

Understanding the Forward Premium Puzzle: A Microstructure Approach

Craig Burnside*, Martin Eichenbaum[†] and Sergio Rebelo[‡]

July 2007

Abstract

High-interest-rate currencies tend to appreciate relative to low-interest-rate currencies. We argue that adverse-selection problems between participants in foreign exchange markets can account for this ‘forward premium puzzle.’ The key feature of our model is that the adverse selection problem facing market makers is worse when, based on public information, a currency is expected to appreciate.

J.E.L. Classification: F31

Keywords: Uncovered interest parity, exchange rates, microstructure.

*Duke University and NBER.

[†]Northwestern University, NBER, and Federal Reserve Bank of Chicago.

[‡]Northwestern University, NBER, and CEPR.

1. Introduction

A classic challenge confronting macroeconomists is how to explain why low-interest-rate currencies tend to depreciate relative to high-interest-rate currencies. An alternative statement of the challenge is that currencies which are at a forward premium tend to depreciate. This empirical regularity, known as the ‘forward premium puzzle,’ represents an egregious deviation from uncovered interest parity. While great strides have been made in documenting the puzzle, very little progress has been made in explaining it.¹ Much of the literature on this puzzle shares two key features. First, the foreign exchange market is modeled as an idealized Walrasian market.² Second, the literature emphasizes risk-based explanations for the forward premium. The first feature is problematic because, the foreign exchange market is actually a decentralized, over-the-counter market in which market makers play a central role (see Lyons (2001) and Sarno and Taylor (2001)). The second feature is also problematic. While risk must surely play a role in exchange rate markets, it has been extremely difficult to tie deviations from uncovered interest parity to economically meaningful measures of risk.³

In this paper we approach the forward premium puzzle from a new angle. Specifically, we take seriously the notion that the foreign exchange market is not Walrasian in nature and that risk is perhaps not at the center of the puzzle. Our analysis emphasizes adverse-selection problems between market makers and traders. To isolate the role of adverse selection we work with a simple model that abstracts entirely from risk considerations.

Our model is based on the microstructure approach developed in Glosten and Milgrom (1985). We assume that spot exchange rates follow an exogenous stochastic process with empirically realistic time-series properties. Our goal is to study the circumstances under which adverse-selection considerations imply that forward premia comove negatively, in population, with changes in exchange rates. We could, of course, make the spot exchange rate endogenous. But doing so in a way that would yield an empirically plausible exchange rate process would greatly complicate the analysis without contributing to our objective of understanding the comovement of forward rates and spot exchange rates.⁴

¹See Engel (1996) for a survey of the early literature. For more recent contributions see Han, Hirshleifer, and Wang (2006), Bacchetta and Van Wincoop (2006) and the references therein.

²Recent exceptions include Corsetti, Dasgupta, Morris, and Shin (2004) and Bacchetta and Van Wincoop (2006).

³See Engel (1996) for a review of the literature up to the mid-1990s and Burnside (2007) for a critical review of recent risk-based explanations of the forward premium puzzle.

⁴In this sense our procedure is similar in spirit to the classic analysis of Hansen and Singleton (1982) who assume that consumption is an exogenous stochastic process and study the comovement between consumption

The basic structure of our model is as follows. Two types of risk-neutral traders enter into forward contracts with competitive, risk-neutral, market makers. Informed traders have more information about exchange rate movements than market makers.⁵ Uninformed traders have the same information as market makers. The uninformed traders follow a behavioral trading rule: they are more likely to buy (sell) the pound forward when, based on public information, the pound is expected to appreciate (depreciate).⁶ While this rule seems natural, we do not derive it from first principles. We assume the rule because it allows us to exposit, in a transparent way, the adverse-selection rationale for the forward premium puzzle.⁷

The presence of informed agents creates an adverse-selection problem for the market maker. When the market maker receives an order he does not know whether it comes from an informed or an uninformed trader. However, he can quote different prices for buy and sell orders and make these prices depend on whether he expects the pound to appreciate or depreciate.

Our main result is that adverse-selection considerations can account for the forward premium puzzle. To be precise, consider an econometrician who regresses the change in the exchange rate on the forward premium using data generated by our model. Denote by $\hat{\beta}$ the econometrician's estimate of the slope coefficient, β . Conditional on a regularity condition holding, the probability limit (plim) of $\hat{\beta}$ is negative. This result obtains whether the econometrician works with the bid forward rate (the rate at which traders can sell the pound forward to the market maker), the ask forward rate (the rate at which traders can buy the pound forward from the market maker), or the average of the ask and bid rates. Our regularity condition requires that agents' ability to forecast exchange rates based on public information be small relative to the private information available to informed traders. This regularity condition has an alternative interpretation: as long as it is difficult to forecast exchange rates using public information and there are informed traders that make positive expected profits then there *must* be a forward premium puzzle.

