FORWARD BIAS, UNCOVERED INTEREST PARITY AND RELATED PUZZLES: THE ROLE OF MONETARY POLICY

John Pippenger
Department of Economics
University of California
Santa Barbara, CA 93106

The forward-bias puzzle and the failure of uncovered interest parity appear to be the result of monetary policy. That is not a new observation, but I believe that I provide the first detailed economic explanation of how monetary policy creates those puzzles. In addition, I show that, unlike alternative explanations of the forward-bias puzzle and the failure of uncovered interest parity like risk premia and the failure of rational expectations, monetary policy can explain the related puzzles.

Author: pipp@ix.netcom.com, 619-423-3618.

Key words: exchange rates; interest rates; risk premia; rational expectations; uncovered and covered interest parity; forward bias; arbitrage.


DRAFT VERSION AS OF 28 JULY 2014. PLEASE DO NOT CITE OR REFER TO THIS DRAFT WITHOUT PERMISSION OF THE AUTHOR.

*I want to thank Jeffrey Frankel, Jim Hamilton, Alan King, David Peel and Doug Steigerwald for their comments and/or suggestions. Any remaining errors are of course mine.
1. Introduction.

Several puzzles are closely related to the forward-bias puzzle and the failure of uncovered interest parity (UIP): (1) the Commodity Puzzle; (2) the Carry Trade Puzzle; (3) the Development Puzzle; (4) the Levels versus Differences Puzzle; (5) the Maturity Puzzle and (6) the Time Dependency Puzzle. Monetary policy appears to provide the best available explanation for all these puzzles.

2. The Puzzles.

2.1. Uncovered Interest Parity.

The theory of uncovered interest parity says that the expected change in the log of the exchange rate, \( E(\Delta s_{t+k}) \), equals the current interest rate differential, \( i_t - i^*_t \). Where \( E(s_{t+k}) \) is the expected log of the domestic price of foreign exchange while \( i_t \) and \( i^*_t \) are risk-free domestic and foreign interest rates with the same maturity as \( s_{t+k} \).

One approach to the theory uses covered interest parity (CIP) and assumes that speculation equates \( E(s_{t+k}) \) and the log of the current forward rate denoted \( f_t \). Since there is a large body of evidence supporting CIP, when UIP fails, it is natural to question the assumption that \( E(s_{t+k}) \) equals \( f_t \). That question appears to be the source of the idea that risk premiums cause UIP to fail.

Another approach uses the expectations version of the Fisher equation, an expectations version of purchasing power parity (EPPP) and the assumption that real interest rates are equal. If the nominal interest rate equals the real rate plus the expected rate of inflation, then \( i_t - i^*_t \) approximately equals the difference in expected rates of inflation plus the difference in real interest rates. If the real differential is zero, then \( i_t - i^*_t \) equals the difference in expected rates of inflation. Assuming that \( E(\Delta s_{t+k}) \) equals the difference in expected inflation, EPPP, produces eq. (1). When UIP fails, this approach suggests that it fails because the Fisher equation fails, EPPP fails, or real rates are not equal.

---

1 This is the approach used in the New Palgrave Dictionary of Economics, Vol. 1 (2008, 451).
In either case, eq. (1) is the standard test equation where estimates of \( b \) should equal 1.0
\[
\Delta s_{t+k} = a + b(i_t - i_t^*) + e_{t+k}
\] (1)
But estimates of \( b \) denoted \( \hat{b} \) are often negative. For a recent example see Aslan and Korap (2010).

If expectations are rational, \( e_{t+k} \) has a zero mean, is uncorrelated and orthogonal to \( i_t - i_t^* \). Since the failure of rational expectations implies that \( e_{t+k} \) and \( i_t - i_t^* \) can be correlated, the failure of rational expectations provides another explanation for why \( \hat{b} \) might not equal 1.0. But that failure does not by itself explain why \( \hat{b} \) are often negative and significant.

2.2. Forward-Bias.

The modern forward-bias puzzle begins with Fama (1984) who splits \( f_t \) into \( E(s_{t+1}) \) and a "premium" denoted \( p_t \).
\[
f_t = E(s_{t+1}) + p_t
\] (2)
Although the subsequent literature usually interprets \( p_t \) as a risk premium, Fama (1984) is cautious about labeling it. He usually refers to \( p_t \) simply as a "premium", not a "risk premium".

Fama’s premium should be interpreted as the expected return to speculation for several reasons. First, \( p_t \) equals \( f_t - E(s_{t+1}) \), which is the expected return. If \( E(s_{t+1}) \) is greater than \( f_t \), a speculator expects to earn \( E(s_{t+1}) \) minus \( f_t \) by buying the foreign currency forward and then selling it later for \( E(s_{t+1}) \). If \( f_t - E(s_{t+1}) \) is positive, speculators expect to make a return by selling the foreign currency forward and covering the sale by buying the currency at the lower \( E(s_{t+1}) \). I discuss speculation in terms of forward contracts to simplify the exposition. Speculators who do not cover an equivalent CIP position expect to earn \( E(s_{t+1}) \) minus \( f_t \). So the expected speculative return is the same either way.

Second, it makes more sense to consider expected speculative returns as the 'cause' of the forward bias. Consider an initial equilibrium where expected returns to speculation and risk premia are both zero. There is no forward bias, UIP holds and speculators hold no forward contracts. An exogenous increase in the expected speculative return \( f_t - E(s_{t+1}) \) causes UIP to fail and induces
speculators to sell the foreign currency forward. That forward position creates a risk premium for risk adverse speculators. But unless an exogenous increase in risk can somehow create an expected speculative return, it cannot create a forward bias or cause UIP to fail. Starting in the same equilibrium, an exogenous increase in risk discourages speculators from taking any position. As a result, expected returns remain zero, there is no forward bias and UIP continues to hold.

A third reason for calling \( p_t \) an expected speculative return is that risk premia cannot explain the 'carry trade', but expected speculative returns can.

A fourth reason is that it is a mistake to assume that risk neutrality implies UIP. Risk neutrality only guarantees that there is no risk premium. Risk neutrality does not guarantee that UIP holds because it does not guarantee that there is no expected speculative return.

Consider an equilibrium in which \( f_t - E(s_{t+1}) \) is zero, speculators hold no forward contracts and there is no risk premium. If an exogenous shock creates an expected speculative return, risk neutrality alone does not guarantee that UIP holds. If portfolio adjustment is costly, in the short run there may not be enough speculative funds available to eliminate the expected return. UIP fails and there is a forward bias. During the transition, adjustment costs can offset expected speculative returns. UIP fails and there is a forward bias even though investors are risk neutral. Claiming that risk neutrality implies that there is no forward bias, and/or that UIP holds, implicitly assumes that adjustment costs are zero.

As Fama points out, eq. (2) implies eq. (3).

\[
f_t - s_t = E(\Delta s_{t+1}) + p_t
\]  

Rearranging eq. (3) and assuming rational expectations produces eq. (4).

\[
\Delta s_{t+1} = (f_t - s_t) - p_t + \epsilon_{t+1}
\]  

Omitting \( p_t \) produces the ‘Fama equation’.
\[ \Delta s_{t+1} = \alpha + \beta(f_t - s_t) + \varepsilon_{t+1} \] (5)

The forward-bias puzzle is that, over a wide variety of time periods when exchange rates are flexible and maturities short, \( \hat{\beta} \) between developed countries are often negative and significant.

2.3. Commodity Puzzle.

One would expect Fama’s premium to be as valid for commodity markets as for foreign-exchange markets. The premium \( f_t - E(s_{t+1}) \) can refer to the domestic price of wool as well as the domestic price of foreign exchange. Eq. (4) would appear to be as relevant for wool, or any other commodity with forward markets, as it is for foreign-exchange markets.

Given the importance of the bias in foreign-exchange markets, looking for the same puzzle in commodity markets would seem an obvious and important thing to do. To the best of my knowledge there have been only two attempts to do so: Fama and French (1987) and Kearns (2007). Using futures indexes, Kearns finds positive \( \hat{\beta} \) for commodities. Using individual futures prices, Fama and French find mostly positive \( \hat{\beta} \) for commodities Frequent negative \( \hat{\beta} \) for flexible exchange rates and mostly positive \( \hat{\beta} \) for commodities is the Commodity Puzzle.

