Notes on a Type of Learning and the Uncovered Interest Parity Puzzle

Larry Christiano

Following is an exploration of the interest rate and exchange rate consequences of the idea that people don't know whether the shock that drives the interest rate is permanent or temporary. Thus, suppose there is a positive money supply shock that drives down the domestic interest rate. Under UIP this must create an anticipated appreciation of the currency, or else all agents will sell assets in the domestic economy and buy foreign assets. In order for people to anticipate an appreciation, there must be an immediate depreciation. How much the depreciation must be depends on the persistence of the underlying shock. If the money shock is completely persistent, then the long-run effect on the exchange rate is to depreciate, and so the depreciation in the present period must be very great, to permit the eventual appreciation. If the shock is temporary, then the long-run exchange rate is unchanged, and the required current depreciation is very small.

Now, suppose that people don't know whether the shock is permanent or temporary. Initially, when they see the interest rate move, they will assume that it reflects whatever shock usually moves the exchange rate. Suppose, for example that the shock moving the money supply is usually temporary, but that the persistent one hits. In this case, people will expect there to be no long-run change in the exchange rate, and so there will only be a small depreciation. However, as time passes, and they see the money supply remaining high, they will slowly revise their beliefs about the long run exchange rate and this will require a higher current depreciation. In this way it is possible, in the wake of a drop in the interest rate, for the exchange rate to depreciate over time.

This note formalizes this argument, and then goes on to show that this line of reasoning does not provide a way to understand the violations of UIP documented in the data. It is still the case that if the domestic interest rate is higher than the foreign, then on average the currency must depreciate afterward. In the data, the result often goes the other way: when the domestic interest rate is high, on average there is an appreciation of the exchange rate (the 'UIP puzzle'). So, uncertainty about the persistence of the shocks underlying movements in the exchange rate does not offer a possible explanation of the UIP puzzle.

1. The Model

Consider the following model:

$$\begin{aligned} E_t e_{t+1} - e_t - R_t &= 0 \text{ (UIP)} \\ \Delta m_t - \Delta p_t + \alpha \Delta R_t &= 0 \text{ (MM)} \\ \Delta p_t - (1-\lambda)\,\Delta p_{t-1} - \lambda \Delta m_t &= 0, \text{ (Sticky Prices)} \end{aligned} \tag{1.1}$$

where m_t denotes the log of the money supply, p_t denotes the log price level, e_t denotes the log of dollars per unit of foreign currency, and R_t denotes the net domestic rate of interest. The first equation is the UIP condition, the second is the money demand equation with output set to zero and interest elasticity equal to α. This equation has been quasi-differenced using the operator, Δ. Thus, $\Delta p_t = p_t - \delta p_{t-1}$, and δ is a number close to, but not necessarily equal to, unity.

The log money stock, m_t, is driven by a permanent, P, and a transitory component, $T:$

$$m_t = m_t^T + m_t^P,$$

where

$$\begin{aligned} m_t^T &= \rho m_{t-1}^T + \varepsilon_t \\ \Delta m_t^P &= \phi \Delta m_{t-1}^P + u_t. \end{aligned}$$

When $\delta = 1$, this is the classic permanent/temporary decomposition. The univariate time series representation of m_t is obtained by first multiplying m_t by $(1-\rho L)(1-\phi L)\Delta:$

$$\begin{aligned} (1-\rho L)(1-\phi L)\Delta m_t &= (1-\phi L)(\varepsilon_t - \delta\varepsilon_{t-1}) + (1-\rho L)u_t && (1.2) \\ &= \eta_t + \theta_1 \eta_{t-1} + \theta_2 \eta_{t-2}, && (1.3) \end{aligned}$$

say, where η_t is the one-step-ahead forecast error in Δm_t based on using only past Δm_t to forecast Δm_t. (The representation involving η_t, (1.3), is called the *Wold representation* of Δm_t.) The following section derives formulas that can be used to compute θ_1, θ_2 and the variance of η_t, σ_η^2, using as input δ, ϕ, ρ, σ_ε^2, σ_u^2. That this is a Wold representation requires that the roots of the moving average representation lie inside the unit circle. The section after that solves the model, and simulates it.