The key to our result is that the market maker's adverse-selection problem is worse when,

and asset prices.

⁵Equivalently, these traders can be thought of as being better at processing information than market makers. See Albuquerque, Bauer, and Schneider (2007) for a recent paper that stresses the importance of differences in investor sophistication for explaining the patterns of international equity flows.

⁶We refer to the local currency in our model as the pound because in Section 4 we use exchange rate data quoted in units of foreign currency per British pound.

⁷We conjecture that it is possible to generate this trading pattern in a model in which exporters have an incentive to hedge exchange rate risk and where the home currency appreciates when the demand for home exports rises.

based on public information, the pound is expected to depreciate. To see why, it is useful to focus on the ask forward rate. Suppose that, on the basis of public information, the pound is expected to depreciate. Then uninformed traders are likely to sell the pound forward. It follows that, if the market maker receives a buy order, he attaches a high probability that the order came from an informed trader who expects the pound to appreciate. Consequently the market maker quotes a high price for the buy order, that is, a high ask forward exchange rate. The forward premium (evaluated at the ask rate) is, on average, high when the pound depreciates. So the model captures the negative correlation that defines the forward premium puzzle.

While the forward premium puzzle is a pervasive phenomenon, it is not uniformly present in the data. Bansal and Dahlquist (2000) show that, in a cross section of countries, estimates of the slope coefficient β are positively related to the average rate of inflation. Our model is consistent with this observation. Suppose that high-inflation countries experience persistent currency depreciations, and that public information is more useful for predicting exchange rates in these countries. When this predictability is sufficiently high, our regularity condition fails, and the plim of $\hat{\beta}$ becomes one. We provide evidence, complementary to that in Bansal and Dahlquist (2000), which is consistent with this property of our model.

In addition to providing an explanation of the forward premium puzzle, our model accounts for two other features of the data that are not obviously related to this puzzle. First, it is well-known that the current spot exchange rate is a better forecaster of the future spot exchange rate than the forward rate. We show that this property always holds when there is a forward premium puzzle. Since our model accounts for the forward premium puzzle it is consistent with this property. Second, we show that estimates of β are negatively related to the volatility of the forward premium, both in the data and in the model.

We conclude by addressing three potential concerns about our explanation of the forward premium puzzle. The first concern is that we require volatility in bid-ask spreads that are much higher than those observed in the data. In fact, our model generates forward bid-ask spreads that are, to a first-order approximation, constant. So, if anything our model understates the volatility of bid-ask spreads. The second concern is that our model requires very volatile forward premia. Given covered interest rate parity, this property would imply that our model generates movements in interest rate spreads that are counterfactually high. The third concern is that our model requires a large fraction of informed traders.

We address the last two concerns in an numerical example. We show that our model can account for the forward premium puzzle even though the volatility of the forward premium and interest rate spreads is roughly one sixth of that in the data. So adverse selection considerations can account for the forward premium puzzle, while adding very little volatility to interest rates. In addition, we show that the fraction of informed traders required to generate the forward premium puzzle is extremely small.

The remainder of this paper is organized as follows. We present our model in Section 2 and discuss its properties in Section 3. Section 4 contains a numerical example. Section 5 discusses the model's predictions for the existence and magnitude of the forward premium puzzle. Section 6 concludes.

2. The Model

In this section we display our model economy. We assume that the spot exchange rate follows an exogenous stochastic process. Forward rates are determined by the interaction between competitive market makers, informed, and uninformed traders. All agents are risk neutral.

2.1. Law of Motion of the Spot Exchange Rate

To simplify, we abstract from bid-ask spreads associated with spot exchange rates. The stochastic process for the growth rate of the spot exchange rate is given by:

$$\frac{S_{t+1} - S_t}{S_t} = \phi_t + \varepsilon_{t+1} + \omega_{t+1}. \quad (2.1)$$

Here S_t denotes the spot exchange rate expressed as foreign currency units per British pound.

The variable ϕ_t represents the change in the exchange rate that is predictable on the basis of time t public information.⁸ At the beginning of time t , all traders observe ϕ_t . For simplicity we assume that this variable is i.i.d. and obeys:

$$\phi_t = \begin{cases} \phi & \text{with probability } 1/2, \\ -\phi & \text{with probability } 1/2, \end{cases}$$

where $\phi > 0$.