2.4 Carry Trade

As pointed out earlier, the term ‘carry trade’ refers to borrowing where rates are ‘low’ and lending where they are ‘high’, which appears to produce profit with little risk. I call this the Carry Trade Puzzle. For some recent articles on the carry trade see Burnside et al (2008), Hochradl and Wagner (2010) and Baillie and Chang (2011).

2.5. Development puzzle.

Apparently Bansal and Dahlquist (2000) were the first to suggest that the forward-bias puzzle was confined largely to developed countries. Frankel and Poonawala (2010) estimate eq. (5) between the U.S. and 21 developed countries and 14 developing countries. The average \( \hat{\beta} \) between developed...
countries and the U.S. is −4.3 while the average \( \hat{\beta} \) between developing countries and the U.S. is 0.003. The difference between −4.3 and 0.003 is the Development Puzzle.

2.6. Levels versus Differences.

The issue of levels versus differences is one of the oldest puzzles associated with the forward-bias puzzle. Researchers initially regressed \( s_{t+1} \) against \( f_t \) and often found coefficients close to one.\(^2\) Then the issue of spurious regressions arose. To avoid that problem, researchers regressed \( s_{t+1} - s_t \) against \( f_t - s_t \). They found coefficients that were routinely negative and almost never close to one. I call those different results the Levels versus Differences Puzzle.

2.7. Maturity puzzle.

Using developed countries, Alexius (2001), Chinn and Meredith (2004) and Chinn (2006) find that \( \hat{\beta} \) are usually negative for maturities of one year or less but that \( \hat{\beta} \) are usually positive for those over one year. More recently Lothian and Wu (2011) find the same pattern using over 200 years of data. I call negative \( \hat{\beta} \) for short maturities and positive \( \hat{\beta} \) for long maturities the Maturity Puzzle.

2.8. Time Dependency Puzzle.

Both \( \hat{\beta} \) and \( \hat{\beta} \) vary widely over time. For an example of the time dependency of \( \hat{\beta} \) see Baillie (2011). For examples of the time dependency of \( \hat{\beta} \) see Han (2004) and Lothian and Wu (2011). I call those time dependent results the Time Dependency Puzzle.

3.0. CIP.

As long as CIP holds, the forward-bias puzzle and failure of UIP are two sides of the same coin. Akram et al (2008) probably provide the best available analysis of covered interest parity.\(^3\)

Overall, the evidence is consistent with the Grossman-Stiglitz view of financial markets where efficiency is not interpreted as a statement about prices being correct at each point in time but the notion that in efficiently-functioning financial markets very short-

---

\(^2\) For a survey of the early literature see Levich (1979).

\(^3\) For additional work on CIP see Fong et al (2010) and the work cited there.
term arbitrage opportunities can arise and invite traders to exploit them, which makes it worthwhile to watch the relevant markets. This is the arbitrage mechanism that restores the arbitrage-free prices we observe on average. Nevertheless, the lack of predictability of arbitrage and the fast speed at which arbitrage opportunities are exploited and eliminated imply that a typical researcher in international macro-finance using data at the daily or lower frequency can safely assume that CIP holds. Akram et al (2008, 238)

Akram et al make it clear that CIP is not an identity or anything like an identity. Identities hold in every possible state of the world. They describe occasional CIP failures. Indeed, as they point out, the success of the theory depends on those occasional failures.

Akram et al (2008) also say that one can safely assume that CIP holds for daily and lower frequency data. They do not distinguish between developed and developing countries nor between short and long maturities. Since they use maturities of one year or less from developed countries, strictly speaking their conclusions only hold under those conditions. Whether or not arbitrage is as effective in developing markets and for longer maturities remains an open issue.

Eq. (6) describes the theory of covered interest parity.

\[ f_t - s_t - (i_t - i_t^*) = d_t \] (6)

Where the forward premium and interest rate differential have the same maturity and \( d_t \) is the deviation from CIP. For good data \( d_t \) should be primarily the result of transaction costs, particularly bid-ask spreads.

For now I ignore \( d_t \) and assume that covered interest parity holds exactly because that simplifies the discussion. I will return to \( d_t \) later.

4.0. Some Evidence.

This section reconfirms and extends some of the relevant evidence.

4.1. Data.

My data cover four intervals: (1) weekly data from the Federal Reserve Bulletin for the 1960s between the U.S. and Canada when rates were pegged, (2) weekly data from the Federal Reserve Bulletin for the early 1970s between the U.S. and Canada when rates were flexible, (3) daily data from
Balke and Wohar (1998) for 1977 to 1993 between the U.S. and U.K. when rates were flexible and (4) weekly data from Einzig (1937) for the early 1920s between the U.S. and U.K. when rates were flexible.4

Quality differs. The best are from Balke and Wohar (1998). The Einzig data are not quite as good. On page 1253 the Federal Reserve Bulletin for October 1964 warns that "…the interest arbitrage incentives shown in these tables provide only an approximate indication of the covered differentials in treasury bill yields in the specified markets."

4.2. Forward Bias and UIP.

All regressions use RATS with Robusterrors. Durbin-Watson statistics are low due to multi-period overlapping horizons.

All the $\hat{\beta}$ in Table 1 are negative and two are significant at 1%. All but one $\hat{\beta}$ is negative with one significant at 5% and two at 1%.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Forward bias and UIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta s_{t+1} = a + b(i_t - i^*<em>t) + e</em>{t+1}$</td>
<td>$\Delta s_{t+1} = \alpha + \beta(f_t - s_t) + \epsilon_{t+1}$</td>
</tr>
<tr>
<td>$\hat{\alpha}$</td>
<td>$\hat{\beta}$</td>
</tr>
<tr>
<td>2.18**</td>
<td>-6.12**</td>
</tr>
<tr>
<td>(0.20)</td>
<td>(0.60)</td>
</tr>
<tr>
<td>0.26**</td>
<td>-0.40</td>
</tr>
<tr>
<td>(0.04)</td>
<td>(0.27)</td>
</tr>
<tr>
<td>-0.43**</td>
<td>-0.91</td>
</tr>
<tr>
<td>(0.12)</td>
<td>(0.50)</td>
</tr>
<tr>
<td>0.61**</td>
<td>-2.45**</td>
</tr>
<tr>
<td>(0.06)</td>
<td>(0.21)</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. * Significant at 5%. ** Significant at 1%.

Estimates like those in Table 1 are 'static'. Cross spectra from $i_t - i^*_t$ to $\Delta s_{t+1}$ show the dynamic interaction between $i_t - i^*_t$ and $\Delta s_{t+1}$. Table 2 shows that cross spectrum using the Balke-Wohar data.5

---

4 In Einzig (1937) and the Bulletin rates are three month. For information on Canadian data see the Federal Reserve Bulletin for October 1964. In Balke and Wohar (1998) rates are one month. See Balke and Wohar for more information. All data are available on request.
Table 2