2. The Univariate Representation of Money Growth

Write the error term in (1.2) in detail:

$$\varepsilon_t - \delta\varepsilon_{t-1} - \phi\varepsilon_{t-1} + \delta\phi\varepsilon_{t-2} + (1-\rho L)u_t = \varepsilon_t - (\delta+\phi)\varepsilon_{t-1} + \delta\phi\varepsilon_{t-2} + u_t - \rho u_{t-1}$$

Let $c_\varepsilon(\tau)$ and $c_u(\tau)$ denote the covariance function of the term involving ε_t's and u_t's, respectively. Note that the former is zero for all $\tau > 2$ and the latter is zero for all $\tau > 1$. Then,

$$\begin{aligned} c_\varepsilon(0) &= E\left[\varepsilon_t - (\delta+\phi)\varepsilon_{t-1} + \delta\phi\varepsilon_{t-2}\right]^2 \\ &= \sigma_\varepsilon^2\left[1 + (\delta+\phi)^2 + (\delta\phi)^2\right] \\ c_\varepsilon(1) &= E\left[\varepsilon_t - (\delta+\phi)\varepsilon_{t-1} + \delta\phi\varepsilon_{t-2}\right]\left[\varepsilon_{t-1} - (\delta+\phi)\varepsilon_{t-2} + \delta\phi\varepsilon_{t-3}\right] \\ &= -\left[(\delta+\phi) + \delta\phi(\delta+\phi)\right]\sigma_\varepsilon^2 = -(\delta+\phi)(1+\delta\phi)\sigma_\varepsilon^2 \\ c_\varepsilon(2) &= \delta\phi\sigma_\varepsilon^2. \\ c_u(0) &= \sigma_u^2\left(1+\rho^2\right) \\ c_u(1) &= -\rho\sigma_u^2. \end{aligned}$$

Then, the covariance function of the whole error is $c(\tau)$ for $\tau \geq 0:$

$$c(\tau) = c_\varepsilon(\tau) + c_u(\tau), \text{ all } \tau.$$

With $c(\tau)$ in hand, we can compute the parameters of the η_t process. We have that:

$$\begin{aligned} c(0) &= E\left[\eta_t + \theta_1\eta_{t-1} + \theta_2\eta_{t-2}\right]^2 \\ &= \sigma_\eta^2\left[1 + \theta_1^2 + \theta_2^2\right] \\ c(1) &= E\left[\eta_t + \theta_1\eta_{t-1} + \theta_2\eta_{t-2}\right]\left[\eta_{t-1} + \theta_1\eta_{t-2} + \theta_2\eta_{t-3}\right] \\ &= \theta_1(1+\theta_2)\sigma_\eta^2 \\ c(2) &= \theta_2\sigma_\eta^2. \end{aligned}$$

Then, solve the following two equations for θ_1 and θ_2 :

$$\begin{aligned}\frac{c(2)}{c(1)} &= \frac{\theta_2}{\theta_1(1+\theta_2)}\\ \frac{c(1)}{c(0)} &= \frac{\theta_1(1+\theta_2)}{1+\theta_1^2+\theta_2^2},\end{aligned}$$

subject to the eigenvalue condition. Finally,

$$\sigma_\eta^2 = \frac{c(1)}{\theta_1(1+\theta_2)}.$$

The eigenvalue condition is expressed as follows. Write:

$$1+\theta_1 L+\theta_2 L^2 = (1-\delta_1 L)(1-\delta_2 L).$$

We require that δ_1, δ_2 be less than unity in absolute value. For every ρ, δ, σ_ε, σ_u, there is always such a θ_1, θ_2.

We have the following connection between the one-step-ahead forecast errors based on the univariate representation of m_t, η_t, and the one-step-ahead forecast errors in the underlying fundamental shocks:

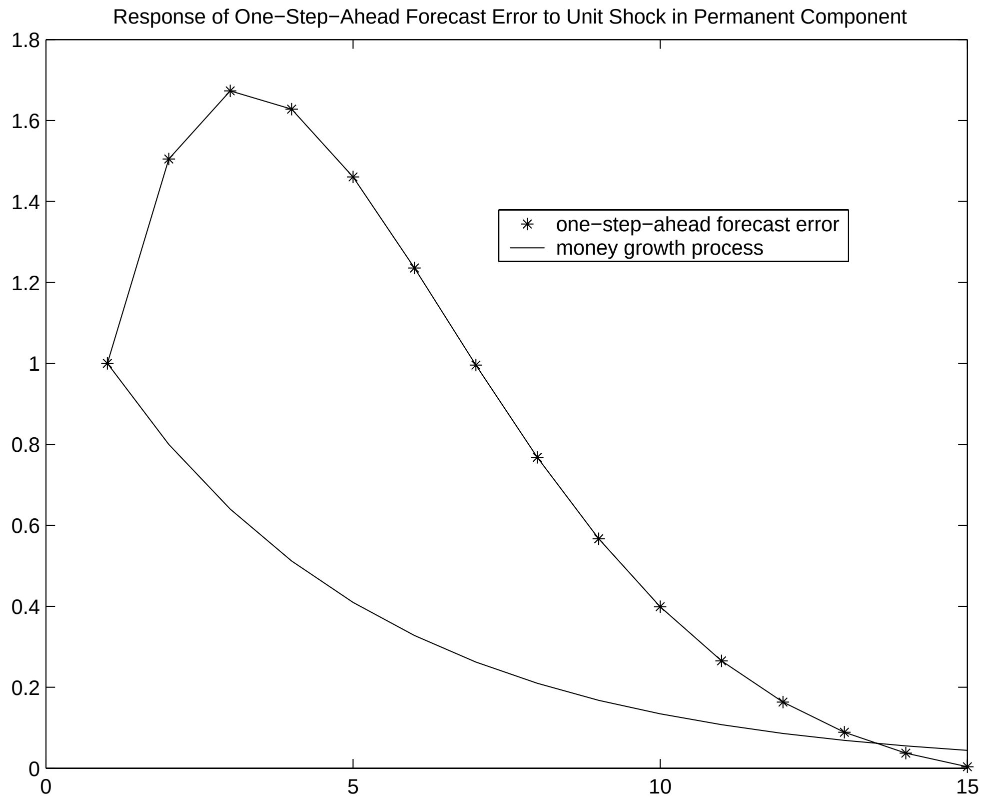
$$\eta_t = -\theta_1\eta_{t-1} - \theta_2\eta_{t-2} + (1-\phi L)(\varepsilon_t - \delta\varepsilon_{t-1}) + (1-\rho L)u_t.$$

Note that if there is a disturbance in u_t, this will produce a serially correlated sequence of one-step-ahead forecast errors in η_t's. The intuition is obvious. A jump in u_t will initially induce a same-magnitude jump in η_t. But, u_t induces a persistent move in m_t, via m_t^P. To the extent that $\sigma_\varepsilon > 0$, so that a rational forecaster will attribute some possibility to the source of the disturbance being ε_t, the dynamic move in m_t will be interpreted as a sequence of same-sign shocks in ε_t. However, as time evolves, such a sequence becomes less likely and eventually it is 'learned' that the source of the shock in fact was u_t. Throughout this period the rational forecaster makes same-sign one-step-ahead forecast errors in forecasting m_t.

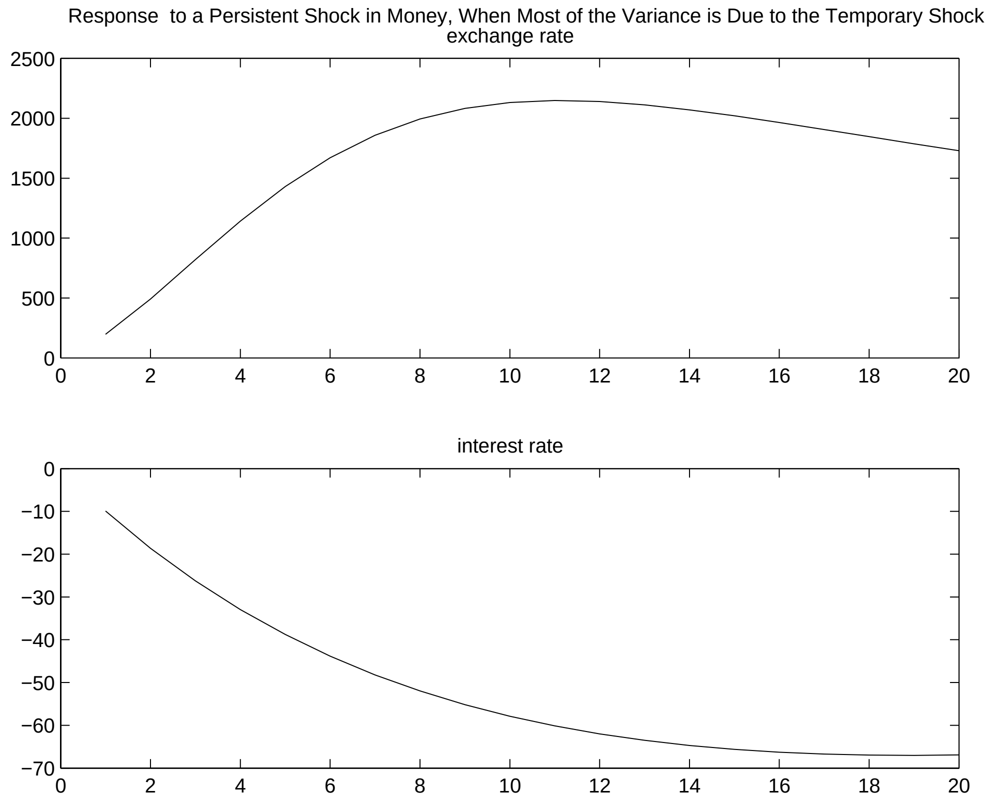
Here is a (rather extreme) example:

$$\sigma_\varepsilon^2 = 1,\ \rho = 0,\ \delta = 0.99,\ \phi = 0.8,\ \sigma_u^2 = 0.01.$$

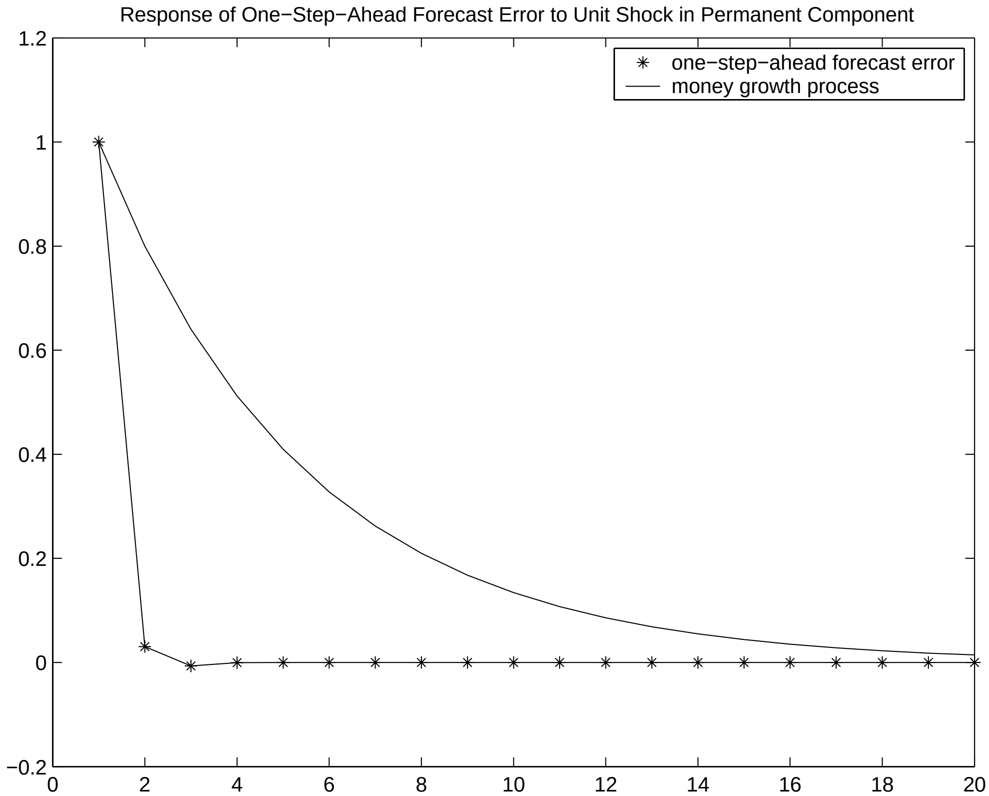
A unit disturbance in u_t pushes up the one-step-ahead forecast error in Δm_t by unity. Since most of the variance in the series comes from ε_t, a rational forecaster will assume the shock came from ε_t. However, since in fact it came from u_t, Δm_{t+1} will be surprisingly high too. So, there will be another shock then, as the forecaster assumes ε_{t+1} must have been high too, and so on. A graph of η_t in response to the unit shock in u_t, as well as the response in Δm_t appears below:



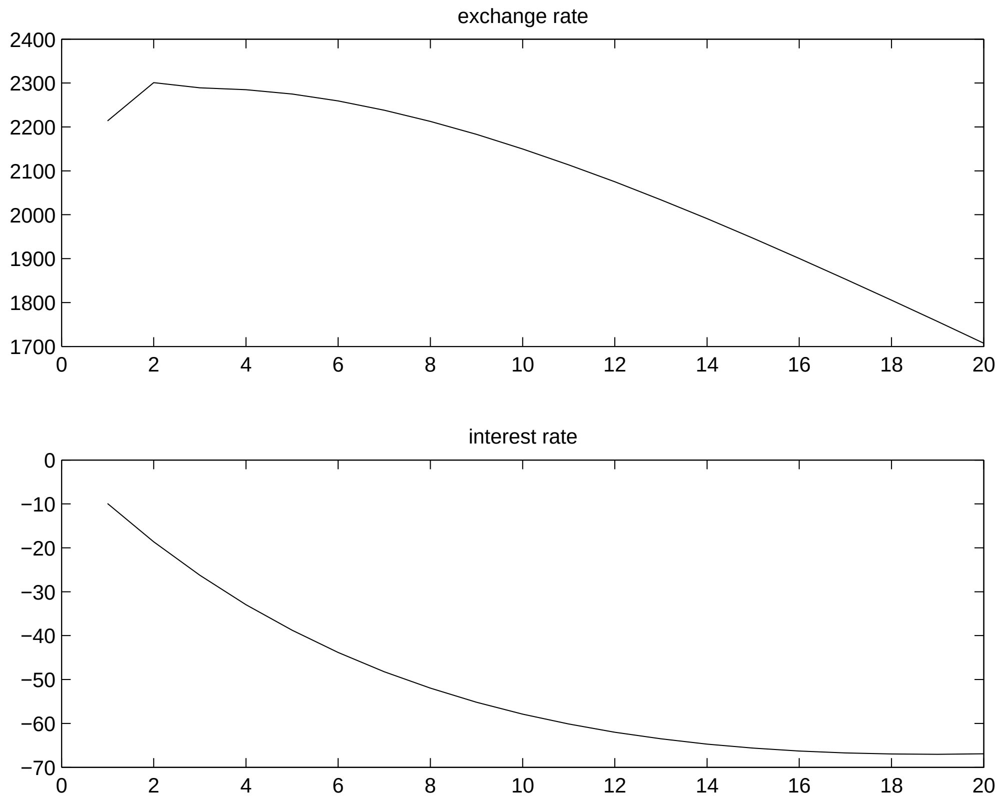
Note how big and long-lasting the one-step-ahead forecast errors are. In some sense, the rational forecaster never learns! They continue to be surprised all along the way. It is interesting to see how the interest rate and exchange rate respond in this case. Note how the exchange rate depreciates with the drop in the interest rate, and it continues to depreciate for 10 periods!



For comparison, consider the opposite extreme, where most of the variance in Δm_t comes from the permanent component. Thus, $\sigma_u^2 = 1$ and $\sigma_\varepsilon^2 = 0.01$. Then,



The picture now is quite different. In the period of the shock, the forecast error is just the surprise in the money growth rate. In the next period, there is a bit of a forecast error, because they assign some probability to the possibility that the shock was temporary. however, after the second period, they've caught on and there is no longer any forecast error. The rational forecaster correctly forecasts the implications of the surprise in Δm_t, because the surprise comes from the place it usually comes from. The following picture shows what happens to the exchange rate and the interest rate in this example:



Note that the exchange rate does depreciate for one period. But, right after that it resorts to the pattern one would expect.