The variable ε_{t+1} is not directly observed at time t , but, as we describe below, one group of traders receives advance signals about its value. This variable is i.i.d. and obeys:

$$\varepsilon_{t+1} = \begin{cases} \varepsilon & \text{with probability } 1/2, \\ -\varepsilon & \text{with probability } 1/2, \end{cases}$$

⁸One obvious source of public information is monetary policy. For example, a country that has persistently high monetary growth rates has predictably high rates of exchange rate depreciation.

where $\varepsilon > 0$.

Finally, none of the agents in the model has information at time t about the value of ω_{t+1} . The presence of this shock allows the model to generate exchange rate volatility that is not tied to either private or public information. The variable ω_{t+1} is i.i.d., mean zero, and has variance σ_ω^2 . The three shocks ϕ_t , ε_{t+1} , and ω_{t+1} are mutually orthogonal.

2.2. Traders and Market Makers

There is a continuum of traders with measure one. A fraction α of the traders are informed. At the beginning of time t , informed traders receive a signal $\zeta_t \in \{\varepsilon, -\varepsilon\}$ which has the following property:

$$\Pr(\zeta_t = \varepsilon | \varepsilon_{t+1} = \varepsilon) = \Pr(\zeta_t = -\varepsilon | \varepsilon_{t+1} = -\varepsilon) = q > 1/2.$$

An alternative to the information-based interpretation of ζ_t is that all agents have the same information set but some agents are simply better at processing information. Informed traders buy (sell) the pound forward when their signal is $\zeta_t = \varepsilon$ ($\zeta_t = -\varepsilon$). In the appendix, we show that this strategy is optimal.

A fraction $1 - \alpha$ of the traders are uninformed. These traders adopt the following behavioral trading rule. They buy pounds forward when they expect the pound to appreciate (ϕ_t is positive). They sell pounds forward when they expect the pound to depreciate (ϕ_t is negative). This rule, together with our assumptions about informed traders, generates the adverse-selection problem underlying our results. We later generalize this trading rule so that uninformed traders are more likely to buy (sell) pounds forward when, based on public information, the pound is expected to appreciate (depreciate).

All trade takes place with market makers. There is free entry into market making, so competition drives the expected profits of market makers to zero.⁹ Each market maker draws one trader per period from a continuum. The trader can submit an order of fixed size x to buy or sell pounds forward. Since agents are risk neutral they typically want to buy or sell an infinite number pounds forward. For simplicity we limit the order size to a finite number, x . We could make the order size finite by assuming that traders are risk averse, but this would greatly complicate the analysis.

⁹The probability of a trader trading more than once at time t is zero. This property rules out strategic considerations.

2.3. The Market Maker's Problem

The market maker does not observe ε_{t+1} or whether a trader is informed. However, he does observe ϕ_t and knows whether a trader wants to buy or sell pounds forward. At time t the market maker posts ask and bid forward rates, $F_t^a(\phi_t)$ and $F_t^b(\phi_t)$, that depend on ϕ_t . To illustrate the nature of the market maker's problem we now derive $F_t^a(\phi)$, the ask forward rate when ϕ_t is positive and uninformed traders expect the pound to appreciate. See the appendix for the derivations of $F_t^b(\phi)$, $F_t^a(-\phi)$, and $F_t^b(-\phi)$.

When $\phi_t = \phi$, the market maker's profit from selling one pound forward, π_{t+1}^m , is:

$$\pi_{t+1}^m = F_t^a(\phi) - S_{t+1}.$$

Here, π_{t+1}^m is denominated in units of the foreign currency. Since the market maker's expected profit is zero it follows that:

$$E(\pi_{t+1}^m | \text{buy}, \phi) = F_t^a(\phi) - E(S_{t+1} | \text{buy}, \phi) = 0.$$

Using equation (2.1) we have:

$$F_t^a(\phi) = S_t [1 + \phi + E(\varepsilon_{t+1} | \text{buy}, \phi)]. \quad (2.2)$$

We now use Bayes' rule to evaluate the market maker's expectation of ε_{t+1} conditional on his information set. This expectation is given by:

$$E(\varepsilon_{t+1} | \text{buy}, \phi) = \Pr(\varepsilon_{t+1} = \varepsilon | \text{buy}, \phi)(\varepsilon) + \Pr(\varepsilon_{t+1} = -\varepsilon | \text{buy}, \phi)(-\varepsilon). \quad (2.3)$$

Bayes' rule implies:

$$\Pr(\varepsilon_{t+1} = \varepsilon | \text{buy}, \phi) = \frac{\Pr(\text{buy} | \varepsilon_{t+1} = \varepsilon, \phi) \Pr(\varepsilon_{t+1} = \varepsilon)}{\Pr(\text{buy} | \phi)}. \quad (2.4)$$