<table>
<thead>
<tr>
<th>Cycles/Month</th>
<th>Coherency Squared</th>
<th>Gain</th>
<th>Phase</th>
<th>Cycles/Month</th>
<th>Coherency Squared</th>
<th>Gain</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.1069</td>
<td>1.52828</td>
<td>0.5584</td>
<td>0.263</td>
<td>0.1244</td>
<td>16.1140</td>
<td>0.5809</td>
</tr>
<tr>
<td>0.026</td>
<td>0.1586</td>
<td>2.0745</td>
<td>0.5806</td>
<td>0.289</td>
<td>0.1579</td>
<td>19.3720</td>
<td>0.5414</td>
</tr>
<tr>
<td>0.053</td>
<td>0.2651*</td>
<td>3.1970*</td>
<td>0.5454†</td>
<td>0.316</td>
<td>0.0203</td>
<td>6.3909</td>
<td>0.3945</td>
</tr>
<tr>
<td>0.079</td>
<td>0.2643*</td>
<td>4.3945*</td>
<td>0.4741†</td>
<td>0.342</td>
<td>0.1078</td>
<td>11.3404</td>
<td>0.0781</td>
</tr>
<tr>
<td>0.105</td>
<td>0.1332</td>
<td>4.5222</td>
<td>0.4696</td>
<td>0.368</td>
<td>0.1227</td>
<td>13.3537</td>
<td>0.1174</td>
</tr>
<tr>
<td>0.132</td>
<td>0.0854</td>
<td>3.9999</td>
<td>0.4443</td>
<td>0.395</td>
<td>0.0540</td>
<td>7.5548</td>
<td>0.2833</td>
</tr>
<tr>
<td>0.158</td>
<td>0.1441</td>
<td>7.7053</td>
<td>0.3975</td>
<td>0.421</td>
<td>0.1123</td>
<td>9.0880</td>
<td>0.3651</td>
</tr>
<tr>
<td>0.184</td>
<td>0.0483</td>
<td>5.5169</td>
<td>0.3484</td>
<td>0.447</td>
<td>0.0557</td>
<td>7.9280</td>
<td>0.2986</td>
</tr>
<tr>
<td>0.211</td>
<td>0.0364</td>
<td>4.8144</td>
<td>0.2610</td>
<td>0.474</td>
<td>0.0038</td>
<td>2.1039</td>
<td>0.4568</td>
</tr>
<tr>
<td>0.237</td>
<td>0.0277</td>
<td>5.7416</td>
<td>0.4969</td>
<td>0.500</td>
<td>0.0572</td>
<td>7.6211</td>
<td>0.6029</td>
</tr>
</tbody>
</table>

* Significantly different from 0.0 at 5% level. † Significantly different from zero but not significantly different from 0.5 at 5% level.

A negative $\hat{b}$ corresponds to a phase angle of 0.5. For cycles less than one year (0.08 cycles/month) and greater than 19 months (0.053 cycles/month) coherence squared and gain are both insignificant. Phase angles are also not significantly different from zero (1.0). Between one year and 19 months both coherence squared and gain are significant and the phase angle is significantly different from zero, but not significantly different from 0.5, which is consistent with a negative $\hat{b}$.

4.3. Levels versus differences.

Table 1 shows the results when using differences. Table 3 reports the results of regressing $s_{t+1}$ against $f_t$. All $\hat{b}$ are significantly greater than zero. For the best data, $\hat{b}$ is close to one.

Table 3
Levels

<table>
<thead>
<tr>
<th>$s_{t+1} = a + B f_t + e_{t+1}$</th>
<th>$\hat{a}$</th>
<th>$\hat{b}$</th>
<th>$\hat{R}^2$/DW</th>
<th>$\hat{a}$</th>
<th>$\hat{b}$</th>
<th>$\hat{R}^2$/DW</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S.-U.K.: 1922-1925 (Flexible)</td>
<td>2.55**</td>
<td>0.58**</td>
<td>0.293</td>
<td>2.18**</td>
<td>0.70**</td>
<td>0.794</td>
</tr>
<tr>
<td>(0.30)</td>
<td>(0.05)</td>
<td>0.069</td>
<td>(0.30)</td>
<td>(0.04)</td>
<td>0.121</td>
<td></td>
</tr>
<tr>
<td>U.S.-Canada: 1961-1969 (Pegged)</td>
<td>2.18**</td>
<td>0.70**</td>
<td>0.794</td>
<td>-1.47**</td>
<td>0.98**</td>
<td>0.952</td>
</tr>
<tr>
<td>(0.30)</td>
<td>(0.04)</td>
<td>0.121</td>
<td>(0.23)</td>
<td>(0.00)</td>
<td>0.071</td>
<td></td>
</tr>
</tbody>
</table>

Standard errors in parentheses. * Significant at 5%. ** Significant at 1%.

---

5 I convert their daily data into 'monthly' data using a month of 22 business days.
4.4. Time Dependency.

Table 4 uses daily Balke-Wohar data to illustrate the time dependency of $\hat{\beta}$ and $\hat{b}$ over the same intervals. I believe that this is the first time that has been done. The first interval uses the first 1,000 observations to obtain $\hat{\beta}_1$ and $\hat{b}_1$. The second interval uses observations 101 to 1,100 to obtain $\hat{\beta}_2$ and $\hat{b}_2$. The third interval uses observation 201 to 1,200 to obtain $\hat{\beta}_3$ and $\hat{b}_3$ and so on. As one might expect, $\hat{\beta}_t$ and $\hat{b}_t$ move together.

Table 4

<table>
<thead>
<tr>
<th>Int.</th>
<th>$\hat{\beta}_t$</th>
<th>$\hat{b}_t$</th>
<th>Int.</th>
<th>$\hat{\beta}_t$</th>
<th>$\hat{b}_t$</th>
<th>Int.</th>
<th>$\hat{\beta}_t$</th>
<th>$\hat{b}_t$</th>
<th>Int.</th>
<th>$\hat{\beta}_t$</th>
<th>$\hat{b}_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.64</td>
<td>0.65</td>
<td>10</td>
<td>-2.60</td>
<td>-2.67</td>
<td>19</td>
<td>-7.89</td>
<td>-8.48</td>
<td>27</td>
<td>-5.30</td>
<td>-5.87</td>
</tr>
<tr>
<td>3</td>
<td>-3.47</td>
<td>-3.62</td>
<td>12</td>
<td>-2.90</td>
<td>-3.02</td>
<td>21</td>
<td>-11.00</td>
<td>-12.40</td>
<td>29</td>
<td>-2.59</td>
<td>-2.42</td>
</tr>
<tr>
<td>4</td>
<td>-3.47</td>
<td>-3.62</td>
<td>13</td>
<td>-5.30</td>
<td>-5.32</td>
<td>22</td>
<td>-10.48</td>
<td>-12.35</td>
<td>30</td>
<td>-4.73</td>
<td>-4.73</td>
</tr>
<tr>
<td>6</td>
<td>-3.08</td>
<td>-3.20</td>
<td>15</td>
<td>-7.80</td>
<td>-8.19</td>
<td>24</td>
<td>-5.00</td>
<td>-7.60</td>
<td>32</td>
<td>-4.99</td>
<td>-4.83</td>
</tr>
<tr>
<td>7</td>
<td>-2.76</td>
<td>-2.85</td>
<td>16</td>
<td>-8.01</td>
<td>-8.49</td>
<td>25</td>
<td>-4.87</td>
<td>-6.47</td>
<td>33</td>
<td>-6.49</td>
<td>-4.71</td>
</tr>
<tr>
<td>8</td>
<td>-2.57</td>
<td>-2.65</td>
<td>17</td>
<td>-7.00</td>
<td>-7.49</td>
<td>26</td>
<td>-7.89</td>
<td>-8.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-2.41</td>
<td>-2.50</td>
<td>18</td>
<td>-7.60</td>
<td>-8.13</td>
<td>27</td>
<td>-5.30</td>
<td>-5.87</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.5. Transaction costs.

Transaction costs help explain the development puzzle. With the exception of Balke and Wohar (1998) and Paya et al (2010), most of the literature ignores transaction costs.

Using U.S.-U.K. data for the early 1920s from Einzig (1937) and the non-linear model described by eq. (7), Paya et al (2010) find that deviations from CIP bias $\hat{\beta}$ toward zero.

$$s_{t+1} - f_t = A + B(f_t - s_t)e^{K(d)^2} + e_{t+1}$$

(7)

They summarize their results as follows:

We examined data for the interwar period for the dollar-sterling exchange rate and found that the degree of bias in the standard Fama regression varies significantly with the deviation from covered interest parity. When deviations are large, the degree of bias is much smaller than implied by the standard Fama regression. Paya et al (2010, 57)
Using my data and eqs. (8) and (9) I replicate their results in Table 5.\footnote{They use the forward premium from Einzig to calculate deviations from CIP. I use the same premia and Einzig's spot rates to retrieve forward rates.}

\begin{equation}
\Delta s_{t+1} = a_1 + b_1(i_t - i_t^*)e^{k_1(d_t)} + v_{t+1}
\end{equation}

(8)

\begin{equation}
\Delta s_{t+1} = a_1 + \beta_1(f_t - s_t)e^{k_2(d_t)} + v_{t+1}
\end{equation}

(9)

Where \( |d_t| \) is the absolute value of \( d_t \) rather than \( d_t \) squared as in Paya et al (2010).

