals.

2. It is possible that the theorem really does not apply to most models after all. In particular if uncovered interest parity does not hold, the behavior of the exchange rate may not depend much on the economic fundamentals as $b \to 1$.

Let me explain by way of an example, using the monetary model. To recall, we started with:

$$m_t - p_t - (m_t^* - p_t^*) = -\lambda(i_t - i_t^*).$$

We then used uncovered interest parity, $i_t - i_t^* = E_t s_{t+1} - s_t$, to write:

$$m_t - m_t^* + q_t - s_t = -\lambda E_t(s_{t+1} - s_t)$$

Instead, though, let's assume there is a deviation from uncovered interest parity, ρ_t, that represents a foreign exchange risk premium or a liquidity premium, or maybe a deviation from rational expectations:

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

Then we have:

$$m_t - m_t^* + q_t - s_t = -\lambda E_t (s_{t+1} - s_t + \rho_t) .$$

Then the exchange rate model is written as:

$$s_t = \frac{1}{1+\lambda}(m_t - m_t^* + q_t) + \frac{\lambda}{1+\lambda}\rho_t + \frac{\lambda}{1+\lambda}E_t s_{t+1}$$

Iterating forward, we get:

$$s_t = \frac{1}{1+\lambda}\sum_{j=0}^{\infty}\left(\frac{\lambda}{1+\lambda}\right)^j E_t x_{t+j} + \frac{\lambda}{1+\lambda}\sum_{j=0}^{\infty}\left(\frac{\lambda}{1+\lambda}\right)^j E_t \rho_{t+j} ,$$

where $x_t \equiv m_t - m_t^* + q_t$ as before.

This can be rewritten as:

$$s_t = (1-b)\sum_{j=0}^{\infty} b^j E_t x_{t+j} + b\sum_{j=0}^{\infty} b^j E_t \rho_{t+j} .$$

As Engel and West note, if the second term is present, then as $b \to 1$, it becomes increasingly important relative to the first term. As $b \to 1$, the behavior of ρ_t dominates. This is problematic for two reasons:

1. ρ_t is unobserved. So we have a model where the main driving force of the exchange rate is not observed.

2. Most people would think ρ_t is stationary. If it is stationary, the Engel-West theorem does not apply (though if its largest root is close to one, the Engel-West theorem still may apply approximately.)

Example: $s_t = (1-b)m_t + b\mu_t + bE_t(s_{t+1})$

$\Delta m_t = \phi \Delta m_{t-1} + u_t; \quad \Delta\mu_t = \gamma\Delta\mu_{t-1} + \varepsilon_t.$

$\Rightarrow$

$$\Delta s_{t+1} = \frac{\phi(1-b)}{1-b\phi}\Delta m_t + \frac{1}{1-b\phi}u_{t+1} + \frac{b\gamma}{1-b\gamma}\Delta\mu_t + \frac{b}{(1-b)(1-b\gamma)}\varepsilon_{t+1}$$

If $\mu_t = 0$, then as $b \to 1$, $\Delta s_{t+1} \to \frac{1}{1-\phi}u_{t+1}$.

If $\mu_t \neq 0$, then as $b \to 1$, $\Delta s_{t+1} \to (\text{large number}) \times \varepsilon_{t+1}$.

NOTE: *s* is <u>not</u> a random walk unless $b \to 1$