To compute $\Pr(\text{buy} | \varepsilon_{t+1} = \varepsilon, \phi)$ we must distinguish between the actions of informed and uninformed traders. When $\phi_t = \phi$ uninformed traders buy the pound forward. When $\varepsilon_{t+1} = \varepsilon$, informed traders receive, with probability q , the signal $\zeta_t = \varepsilon$. Given this signal they buy the pound forward. Since there are $1 - \alpha$ uninformed traders and α informed traders, it follows that:

$$\Pr(\text{buy} | \varepsilon_{t+1} = \varepsilon, \phi) = 1 - \alpha + \alpha q. \quad (2.5)$$

This equation allows us to compute the numerator of (2.4). Turning to the denominator, we have that:

$$\Pr(\text{buy}|\phi) = \Pr(\text{buy}|\varepsilon_{t+1} = \varepsilon, \phi) \Pr(\varepsilon_{t+1} = \varepsilon) + \Pr(\text{buy}|\varepsilon_{t+1} = -\varepsilon, \phi) \Pr(\varepsilon_{t+1} = -\varepsilon). \quad (2.6)$$

Therefore we need to compute $\Pr(\text{buy}|\varepsilon_{t+1} = -\varepsilon, \phi)$. Uninformed traders buy the pound forward because $\phi_t = \phi$. Since $\varepsilon_{t+1} = -\varepsilon$ informed traders receive, with probability $1 - q$, the signal $\zeta_t = \varepsilon$. Given this signal they buy the pound forward, so:

$$\Pr(\text{buy}|\varepsilon_{t+1} = -\varepsilon, \phi) = 1 - \alpha + \alpha(1 - q). \quad (2.7)$$

Equations (2.5), (2.6) and (2.7) imply:

$$\Pr(\text{buy}|\phi) = (1 - \alpha + \alpha q) \frac{1}{2} + [1 - \alpha + \alpha(1 - q)] \frac{1}{2} = 1 - \frac{\alpha}{2}. \quad (2.8)$$

Substituting (2.5) and (2.8) into (2.4) we obtain:

$$\Pr(\varepsilon_{t+1} = \varepsilon|\text{buy}, \phi) = \frac{1 - \alpha(1 - q)}{2 - \alpha}. \quad (2.9)$$

Since

$$\Pr(\varepsilon_{t+1} = -\varepsilon|\text{buy}, \phi) = 1 - \Pr(\varepsilon_{t+1} = \varepsilon|\text{buy}, \phi),$$

we have

$$\Pr(\varepsilon_{t+1} = -\varepsilon|\text{buy}, \phi) = \frac{1 - \alpha q}{2 - \alpha}. \quad (2.10)$$

Equations (2.9) and (2.10) imply that when the market maker receives a buy order he attaches a higher probability to $\varepsilon_{t+1} = \varepsilon$ than to $\varepsilon_{t+1} = -\varepsilon$. The intuition is that uninformed traders' actions are not influenced by ε_{t+1} while informed traders are more likely to buy when $\varepsilon_{t+1} = \varepsilon$ than when $\varepsilon_{t+1} = -\varepsilon$.

Using (2.9), (2.10), and (2.3) we obtain:

$$E(\varepsilon_{t+1}|\text{buy}, \phi) = \frac{\alpha}{2 - \alpha}(2q - 1)\varepsilon. \quad (2.11)$$

It follows that $E(\varepsilon_{t+1}|\text{buy}, \phi)$ is zero when there is no private information, $\alpha = 0$ or $q = 0.5$. However, in the presence of private information $E(\varepsilon_{t+1}|\text{buy}, \phi)$ is positive.

Equation (2.2) implies:

$$F_t^a(\phi) = S_t \left[1 + \phi + \frac{\alpha}{2 - \alpha}(2q - 1)\varepsilon \right]. \quad (2.12)$$

When there is no private information, $F_t^a(\phi)$ is independent of ε because a buy order conveys no information about future exchange rates. When there is private information, $\alpha > 0$ and $q > 0.5$, $F_t^a(\phi)$ is an increasing function of the fraction of informed traders, α , and of ε , which is the standard deviation of ε_{t+1} .