Table 5
Effects of Transaction Costs

<table>
<thead>
<tr>
<th></th>
<th>( \Delta s_{t+1} = a_1 + b_1(i_t - i_t^*)e^{k_1(d_t)} + v_{t+1} )</th>
<th>( \Delta s_{t+1} = a_1 + \beta_1(f_t - s_t)e^{k_2(d_t)} + v_{t+1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{a}_1 )</td>
<td>( \hat{b}_1 )</td>
<td>( \hat{\beta}_1 )</td>
</tr>
<tr>
<td>U.S.-U.K.: 1922-1925</td>
<td>2.16**</td>
<td>-9.60**</td>
</tr>
<tr>
<td>Flexible</td>
<td>(0.20)</td>
<td>(1.38)</td>
</tr>
<tr>
<td>U.S.-Canada: 1961-</td>
<td>0.26**</td>
<td>-2.24**</td>
</tr>
<tr>
<td>1969 (Pegged)</td>
<td>(0.04)</td>
<td>(0.81)</td>
</tr>
<tr>
<td>U.S.-Canada: 1970-</td>
<td>-0.23</td>
<td>0.67</td>
</tr>
<tr>
<td>1973 (Flexible)</td>
<td>(0.12)</td>
<td>(16.93)</td>
</tr>
<tr>
<td>(Flexible)</td>
<td>(0.06)</td>
<td>(4.91)</td>
</tr>
<tr>
<td>( \hat{a}_2 )</td>
<td>( \hat{b}_2 )</td>
<td>( \hat{\beta}_2 )</td>
</tr>
<tr>
<td>Flexible</td>
<td>(0.236)</td>
<td>(1.65)</td>
</tr>
<tr>
<td>U.S.-Canada: 1961-</td>
<td>0.28**</td>
<td>-1.77**</td>
</tr>
<tr>
<td>1969 (Pegged)</td>
<td>(0.05)</td>
<td>(0.47)</td>
</tr>
<tr>
<td>U.S.-Canada: 1970-</td>
<td>-0.19</td>
<td>1.21</td>
</tr>
<tr>
<td>1973 (Flexible)</td>
<td>(0.10)</td>
<td>(1.15)</td>
</tr>
<tr>
<td>(Flexible)</td>
<td>(0.06)</td>
<td>(4.56)</td>
</tr>
<tr>
<td>( \hat{R}^2/DW )</td>
<td>( \hat{R}^2/DW )</td>
<td>( \hat{R}^2/DW )</td>
</tr>
<tr>
<td>U.S.-U.K.: 1922-1925</td>
<td>0.307</td>
<td>0.232</td>
</tr>
<tr>
<td>Flexible</td>
<td>(1.14)</td>
<td>(1.65)</td>
</tr>
<tr>
<td>U.S.-Canada: 1961-</td>
<td>0.008</td>
<td>0.014</td>
</tr>
<tr>
<td>1969 (Pegged)</td>
<td>(0.05)</td>
<td>(0.47)</td>
</tr>
<tr>
<td>U.S.-Canada: 1970-</td>
<td>0.000</td>
<td>-7.69</td>
</tr>
<tr>
<td>1973 (Flexible)</td>
<td>(0.10)</td>
<td>(1.41)</td>
</tr>
<tr>
<td>(Flexible)</td>
<td>(0.06)</td>
<td>(4.56)</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. * Significant at 5%. ** Significant at 1%.

Five of the eight \( \hat{k}_i \) in Table 5 are negative and significant. The one positive \( \hat{k}_i \) is insignificant. \( \hat{b}_i \) represent what \( \hat{b} \) would be if \( d_t \) were zero. Three of the four \( \hat{b} \) in Table 1 are closer to zero than corresponding \( \hat{b}_i \). \( \hat{\beta}_i \) represent what \( \hat{\beta} \) would be if \( d_t \) were zero. All \( \hat{\beta} \) in Table 1 are closer to zero than corresponding \( \hat{\beta}_i \). Deviations from CIP appear to bias \( \hat{b} \) and \( \hat{\beta} \) toward zero. For the Balke-Wohar data, \( d_t \) represent primarily transaction costs. For that data, \( \hat{k}_j \) and \( \hat{k}_2 \) are both significant at 1%.

5. Monetary Policy and UIP.

This section provides an economic explanation for the failure of uncovered interest parity. That explanation also helps explain the forward-bias puzzle and related puzzles. Monetary policy is the
culprit. That idea is not new, but this is the most detailed explanation of how monetary policy can produce negative $\hat{b}$. In addition, to the best of my knowledge, with the exception of Chinn and Meredith (2004), it is also the first time monetary policy has been used to explain the related puzzles.

McCallum (1994, 123) apparently was the first to suggest that monetary policy is the culprit. He uses a simple model to show how a combination of stabilizing interest rates and leaning against the wind can produce negative $\hat{b}$. Chinn and Meredith (2004, 420-23) follow up on McCallum's suggestions and build a "richer" model that produces the maturity puzzle. Other work suggesting that monetary policy is relevant includes Anker (1999), Kirkos (2002) and Lafuente and Ruiz (2006) while Mark and Moh (2007) show that intervention intensifies the forward premium.

In addition a substantial literature uses vector auto-regression to analyze how monetary shocks affect UIP. It includes Eichenbaum and Evans (1995), Grilli and Roubini (1996), Cushman and Zha (1997), Kim and Roubini (2000), Faust and Rogers (2003), Scholl and Uhlig (2008), Bjørnland (2009), Bouakez and Normandin (2010) and Heinlein and Krolzig (2012). Most find that monetary shocks create deviations from UIP.

Building on this theoretical and empirical work, I show how three common forms of monetary policy affect $\hat{b}$: (1) interest rate stabilization, (2) exchange rate stabilization and (3) macro stabilization. (2) and (3) can produce negative $\hat{b}$. (1) primarily affects the absolute value of $\hat{b}$.

5.1. The Model.

The model in Table 6 is not intended to be realistic. It is meant only to show how monetary policy affects $\hat{b}$. Even though the Bank of England rather than the Fed normally intervenes in the foreign exchange market, I discuss official intervention as though the Fed intervened.

Eq. (1) in Table 6 describes CIP. Like the rest of the model it ignores transaction costs. Their effects are discussed later.
\begin{align*}
\Delta_{t+1} = \pi_{t+1} &= C(B\bar{t}_t + Du_t - FX\Delta s_t) + x_{t+1} \quad \text{(VI)} \\
E(\pi_{t+1} | C_e_t) &= C(B\bar{t}_t + Du_t - FX\Delta s_t) \quad \text{(VII)}
\end{align*}

Additional definitions:

- \( u_t \): Actual rate of unemployment minus natural rate.
- \( u_t = Uu_{t-1} + w_t \) \( 0 \leq U < 1 \)

Additional restrictions: \( D, X, \Lambda, \alpha \) and \( h \) are all \( \geq 0 \). \( \lambda \) is positive but less than 1.0 while \( 0 \leq F \leq 1 \). \( w_t, x_t \) and \( \bar{r}_t \) are orthogonal white noise. \( Z = \alpha(1+F) \) and \( H = h(1-F) \).

Equations (II) and (III) describe the foreign exchange markets. CIP dictates their form. They take no position on the issue of whether spot markets dominate forward markets or the reverse. \( f_t \) depends on \( \lambda \bar{t}_t \) and \( s_t \) depends on \( (1-\lambda) \bar{t}_t \). The larger \( \lambda \), the more \( s_t \) dominates \( f_t \). Later sensitivity tests suggest that \( \hat{\beta} \) is not very sensitive to \( \lambda \).

Central banks often intervene directly in spot exchange markets. \( ZX\Delta s_t \) in eqs. (II) and (III) captures the most common form of intervention, leaning against the wind. Neely (2001, 4-6) reports that almost 90% percent of central banks sometimes or always lean against the wind and that 40% fully sterilize while only 30% never sterilize.

\( X \) describes the policy response to a given \( \Delta s_t \). The Fed leans against the wind by selling sterling (buying dollars) as the price of sterling rises. \( Z \) describes how that intervention affects \( s_t \).