3. Solving and Simulating the Model

To simulate this model, it is appropriate to do so using the univariate time series representation for Δm_t :

$$\Delta m_t = (\rho + \phi)\,\Delta m_{t-1} - \rho\phi\Delta m_{t-2} + \eta_t + \theta_1\eta_{t-1} + \theta_2\eta_{t-2}.$$

It is convenient to express Δm_t as a vector first order autoregression:

$$\begin{pmatrix} \Delta m_t \\ \Delta m_{t-1} \\ \eta_t \\ \eta_{t-1} \end{pmatrix} = \begin{bmatrix} \rho+\phi & -\rho\phi & \theta_1 & \theta_2 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} \Delta m_{t-1} \\ \Delta m_{t-2} \\ \eta_{t-1} \\ \eta_{t-2} \end{pmatrix} + \begin{pmatrix} \eta_t \\ 0 \\ \eta_t \\ 0 \end{pmatrix},$$

or,

$$s_t = P s_{t-1} + \begin{pmatrix} \eta_t \\ 0 \\ \eta_t \\ 0 \end{pmatrix},$$

where η_t is iid over time, with variance σ_η. Let

$$z_t = \begin{pmatrix} e_t \\ \Delta p_t \\ R_t \end{pmatrix},$$

and write the system of three equations in matrix form as follows:

$$E_t \left[\alpha_0 z_{t+1} + \alpha_1 z_t + \alpha_2 z_{t-1} + \beta_0 s_{t+1} + \beta_1 s_t \right] = 0,$$

where

$$\begin{aligned}
\alpha_0 &= \begin{bmatrix} \beta & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; \alpha_1 = \begin{bmatrix} -1 & 0 & -\beta \\ 0 & -1 & \alpha \\ 0 & 1 & 0 \end{bmatrix}, \; \alpha_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -\delta\alpha \\ 0 & -(1-\lambda) & 0 \end{bmatrix} \\
\beta_0 &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \; \beta_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -\lambda & 0 & 0 & 0 \end{bmatrix}
\end{aligned}$$

Given this solution and given a random draw on η_t, the system can be simulated for T observations:

$$\begin{aligned}
s_t &= P s_{t-1} + \begin{pmatrix} \eta_t \\ 0 \\ \eta_t \\ 0 \end{pmatrix} \\
z_t &= A z_{t-1} + B s_t.
\end{aligned}$$

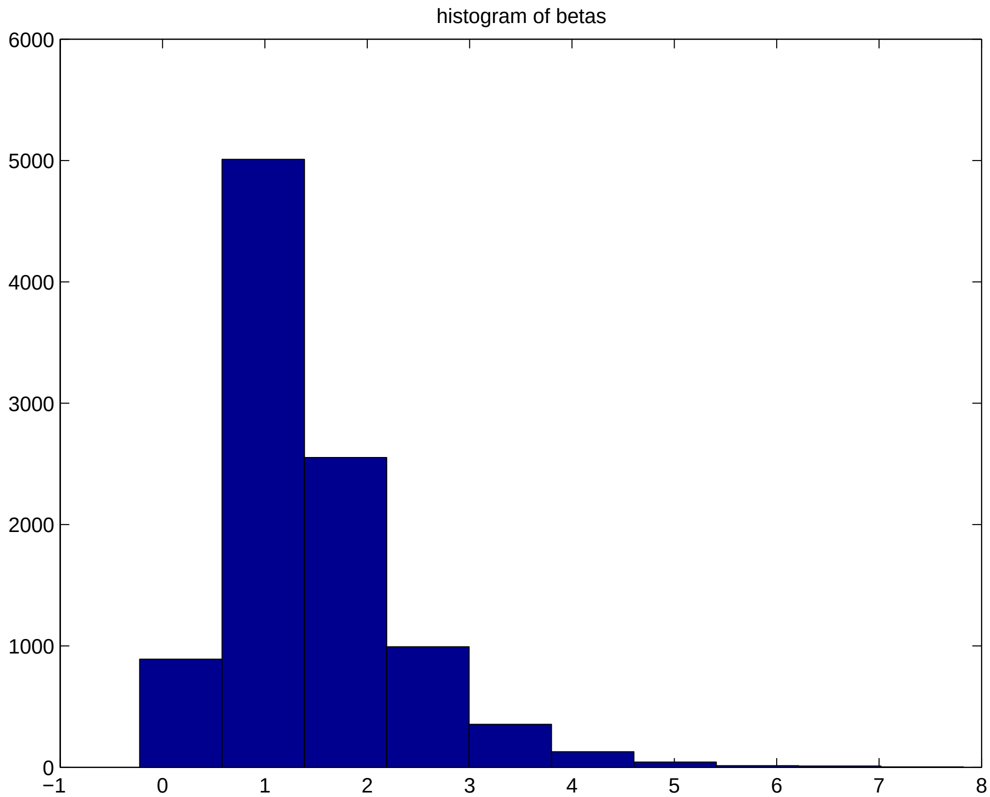
The change in the exchange rate, $e(t+1) - e(t)$, can be recovered from the first element in z_{t+1} and z_t and the interest rate, R_t, can be obtained from the third element of z_t. Then, the slope in the regression of $e(t+1) - e(t)$ on R_t can be computed from:

$$\hat{\beta} = \frac{cov\left(e\left(t+1\right) - e\left(t\right), R_t\right)}{var\left(R_t\right)}.$$