In the appendix we show that the full set of forward rates, $F_t^b(\phi)$, $F_t^a(-\phi)$, and $F_t^b(-\phi)$ is given by:

$$\begin{aligned} F_t^a(\phi_t) &= \begin{cases} S_t [1 + \phi + (2q - 1)\varepsilon\alpha/(2 - \alpha)] & \text{if } \phi_t = \phi, \\ S_t [1 - \phi + (2q - 1)\varepsilon] & \text{if } \phi_t = -\phi, \end{cases} \\ F_t^b(\phi_t) &= \begin{cases} S_t [1 + \phi - (2q - 1)\varepsilon] & \text{if } \phi_t = \phi, \\ S_t [1 - \phi - (2q - 1)\varepsilon\alpha/(2 - \alpha)] & \text{if } \phi_t = -\phi. \end{cases} \end{aligned} \quad (2.13)$$

3. Properties of the Model

In this section we analyze the properties of our model and deduce its implications for the variability of bid-ask spreads, the population values of the slope coefficient in the forward premium regression, and the mean-squared errors associated with different exchange rate forecasts.

3.1. Bid-Ask Spreads

It is well-known that bid-ask spreads in forward markets display very low levels of volatility (see also Section 4). A natural question is whether our model is consistent with this fact. The following proposition establishes that it is.

Proposition 3.1. *To a first-order approximation, the bid-ask spread in forward markets is independent of ϕ_t and constant over time:*

$$\ln [F_t^a(\phi_t)/F_t^b(\phi_t)] \cong \frac{2}{2 - \alpha} (2q - 1) \varepsilon. \quad (3.1)$$

The proof of this proposition follows directly from equations (2.12) and (2.13).

According to equation (3.1) the severity of the adverse-selection problem in forward markets is reflected in the level of the bid-ask spread but not in its volatility. When there is no adverse selection ($q = 1/2$) the bid-ask spread collapses to zero. The bid-ask spread is increasing in the precision of the signal received by the informed agents. The spread is also increasing in ε , which is the standard deviation of ε_{t+1} .

3.2. The Forward Premium Regression

We now state the main result of our paper. The following proposition establishes the conditions under which our model can account, in population, for the forward premium puzzle.

Proposition 3.2. *Consider the regression equation:*

$$\frac{S_{t+1} - S_t}{S_t} = a + \beta \frac{F_t - S_t}{S_t} + \xi_{t+1}, \quad (3.2)$$

where ξ_{t+1} is the regression error. Suppose that the data is generated by our model. Then the probability limit of $\hat{\beta}$, the least-squares estimator for β , is:

$$\text{plim } \hat{\beta} = \frac{\phi}{\phi - (1 - \alpha)(2q - 1)\varepsilon/(2 - \alpha)}.$$

This result holds whether F_t is measured using the ask rate, the bid rate, or an average of the two. If

$$\phi < \frac{1 - \alpha}{2 - \alpha} (2q - 1)\varepsilon, \quad (3.3)$$

then $\text{plim } \hat{\beta} < 0$.

Proof: See appendix.

To understand why proposition 3.2 holds regardless of how F_t is measured, note that (2.13) implies:

$$\frac{F_t^a(\phi_t) - S_t}{S_t} = \frac{F_t^b(\phi_t) - S_t}{S_t} + \frac{2}{2 - \alpha} (2q - 1)\varepsilon,$$

so the right-hand side variable in regression (3.2) is always the same up to an additive constant.

To provide intuition for why $\text{plim } \hat{\beta}$ can be negative we begin by considering two extreme cases. In the first case, all traders are uninformed, so market makers do not face an adverse-selection problem. In the second case all traders are informed. In this case there is an adverse-selection problem but its severity is not related to the value of ϕ_t . In both cases $\text{plim } \hat{\beta} = 1$. To obtain $\text{plim } \hat{\beta} < 0$ it is critical that there is less of an adverse-selection problem when, based on public information, the currency is expected to appreciate ($\phi_t = \phi$).

All traders are uninformed ($q = 1/2$) In this case market makers do not face an adverse-selection problem. Equation (2.13) implies that the bid-ask spread is zero, $F_t^a(\phi_t) = F_t^b(\phi_t)$, and the forward premium is:

$$\frac{F_t^a(\phi_t) - S_t}{S_t} = \frac{F_t^b(\phi_t) - S_t}{S_t} = \phi_t. \quad (3.4)$$

Since ε_{t+1} , ω_{t+1} , and ϕ_t are orthogonal, equations (2.1) and (3.4) imply that the forward rate is equal to the expected value of the future spot exchange rate. Consequently, $\text{plim } \hat{\beta} = 1$.