Both \( Z \) and the change in the monetary base caused by a given \( X \Delta s_t \), depend on sterilization. The Fed sterilizes by selling (buying) U.S. T bills as it buys (sells) sterling. With complete
sterilization, intervention does not affect the monetary base and $F$ is zero. With no sterilization, $F$ is one and each dollar's worth of sterling that the Fed sells reduces the monetary base by one dollar.

There is a general consensus that sterilization reduces $Z$. There is a less general consensus that $Z$ is positive even when sterilization is complete.\(^7\) Table 6 assumes that $Z$ is positive even when $F$ is zero.

$\bar{P}_t$ describes the relative price levels and captures the influence of purchasing power parity.

There are several versions of purchasing power parity, but the most common is based on arbitrage and the law of one price (LOP).\(^8\) If one takes the role of arbitrage and the LOP seriously, then the prices in $\bar{P}_t$ should not be retail because arbitrage is not possible at the retail level. If one grocery store sells seedless red grapes at $0.99 per pound and another store across the street sells them for $2.00 a pound, an arbitrager cannot buy them at $0.99 and sell them for $2.00. Arbitrage also is not possible in most wholesale markets. Brand names and marketing contracts usually prevent arbitrage. But there are exceptions like wholesale markets for fresh fish and fresh flowers. It is possible to buy live lobsters at the wholesale market in Boston on Monday and simultaneously contract to sell them in LA later that week. But that is not true for frozen lobster or frozen fish where there are brand names and no similar wholesale markets.

I call the version of purchasing power parity that takes arbitrage and the law of one price seriously the arbitrage version of purchasing power parity or APPP. In APPP all prices are from markets where true arbitrage is possible.\(^9\) Although the number of forward commodity markets is small relative to the number of retail or wholesale markets, the absolute number is still large. It includes a wide range of petroleum products, metals and agricultural products.

---

\(^7\) Phillips and Pippenger (1993) find evidence of a temporary effect for intervention. Although she attributes it to a signaling effect, in her early review of the literature Edison (1993, 54) refers to a temporary effect. In their review of the later literature, Beine et al. (2003,892) also mention a temporary effect. More recently Fatum (2008) finds evidence of a temporary effect. Pippenger and Phillips (1973) probably provide the best estimate of $Z$ because they are able to avoid simultaneity by assuming that, in the absence of intervention, the exchange rate is a random walk; an assumption that Pippenger (2008a) supports.

\(^8\) See Rogoff (1996).

\(^9\) For some evidence that arbitrage is effective in such markets see Phillips and Pippenger (2008) and Pippenger (2008b).
Of course, if you prefer to interpret $\bar{P}_t$ as describing conventional PPP, you can do so.

If both central banks target their price level, they stabilizes $\bar{P}_t$ and the exchange rate. If either central bank targets inflation, $\bar{P}_t$ has a unit root because it is the sum of random shocks as at least one central bank allows random changes in the price level to accumulate.

Eq. (IV) describes how monetary policy affects liquidity. A positive $B$ implies interest stabilization, which reduces $\sigma_t^2$. That policy increases (decreases) liquidity when $\bar{r}_t$ is positive (negative). A positive $D$ implies macro-stabilization. The Fed buys T bills when unemployment is above the 'natural rate' and sells T bills when unemployment is below the natural rate. The first increases liquidity the second reduces it. A positive $X$ implies intervention. With full sterilization, intervention does not affect liquidity. Without full sterilization, selling (buying) sterling reduces (increases) liquidity.

Eq. (V) describes $\bar{r}_t$. As long as $B$, $D$ and $X$ are all zero, $\bar{r}_t$ equals the difference in expected inflation plus $\bar{r}_t$. $\Lambda$ describes how liquidity affects $\bar{r}_t$. Other things equal, the more liquid are households and firms the lower is $\bar{r}_t$. Liquidity effects disappear as firms and households adjust their portfolios and prices change. Liquidity effects presumably decline as maturity increases.

Eq. (VI) describes actual inflation. It depends on $x_{t+1}$ and how policy affects the monetary base.

Eq. (VII) describes expected inflation. With rational expectations, expected inflation equals actual inflation minus the white noise $x_{t+1}$. $C$ depends on the time horizon; the shorter the time horizon the smaller $C$. $C$ also depends on the monetary regime. In a highly stable monetary regime like post WWII Germany $C$ is small. In an inflationary regime like Germany in the early 1920s $C$ is large.
5.2. Monetary policy and \( \hat{b} \).

The Appendix shows how all three policies interact to determine \( \hat{b} \). To provide some economic insight, I discuss each policy separately.

5.2.1. No stabilization. For the model in Table 6 to be useful it should produce sensible results in the absence of monetary policy. Eq. (10) describes \( \hat{b} \) when \( B, D \) and \( X \) are all zero.

\[
\hat{b} = (\lambda - 1)(R - 1) \quad (10)
\]

\( \hat{b} \) is positive but less than one.

5.2.2. Stabilizing interest rates. Eq. (11) shows the effects of just stabilizing interest rates.

\[
\hat{b} = \{CB + (\lambda - 1)(R - 1)\} \frac{\sigma^2}{\sigma^2_r} = CB + (\lambda - 1)(R - 1) \quad (11)
\]

Just stabilizing interest rates increases \( \hat{b} \).

5.2.3. Macro stabilization. Eq. (12) describes the solution for \( \hat{b} \) when just \( D \) is positive.

\[
\hat{b} = \{[C + (C - \Lambda)(\lambda - 1)(U - 1)](C - \Lambda)D^2\sigma^2_u + [(\lambda - 1)(R - 1)]\sigma^2_r\}/\{(C - \Lambda)^2D^2\sigma^2_u + \sigma^2_r\} \quad (12)
\]

If \( C - \Lambda \) positive, then \( \hat{b} \) is positive. But if liquidity effects dominate inflationary effects, which is likely under a stable monetary regime, then \( C - \Lambda \) is negative. In that case \( \hat{b} \) is negative when \( C \) is greater than \( |(C - \Lambda)(\lambda - 1)(U - 1)| \) and \(|[C + (C - \Lambda)(\lambda - 1)(U - 1)](C - \Lambda)D^2\sigma^2_u| \) is greater than \( (\lambda - 1)(U - 1)\sigma^2_r \).

Eq. (13) shows how macro stabilization and interest rate stabilization interact.

\[
\hat{b} = \left\{ \{C(C - \Lambda)[1 + B(\Lambda - C)] + B\Theta(C - \Lambda) + (\lambda - 1)(C - \Lambda)(U - 1)\}D^2\sigma^2_u + \{CB + (\lambda - 1)(R - 1)\}\sigma^2_r \right\} / \{(C - \Lambda)^2D^2\sigma^2_u + \sigma^2_r\} \quad (13)
\]

\( d\hat{b}/dB \) depends on the values of various parameters, particularly \( C \) and \( \Lambda \). As the simulations below show, ‘reasonable’ values for the various terms produce negative \( \hat{b} \) and negative \( d\hat{b}/dB \).
The economics is simpler than the math. Start in equilibrium with \( s_t, f_t, \bar{t}_t \) and \( u_t \) all zero. Then the unemployment rate rises and the Fed buys T bills. If that purchase lowers \( \bar{t}_t \) and creates inflation, and that inflation produces a positive \( \Delta s_{t+1} \), then macro stabilization produces a negative \( \sigma_{\Delta s_{t+1}} \). As long as stabilizing interest rates reduces \( \sigma^2 \) by more than it reduces \( |\sigma_{\Delta s_{t+1}}| \), it increases \( |b| \).

5.2.4. Exchange rate stabilization. Eq. (14) describes \( \hat{b} \) when the Fed just leans against the wind.

\[
\hat{b} = \left\{ \begin{array}{l}
\Pi(\lambda-1)\{1-\Phi\} \{1+Z\Pi(\lambda-1)\Omega\}\{1+Z\Pi(\lambda-1)(1-\Phi R)\} \\
+\Omega\pi^2\{1-\Phi^2\}\sigma^2_t \\
+\Omega\pi^2\pi^2\{1-\Phi^2\}\sigma^2_t \\
+\Omega\pi^2\pi^2\{1-\Phi^2\}\sigma^2_t \\
\end{array} \right. 
\]

(14)

Where \( \Omega = [(\lambda-C)FX-HX], \pi = 1/(1+Z\Pi(\lambda-1)\Omega) \) and \( \Phi = [ZX-CFX-(\lambda-1)\Omega]/[1+Z(\lambda-1)\Omega] \).