This was done for the following parameterization:

$$\alpha = 0.1, \; \sigma_\varepsilon^2 = .1, \; \rho = 0, \; \delta = 0.99, \; \phi = 0.9, \; \sigma_u^2 = 0.01, \; \lambda = 0.01$$

After doing 10,000 replications with $T = 200$ data points, the following histogram of $\hat{\beta}$'s resulted:



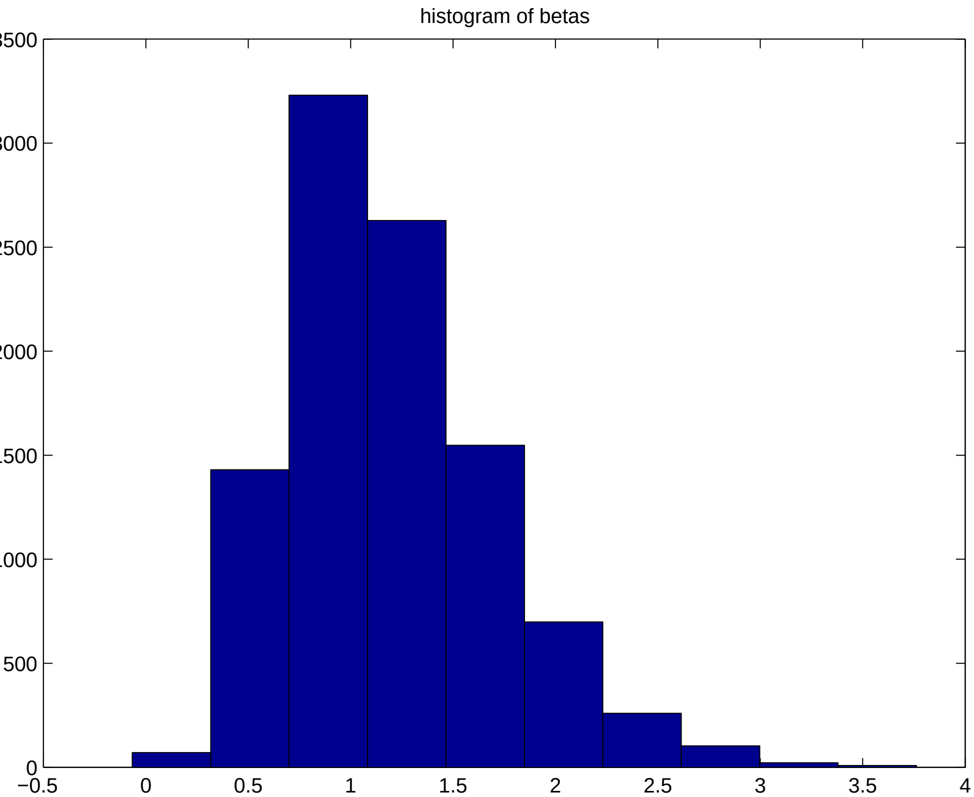
Interestingly, here there is actually an upward bias. The mean, across 10,000 replications, of $\hat{\beta}$, is 1.42. The standard deviation huge, 0.85. However, this should not be interpreted to suggest that negative $\hat{\beta}$'s are possible. As the histogram makes clear, the distribution is skewed to the right. Very high numbers are possible, but not very low numbers, such as the ones estimated in the data. Interestingly,

Note that there is little bias here. Consider now the 'best case scenario' for the big bias, the example in which

$$\beta = .99, \ \delta = 0.99, \ \alpha = .1, \ \lambda = 0.01, \ \rho = 0, \ \phi = 0.8, \ \sigma_u^2 = 0.01, \ \sigma_\varepsilon^2 = 1$$

In this case, the mean (across 10,000 simulations) of $\hat{\beta}$ is 1.20, and the standard deviation is

0.51. The histogram of the $\hat{\beta}$'s is



The point is that the considerations raised here do nothing to resolve the UIP puzzle.