All traders are informed ($\alpha = 1$, $q > 1/2$) In this case the direction of an order completely reveals the signal received by traders, so that all agents have the same expectation about ε_{t+1} and S_{t+1} . Equation (2.13) implies that the forward premium is:

$$\begin{aligned}\frac{F_t^a(\phi_t) - S_t}{S_t} &= \phi_t + (2q - 1)\varepsilon, \\ \frac{F_t^b(\phi_t) - S_t}{S_t} &= \phi_t - (2q - 1)\varepsilon.\end{aligned}\tag{3.5}$$

The intuition underlying (3.5) is particularly transparent when $q = 1$. In this case all agents except the market maker have perfect information about ε_{t+1} . When an agent wants to buy (sell) the pounds forward the market maker can deduce that $\varepsilon_{t+1} = \varepsilon$ ($\varepsilon_{t+1} = -\varepsilon$). Consequently, the forward rate fully reflects the realized value of ε_{t+1} . Since ε_{t+1} , ω_{t+1} , and ϕ_t are orthogonal, (2.1) and (3.5) imply that $\text{plim } \hat{\beta} = 1$ regardless of whether we use F_t^a or F_t^b in regression (3.2). This result holds in the more general case where $q > 1/2$.

Informed and uninformed traders ($\alpha < 1$, $q > 1/2$) The central feature of this case is that a market maker faces less adverse selection when $\phi_t = \phi$. To make the intuition for this case as transparent as possible suppose that $q = 1$. When $\phi_t = -\phi$ *only* informed agents buy the pound forward. It follows that when the market maker receives a buy order he can infer with certainty that the buyer is informed and that $\varepsilon_{t+1} = \varepsilon$. Consequently, $F_t^a(-\phi)$ fully reflects the fact that $\varepsilon_{t+1} = \varepsilon$:

$$F_t^a(-\phi) = S_t(1 - \phi + \varepsilon).\tag{3.6}$$

In contrast, when $\phi_t = \phi$, both uninformed and informed agents buy the pound forward. It follows that a buy order can come from either an uninformed agent responding to $\phi_t > 0$ or from an informed agent who knows that $\varepsilon_{t+1} = \varepsilon$. With $q = 1$, equation (2.11) implies that:

$$E(\varepsilon_{t+1}|\text{buy}, \phi) = \frac{\alpha}{2 - \alpha}\varepsilon < \varepsilon.\tag{3.7}$$

Consequently, the forward rate is given by (see (2.13)):

$$F_t^a(\phi) = S_t \left(1 + \phi + \frac{\alpha}{2 - \alpha}\varepsilon \right).\tag{3.8}$$

In this equation, the coefficient on ε is less than one.

Comparing (3.6) with (3.8) and imposing (3.3) in proposition 3.2 we have:

$$F_t^a(\phi) < F_t^a(-\phi).$$

The ask forward rate is actually *lower* when, based on public information, the pound is expected to *appreciate*. Consequently, the covariance between $(S_{t+1} - S_t)/S_t$ and $[F_t^a(\phi_t) - S_t]/S_t$ is negative. It follows that $\text{plim } \hat{\beta} < 0$.

Condition (3.3) plays an important role in our results. This condition requires that the predictability of exchange rate changes based on public information be relatively small. This restriction is consistent with the large literature that documents how difficult it is to predict exchange rates.

We now provide an alternative interpretation of condition (3.3). The expected profits of informed traders are:

$$\pi_i^e = \frac{1 - \alpha}{2 - \alpha} (2q - 1) \varepsilon S_t. \quad (3.9)$$

The expected profits of the uninformed traders are given by:

$$\pi_u^e = -\frac{\alpha}{2 - \alpha} (2q - 1) \varepsilon S_t.$$

Aggregate trader profits are zero:

$$\alpha \pi_i^e + (1 - \alpha) \pi_u^e = 0.$$

As is standard in models with informed and uninformed agents, informed agents make positive expected profits, whereas uninformed agents make negative expected profits. Suppose that α is close to zero (there are very few informed traders) and q is close to 1/2 (private information is very noisy). Then the expected loss of each uninformed trader is vanishingly small.

Using (3.9) we can rewrite (3.3) as:

$$\phi < \pi_i^e / S_t.$$

Since π_i^e is positive there is always a ϕ such that this condition is satisfied. Put differently, as long as it is difficult to forecast exchange rates using public information and there are informed traders who make positive expected profits, there must be a forward premium puzzle.

Up to now we have assumed that uninformed agents buy (sell) the pound forward when they expect the pound to appreciate (depreciate). What is critical for our results is that these traders are more likely to buy pounds forward when, based on public information, the pound is expected to appreciate. The following proposition formalizes this point.