To simplify the discussion of eq. (14), and to be consistent with Neely (2001), I assume that intervention is fully sterilized and \( F \) is zero. With that assumption, \( \Omega \) is negative and \( \Phi \) equals \( [ZX-(\lambda-1)\Omega]/\Pi \). Since \( Xs_t \) is only a very small fraction of the stock of domestic and foreign T bills, \( \Omega \) should be close to zero. An \( \Omega \) close to zero implies that \( \Pi \) is positive even when \( Z \) is zero.

Since the coefficient for \( \sigma^2_t \) in the numerator is positive, for \( \hat{b} \) to be negative \( \Phi \) must be positive. For \( \Phi \) to be positive when \( \Pi \) is positive, \( ZX \) must be greater than \( (\lambda-1)\Omega \), which is very small because \( (\lambda-1) \) is less than one and \( \Omega \) is close to zero.

To see the economics behind how leaning against the wind produces a negative \( \hat{b} \), consider the following mental experiment: \( r_t \) and \( x_t \) have been zero long enough for the system to achieve equilibrium. Then a positive \( x_t \) produces a positive \( \Delta s_t \) and the Fed leans against the wind by selling sterling today as the rate rises. To fully sterilize that sale, the Fed buys an equivalent amount of domestic T bills. The monetary base is unchanged, but that purchase reduces the stock of U.S. T bills whose prices rise and rates fall. To maintain its sterling deposits at the Bank of England, the Fed also
sells sterling T bills. That sale increases the stock of sterling bills whose price falls and rates rise.

Those transactions produce a negative $\Delta s_t$.

With $\Delta s_t$ negative, a negative $\hat{b}$ requires a positive $\Delta s_{t+1}$. Phillips and Pippenger (1993), Pippenger (2003) and Fatum (2008) find that leaning against the wind produces a positive $\Delta s_{t+1}$ because leaning against the wind introduces positive autocorrelation into $\Delta s_t$. Of course leaning against the wind is not sufficient to produce negative $\hat{b}$. That requires that the effects of leaning against the wind dominate. In eq. (14) that domination requires that $|\Omega \Pi^2 \{\Phi/(1-\Phi^2)\} \sigma_\tau^2|$ is larger than

$$\left[\Pi(\lambda-1)\{R^{-\circ}(1-\Phi)/(1-\Phi R)\} + IT^2 \Omega(\lambda-1)^2 \{(1-R)(\Phi-I-R-\Phi R)/(1-\Phi R)(1+\Phi)\}\right] \sigma_\tau^2$$

Eq. (15) describes $\hat{b}$ when $X$ and $B$ are both positive.

$$\hat{b} = \left\{\left[CB \Theta^2 \Pi \{1/(1-\Phi R)\} + \Theta^3 \Pi(\lambda-1)\{R-[1-\Phi)/(1-\Phi R)\}\right] + \Omega(CB) \Theta^3 \{(R+\Phi)/(1-\Phi R)\} + \Omega CB \Theta^3 \Pi^2(\lambda-1)\{R^2(1+\Phi)-R\}/[(1-\Phi R)(1+\Phi)]\right\} \sigma_\tau^2$$

$$\left.+\Omega \Theta \Pi^2 \{(1-\Phi^2)\} \sigma_\tau^2\right\} \left\{\left[\Theta^2 + 2\Theta^3 \Pi \Omega(\lambda-1)\{1/(1-\Phi)\} + 2\Theta^3 \Pi \Omega \{1/(1-\Phi R)\}\right] \sigma_\tau^2\right\} \left\{\left[1/(1-\Phi^2)\right] \sigma_\tau^2\right\} \left\{\left[1/(1-\Phi R)\right] \sigma_\tau^2\right\}$$

Where $\Omega$ equals $(1-A)FX-HX$, $\Theta$ equals $1/[1+B(1+A-C)]$, $\Pi$ equals $1/(1+Z-\Theta (\lambda-1)\Omega)$ and $\Phi$ equals $[ZX+\Theta CB \Omega-CFX-\Theta(\lambda-1)\Omega]/[1+ZX-\Theta (\lambda-1)\Omega]$. Given a positive $\Phi$, the condition for a negative $\hat{b}$ is that $|\Omega \Pi^2 \{\Phi/(1-\Phi^2)\} \sigma_\tau^2|$ is greater than

$$\left\{\left[CB \Theta^2 \Pi \{1/(1-\Phi R)\} + \Theta^3 \Pi(\lambda-1)\{R-[1-\Phi)/(1-\Phi R)\}\right] + \Omega(CB) \Theta^3 \{(R+\Phi)/(1-\Phi R)\} + \Omega CB \Theta^3 \Pi^2(\lambda-1)\{R^2(1+\Phi)-R\}/[(1-\Phi R)(1+\Phi)]\right\} \sigma_\tau^2$$

$$\left.+\Omega CB \Theta^3 \Pi^2(\lambda-1)\{1/(1-\Phi R)\} + \Omega \Theta \Pi^2(\lambda-1)^2 \{(1-\Phi R)/(1-\Phi R)\} \right\} \sigma_\tau^2$$
Once again that condition depends on the parameters as well as $\sigma_x^2$ and $\sigma_t^2$. As the next subsection shows, ‘reasonable’ values produce negative $\hat{h}$ and negative $d\hat{h}/dB$.

5.2.5. Interaction and sensitivity. Testing for interaction and sensitivity requires imposing some structure on the model. Unfortunately most of the terms in Table 6 are not easily estimated and attempting to do so would exceed the objectives of this paper. But a few terms are easily identifiable. FRED provides monthly data for the U.S unemployment rate, the exchange rate and nominal interest rate differential between the U.S and U.K. over the interval used by Balke and Wohar (1998).

<table>
<thead>
<tr>
<th>Table 7</th>
<th>Interaction and Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>$\sigma_x^2$</td>
</tr>
<tr>
<td>20</td>
<td>0.0021</td>
</tr>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>1.0*</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>

*Actually 0.99999 because $R$ equal to 1.0 implies that $\sigma_t^2$ is infinite.

Using that information, Table 7 assume that $\sigma_x^2$ equals 0.00018, $U$ equals 0.95, $\sigma_t^2$ equals 0.000063 and $\sigma_{\Delta s}^2$ equals 0.0013.
The first row in Table 7 lists the 12 remaining ‘free’ parameters. The second row describes their base line values. These values are not unique. Others could choose equally plausible values. But the objective of Table 7 is only to show that reasonable values can explain negative $\hat{b}$ and to provide some insight into how sensitive $\hat{b}$ might be to those values. The last column of the second row reports the base line $\hat{b}$ implied by the base line values, −1.36.

Reducing $B$ from 20 to zero while leaving all other parameters unchanged as indicated by blank boxes, increases $\hat{b}$ to −0.46. Doubling $B$ to 40 reduces $\hat{b}$ to −2.25. This supports McCallum (1994).

Reducing $\sigma_i$, $h$, $X$ and $a$ one at a time to 0.0 produces positive $\hat{b}$. Increasing $R$ from 0.9 to 0.99999 also produces a positive $\hat{b}$. Sterilization is crucial. Initially $F$ is zero. Partial sterilization, $F$ equal to 0.5, produces a $\hat{b}$ equal to 0.2. Full sterilization, $F$ equal to 1.0, produces a $\hat{b}$ equal to 2.4 even though macro stabilization is unchanged. As expected, $\hat{b}$ is sensitive to $C$. With $C$ equal to 0.0 and $\Lambda$ equal to 0.11, $\hat{b}$ equals −1.62. But a $C$ equal to 1.0 with $\Lambda$ equal to 0.11 produces a $\hat{b}$ equal to 9.29.