Proposition 3.3. *Suppose that the data are generated by a version of our model in which, with probability $v > 1/2$, uninformed traders sell (buy) the pound forward when $\phi_t > 0$ ($\phi_t < 0$). Then $\text{plim } \hat{\beta}$ in regression (3.2) is:*

$$\text{plim } \hat{\beta} = \frac{\phi}{\phi - \alpha(z - 1)(2q - 1)\varepsilon / [z(2 - z)]},$$

where $z = 2v(1 - \alpha) + \alpha$. This results holds whether F_t is measured using the ask rate, the bid rate, or an average of the two. Suppose that:

$$\phi < \frac{\alpha(z - 1)}{z(2 - z)}(2q - 1)\varepsilon. \quad (3.10)$$

Then $\text{plim } \hat{\beta} < 0$.

Proof: See appendix.

Since the right hand side of (3.10) is increasing in v , the higher v is the more likely it is that condition (3.10) holds. When $v = 1$ condition (3.10) reduces to condition (3.3).

A testable implication of our model is that the magnitude of the forward premium puzzle declines as the forward premium becomes more volatile. This property is summarized in the following proposition:

Corollary 3.4. *If condition (3.10) holds then $\text{plim } \hat{\beta}$ can be written as:*

$$\text{plim } \hat{\beta} = \frac{-\phi}{\text{Standard Deviation} [(F_t - S_t)/S_t]}.$$

Proof: See appendix.

Suppose that we are willing to assume that the parameter ϕ is roughly the same across a group of countries. Then, we can assess the prediction provided by the previous corollary without taking a stand on the values of the model's underlying parameters. We discuss our results in Section 5.

3.3. Forecasting the Future Spot Exchange Rate

We now consider another prediction of our model that involves the relative performance of two alternative forecasts of S_{t+1}/S_t . The first forecast is based on the forward exchange rate: $E(S_{t+1}/S_t) = F_t/S_t$. The second forecast assumes that S_t is a martingale, so that: $E(S_{t+1}/S_t) = 1$. Since forward rates are forward looking it is natural to expect the first forecast to outperform the second forecast. However, as we show in the next section, the opposite is true in our data set.

There is a close connection between the forward premium puzzle and the fact that the spot exchange rate outperforms the forward rate in predicting the future spot rate. Define the mean-square-forecast error of S_{t+1}/S_t based on the average of the bid and ask forward exchange rates:

$$MSE_F = E [(S_{t+1} - F_t)/S_t]^2. \quad (3.11)$$

Also define the mean-square-forecast error of S_{t+1}/S_t based on the spot exchange rate as:

$$MSE_S = E [(S_{t+1} - S_t)/S_t]^2. \quad (3.12)$$

It is straightforward to show that:

$$MSE_F - MSE_S = \text{var} [(F_t - S_t)/S_t] (1 - 2\beta). \quad (3.13)$$

Equation (3.13) implies that models embodying uncovered interest rate parity ($\beta = 1$) are inconsistent with the finding that $MSE_F > MSE_S$. Since our model can generate $\beta < 0$, it implies that $MSE_F > MSE_S$.

4. A Numerical Example

In this section We conclude by addressing three potential concerns about our explanation of the forward premium puzzle. The first concern is that we require volatility in bid-ask spreads that are much higher than those observed in the data. In fact, our model generates forward bid-ask spreads that are, to a first-order approximation, constant. So, if anything our model understates the volatility of bid-ask spreads. The second concern is that our model requires very volatile forward premia. Given covered interest rate parity, this property would imply that our model generates movements in interest rate spreads that are counterfactually high. The third concern is that our model requires a large fraction of informed traders.

The ability of our model to account for the forward premium puzzle would not be very interesting if we had to assume that forward premia are much more volatile than the data. Given covered interest rate parity, this shortcoming would translate into counterfactually volatile interest rate movements. The plausibility of our model would also be seriously be called into question if we had to assume that a large fraction of traders is informed. In this section we provide a numerical example, loosely calibrated to the data, to demonstrate that the model can generate large negative values for the plim of β in regression (3.2) with low volatility in forward premia and a small value of α .

We begin by describing some basic properties of the data. Our data set, obtained from Datastream, consists of daily observations on dealer quotes of bid and ask spot exchange rates and 1-month forward exchange rates. We convert daily data into non-overlapping monthly observations.¹⁰ Our sample period is January 1976 to December 1998 for the Euro-legacy currencies and January 1976 to December 2005 for all other currencies. The countries included in our data set are Belgium, Canada, France, Germany, Italy, Japan, the Netherlands, Switzerland, the U.K., and the U.S. All exchange rates are quoted in units of foreign currency per British pound.