Again I want to emphasized that Table 6 only shows that reasonable values can produce negative $\hat{b}$. That does not mean that the actual values do so. Determining whether or not that is the case requires research far beyond the objective of this paper.

5.3. Forward bias.

The liquidity effect of responding to high unemployment tends to lower $\bar{T}_t$ while the inflationary effect tends to raise $\Delta s_{t+1}$. A fully sterilized response to a rise in current exchange rates tends to introduce positive serial correlation into $\Delta s_t$ while the associated changes in the relative stock of T bills reduce $\bar{T}_t$. These responses to monetary policy do not depend on CIP. I could drop CIP from Table 6 and the economic explanation for negative $\hat{b}$ would not change.
But without CIP it would be difficult to explain the failure of UIP. As long as CIP holds, the $\hat{\beta}$ implied by eq. (5) equals $\hat{b}$ because $(f_t - s_t)$ equals $\tilde{r}_t$. This role for CIP is important because it helps explain some of the related puzzles, particularly the commodity puzzle.

6. Related puzzles.

A viable explanation for the failure of UIP and the forward-bias puzzle should also explain, or at least be consistent with, the related puzzles. Otherwise it is suspect.

6.1. Commodity puzzle. With flexible exchange rates between the U.S and other developed countries, for short maturities $\hat{\beta}$ are routinely negative. That is not the case in commodity markets. CIP provides a simple explanation for the Commodity Puzzle.

When monetary policy produces negative $\hat{b}$, CIP transmits that negativity to the $\hat{\beta}$ in foreign exchange markets. The $\hat{\beta}$ in commodity markets are not routinely negative because nothing similar transmits that negativity to commodity $\hat{\beta}$.

6.2. Carry Trade.

Monetary policies produce the expected returns that motivate the carry trade. Commercial banks and other institutions with relatively low transaction costs take positions based on these expected returns. Those positions create risk premiums. When the Fed leans against the wind and reduces short-run movements in exchange rates it reduces the risk and encourages the carry trade.

As an example, start in equilibrium with $\tilde{r}_t$ and the expected change in the exchange rate zero. Then the Fed responds to a rise in unemployment by buying T bills. The liquidity effect dominates the inflationary effect and $\tilde{r}_t$ falls. The increase in the monetary base puts upward pressure on spot exchange rates. Even if expected future spots rate do not rise, there is still an expected return from borrowing where interest rates are low and lending where they are high. If the Fed also leans against
the wind, it reduces the possibility that exchange rates might fall and offset the negative $\tilde{r}_n$, which reduces the risk associated with the carry trade.

6.3. The Development Puzzle.

The average $\hat{\beta}$ between developed countries in Table 1 in Frankel and Poonawala (2010) is $-4.3$ while the average $\hat{\beta}$ between the U.S. and developing countries in Table 2 is $0.003$; suggesting that the forward exchange rate is a less biased indicator of the future expected spot rate in emerging market economies. While providing no explanation, Frankel and Poonawala point out that the source of the forward discount bias probably does not lie entirely in the exchange risk premium.

Table 6 and the earlier discussion suggest three related explanations for the Development Puzzle: (1) transaction costs are usually higher in developing than developed markets at least partly because volume is usually lower, (2) central banks in developed and developing countries follow different monetary policies and (3) there is more government intervention in developing than in developed markets, particularly in foreign exchange markets.

With respect to transaction costs, transaction costs, particularly those associated with CIP, tend to bias $\hat{\beta}$ toward zero. That bias helps explain why the average absolute value of $\hat{\beta}$ in Frankel and Poonawala's Table 1 is 4.6 while it is only 0.67 in their Table 2.

With respect to monetary policy, while there is no easy way to compare how central banks in developing and developed countries respond to macro issues like unemployment, the IMF’s annual publication *Exchange Arrangements and Exchange Restrictions* provides an easy way to compare foreign exchange policies. Table 8 uses that publication for 1996 and 2004 to describe the foreign exchange policies of developed and developing countries. To simplify the table, and avoid any potential double counting, all countries that initially adopted the euro and were either part of the ERM or floated independently before that adoption are lumped together under the euro.
The 2000 edition defines an independent float as follows:

The exchange rate is market determined, with any foreign exchange intervention aimed at moderating the rate of change and preventing undue fluctuations in the exchange rate.

It appears that, except for Greece, all of the developed countries in Frankel and Poonawala (2010) floated against the dollar from 1996 to 2004. It might not look that way in Table 8, but Norway did so. The 1996 edition describes the managed float for Norway as follows: "The exchange rate of the Norwegian krone is determined by market forces, however, the Bank of Norway (BN) intervenes to maintain stability."

Greece appears to be the only developed country in Frankel and Poonawala (2010) that did not float against the U.S. dollar for the entire period. It is also the only developed country with a positive $\hat{\beta}$. The 1996 edition describes Greece's managed float very differently:

The exchange rate for the drachma is determined in daily fixing sessions in which the Bank of Greece (BOG) and authorized commercial banks participate. Greece is a member of the EMS. The drachma is included in the ECU basket, but Greece does not participate in the ERM of the EMS.

Only four developing countries floated against the dollar for the entire period: India, Mexico, the Philippines and South Africa. Three of those four $\hat{\beta}$ are negative as is the average at $-0.71$. 

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>IF</td>
<td>IF</td>
<td>Greece</td>
<td>MF</td>
<td>Euro</td>
<td>Sweden</td>
<td>IF</td>
<td>IF</td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>IF</td>
<td>IF</td>
<td>Japan</td>
<td>IF</td>
<td>IF</td>
<td>Switzerland</td>
<td>IF</td>
<td>IF</td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>ERM</td>
<td>ERM II</td>
<td>New Zealand</td>
<td>IF</td>
<td>IF</td>
<td>U.K.</td>
<td>IF</td>
<td>IF</td>
<td></td>
</tr>
<tr>
<td>Euro</td>
<td>IF</td>
<td>IF</td>
<td>Norway</td>
<td>MF</td>
<td>IF</td>
<td>U.S.</td>
<td>IF</td>
<td>IF</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Czech Rep.</td>
<td>P</td>
<td>MF</td>
<td>Kuwait</td>
<td>PB</td>
<td>P</td>
<td>South Africa</td>
<td>IF</td>
<td>IF</td>
<td></td>
</tr>
<tr>
<td>Hong Kong</td>
<td>CB</td>
<td>CB</td>
<td>Mexico</td>
<td>IF</td>
<td>IF</td>
<td>Thailand</td>
<td>P</td>
<td>MF</td>
<td></td>
</tr>
<tr>
<td>Hungry</td>
<td>CP</td>
<td>P</td>
<td>Philippines</td>
<td>IF</td>
<td>IF</td>
<td>Turkey</td>
<td>MF</td>
<td>IF</td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>IF</td>
<td>IF</td>
<td>Saudi Arabia</td>
<td>P</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indonesia</td>
<td>MF</td>
<td>MF</td>
<td>Singapore</td>
<td>PB</td>
<td>MF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fixed (F); Independent Float (IF); Managed Float (MF); Pegged (P); Pegged to Basket (PB); Crawling Peg (CP) Target Band (TB); Currency Board (CB). #Adopted euro January 2001.
With respect to forward exchange markets, the 2000 edition of *Exchange Arrangements and Exchange Restrictions* indicates that forward markets for developed countries operated freely. That is not the case for the developing countries. Almost all developing countries interfered in forward markets. The 2000 edition describes the forward markets for the four developing countries with an independent float as follows:

India (first paragraph) "Ads (authorized dealers) are allowed to deal forward in any permitted currency. Forward purchases or sales of foreign currencies against rupees with banks abroad are prohibited. The RBI may enter into swap transactions with Ads, under which it buys or sells spot dollars and sells or buys forward dollars for maturities available in the market."

Mexico "Commercial banks are allowed to enter into forward transactions that involve currencies in over-the-counter markets and formal markets recognized by the BOM (currently, the Chicago Mercantile Exchange, the Chicago Board Options Exchange, the Mid America Commodity Exchange, the Mexican Derivatives Exchange, and the Commodity Exchange Inc.)"

Philippines (first sentence) "All forward transactions to purchase foreign exchange from nonresidents, including renewals thereof, require prior clearance by the BSP."

South Africa (first sentence): "Subject to certain limitations, authorized dealers are permitted to conduct forward exchange operations, including cover for transactions by nonresidents."

Mexico and South Africa have the least restricted forward markets and both have negative $\beta$.

Monetary policies with respect to foreign exchange differ between developed and developing countries. Determining whether or not those and other policy differences fully explain the Development Puzzle requires research beyond the objectives of this paper.

6.4. Levels versus Differences.

The model in Table 6 provides a simple explanation for this puzzle. CIP implies that

$$\Delta f_{t+1} = \Delta s_{t+1} + \Delta \tau_{t+1},$$

which implies that $f_{t+1} = f_t + \Delta s_{t+1} + \Delta \tau_{t+1}$. Therefore $s_{t+1} = f_{t+1} - \tau_{t+1}$ can be written as

$$s_{t+1} = f_t + \Delta s_{t+1} - \tau_{t+1},$$

which produces the following test equation where CIP implies that $g_1$ equals one:
\[ s_{t+1} = g_0 + g_1 f_t + \Delta s_{t+1} - \bar{t}_t \]  

(16)

As long as \( s_t \) and \( f_t \) have unit roots, and \( \Delta s_{t+1} \) and \( \bar{t}_t \) are stationary, super consistency implies that \( \hat{g}_t \) approaches one rapidly. \( \hat{\beta} \) are negative for changes because CIP implies that \( \hat{\beta} \) equal \( \hat{\beta} \).

6.5. Maturity.

Using a simple model where the Fed follows a Taylor rule, Chinn and Meredith (2004) show that monetary policy can explain the maturity puzzle. The model in Table 6 provides an economic explanation for their finding. As maturity increases, leaning against the wind affects exchange rates less and less. Indeed the rationale for leaning against the wind is to moderate short-run movements without affecting longer run movements. As maturity increases liquidity effects decrease and inflationary effects increase. In addition, as maturity increases the effects of sterilization on relative interest rates decline because the sterilization uses T bills. Finally, as maturity increases the effects on longer maturities of stabilizing short-term interest rates decrease.

Eq. (17) shows the solution for \( \hat{\beta} \) as \( B, A, a, \) and \( H \) go to zero.

\[
\hat{\beta} = \frac{\{C^2 + C^2(\lambda - 1)(U - 1)\}D^2\sigma_u^2 + \{(\lambda - 1)(R - 1)\}\sigma_r^2}{\{C^2D^2\sigma_u^2 + \sigma_r^2\}} \]  

(17)

\( \hat{\beta} \) is positive. How positive depends on the parameters. As \( \sigma_r^2 \) goes to zero, \( \hat{\beta} \) goes to \( 1 + (\lambda - 1)(U - 1) \).

6.6. Time Dependency.

Table 6 describes a simple model in which the relevant parameters and variances are constant. As a result, monetary policy is constant and how the public responds to those policies is constant. But central banks change policies over time and the public responds to those changes. Those changes can produce time dependency.

For example, Table 6 assumes that \( F \) is constant while the discussion assumes that \( F \) is zero. While most banks routinely sterilize, the degree of sterilization can vary over time. In addition, central banks buy more T bills when \( u_t \) is +3% than they sell when it is −3%. They also lean against the wind
more when foreign exchange markets are ‘turbulent’ than when they are ‘calm’. That variation can produce time dependency.

But the fact that monetary policy can explain time dependency does not imply that it does so. Determining whether or not it does so requires research far beyond the objectives of this paper.

6.7. Alternative explanations.


Risk premiums and the failure of rational expectations are not consistent with the related puzzles. To be consistent, risk would have to be lower, or expectations more rational in developing markets. The same would have to be true for longer maturities and for commodity markets. None of which seems likely.

The other alternatives also do not seem to provide viable explanations for the related puzzles. For example, how can perpetual learning, infrequent portfolio decisions or adverse selection explain the maturity, development or commodity puzzles? If advocates of those and other alternatives want their alternative to be taken seriously then they will need to show how their alternative can explain the related puzzles. Until then, monetary policy appears to be the most consistent with all the evidence.

But that consistency is not the same as showing that monetary policy predicts, or at least 'postdicts', $\hat{b}$ and $\hat{\beta}$.
7. Summary and Conclusions.

Section 2 reviews the failure of UIP, the forward-bias puzzle and the six related puzzles. Section 3 reviews covered interest parity which plays a key role in Section 5. Section 4 presents some new relevant evidence. Section 5 shows that monetary policy can explain the failure of UIP and the forward-bias puzzle. Section 6 shows that monetary policy can also explain the six related puzzles.

As far as I am aware, monetary policy is the only explanation for the failure of UIP and the forward-bias puzzle that also can explain the six related puzzles. The two most common explanations for the failure of UIP and the forward-bias puzzle, risk premiums and the failure of rational expectations, cannot do so because they appear to be inconsistent with most of the six related puzzles.

But showing that monetary policy can explain all these puzzles is very different from showing that monetary policy does explain all these puzzles. Doing that is far beyond the objectives of this paper. That is left for others to either confirm or reject.
References


Peel, D., A., 2002. Covered Interest Rate Arbitrage in the Interwar Period and the Keynes-Einzig Conjecture, Journal of Money, Credit and Banking 34 51-75,


APPENDIX

\[\sigma_{\Delta_{\lambda_1}/\sigma_1}^2 = \left\{ \left[ \text{CPI}(C-\Lambda)[1+B(\Theta(C-\Lambda))\{1/(1-\Phi U)\} + \Theta^2 \Pi(\lambda-1)(C-\Lambda)^2 \{1-[(1-\Phi)/(1-\Phi U)]\} \right] \right. \]

\[+ \Omega^2 \Pi^2 \{1+B(\Theta(C-\Lambda))\} \{\Phi(1-\Phi)/(1-\Phi^2)\} \]

\[+ \Omega^2 C \Theta^2 \Pi^2 (\lambda-1)(C-\Lambda)[1+B(\Theta(C-\Lambda))] \{U^2(1+\Phi-\Phi)/[1/(1-\Phi U)(1+\Phi)]\} \]

\[+ \Omega^2 C \Theta^2 \Pi^2 (\lambda-1)(C-\Lambda)[1+B(\Theta(C-\Lambda))] \{1-U/[1/(1-\Phi U)(1+\Phi)]\} \]

\[+ \Omega^2 \Theta^2 \Pi^2 (\lambda-1)(C-\Lambda)^2 \{[1-U](\Phi-1-U)/[1/(1-\Phi U)(1+\Phi)]\} \]

\[+ \Omega^2 \Theta^2 \Pi^2 (\lambda-1)(C-\Lambda)^2 \{[1-U](\Phi-1-U)/[1/(1-\Phi U)(1+\Phi)]\} \right\} D^2 \sigma_2 \]

\[+ \left\{ \left[ \Theta^2 (C-\Lambda)^2 + 2 \Omega^2 \Theta^2 \Pi \Theta(C-\Lambda) \{U/(1-\Phi U)\} + 2 \Omega^2 \Theta^2 \Pi \lambda \lambda (C-\Lambda)^2 \{1-U/(1-\Phi U)\} \right] \right. \]

\[+ (\Omega \Theta \Pi C) \left[ 1+B(\Theta(C-\Lambda)) \right] \{1+\Phi U/[1/(1-\Phi U)(1-\Phi^2)]\} \]

\[+ \left[ \Theta^2 (C-\Lambda)^2 \{1+B(\Theta(C-\Lambda))\} \{1+\Phi U/[1/(1-\Phi^2)]\} \right] \right\} \frac{1}{\Omega \Theta \Pi} \left\{ \frac{1}{1/(1-\Phi^2)} \right\} \sigma_2^2 \]

Where \(1/[1+B(\Lambda-C)]\) equals \(\Theta\), \(\Omega\) equals \([\Lambda-C]FX-HX\), \(1/(1+ZX-\Theta(\lambda-1)\Omega)\) equals \(\Pi\), and \([ZX+\Theta CB\Omega-CFX-\Theta(\lambda-1)\Omega]/[1+ZX-\Theta(\lambda-1)\Omega] \right\} \right\} \Phi\).