Table 1 summarizes some basic properties of the data. Column 1 reports the time series median bid-ask spread on one-month forward contracts against the pound. The numbers range from a high of 0.34 percent for the Dutch guilder to a low of 0.07 percent for the U.S. dollar. The average bid-ask spread across currencies is about 0.23 percent. Column 2 reports the standard deviation of the bid-ask spread for each of the nine currencies. From this column we see that the bid-ask spread volatility is quite low. The average, across currencies, of the standard deviation of the bid-ask spread is roughly 0.13 percent. Columns 3 and 4 report the standard deviations of the monthly rate of depreciation, $\delta_{t+1} = (S_{t+1} - S_t) / S_t$, and the forward premium, $f_t = (F_t - S_t) / S_t$, where S_t and F_t are measured as the average of bid and ask rates. We denote by σ_δ and σ_f the standard deviation of δ_{t+1} and f_t , respectively. The average values of σ_δ and σ_f are 2.9 percent and 0.26 percent, respectively. Column 5 reports the ratio of these standard deviations. The average ratio of the two volatilities is roughly 12, indicating that changes in exchange rates are extraordinarily volatile relative to the forward premium. Column 6 reports our estimates of β , the slope coefficient in regression (3.2). Consistent with the literature, our estimates of β are negative for each of the nine

¹⁰See the appendix in Burnside, Eichenbaum, Kleshchelski, and Rebelo (2006) for details.

countries in our sample. The average point estimate of β across countries is -1.75 . Finally, column 7 reports the ratio of the sample analogues to MSE_F and MSE_S defined in (3.11) and (3.12). Consistent with equation (3.13), this ratio is always above one, with an average value of 1.032.

In our numerical example we set the percentage of informed traders to a small number, $\alpha = 0.0001$. For simplicity we assume that $v = 1$, so that uninformed agents always buy (sell) the pound forward when $\phi_t > 0$ ($\phi_t < 0$). We then choose values of σ_ω , ϕ , and ε so as to match, with the lowest possible value of q , the average monthly value of σ_δ (0.029), the average estimate of β (-1.75), and the average monthly bid-ask spread in forward markets (0.0023). The resulting value of q is 0.54. The corresponding values of σ_ω , ϕ , and ε are 0.0037, 0.00073, and 0.0287, respectively. Clearly, it is possible to account for the forward premium puzzle without assuming that there is a large fraction of informed traders who receive very precise signals about future spot exchange rates.

In our calibration the process ε_{t+1} , about which informed traders receive signals, generates most of the volatility in the exchange rate. However, the quality of private information about ε_{t+1} is relatively low, as q is close to $1/2$. Alternatively, we can match the same set of moments in the data (σ_δ , $\hat{\beta}$, and the average bid-ask spread) by assuming that ω_{t+1} generates most of the volatility in the exchange rate, as long as we also assume a larger value of q . In this case, private information about ε_{t+1} is of higher quality, but this information is less useful in forecasting S_{t+1} . In both cases, informed traders have only a limited ability to forecast S_{t+1} .

We now turn to the model's implication for the volatility of interest rate spreads. The average monthly volatility of the forward premia is 0.04 percent in our model and on average 0.26 percent across the nine countries in our sample (see Table 1). Given covered interest parity this result implies that interest rate spreads are much more volatile in our model than in the data. So the adverse selection mechanism in our model can generate the forward premium puzzle, while adding very little volatility to interest rates.

5. Is There Always a Forward Premium Puzzle?

According to our model $\text{plim } \hat{\beta} < 0$ only when condition (3.10) is satisfied. As ϕ gets very large most of the variation in exchange rates can be predicted on the basis of public information and $\text{plim } \hat{\beta}$ converges to one. This property is desirable because the forward

premium puzzle is not uniformly present in the data. Bansal and Dahlquist (2000) show that, in a cross section of countries, estimates of β are positively related to the average rate of inflation. Suppose that the predictability of the spot exchange rate based on public information is higher in high inflation countries. Then, our model is consistent with the Bansal and Dahlquist (2000) findings. Figure 1 provides complementary evidence to Bansal and Dahlquist (2000). The first panel of Figure 1 displays the cross-sectional relation between the average monthly inflation rate in the period 1976–1998 and estimates of $\hat{\beta}$ for a group of 15 countries. The average annual rate of inflation across the countries included in Table 1 is equal to 4.6 percent. The larger data set being used here includes Denmark, Ireland, Norway, Portugal, Spain, and Sweden, countries whose average annual rate of inflation was 8.2 percent between 1976 and 1998. Not surprisingly, there is a positive cross-sectional relation between the average rate of inflation and the average rate of depreciation over this sample period (see the second panel of Figure 1).

We conclude by showing that the data are consistent with the prediction associated with the corollary to proposition 3.3. According to this corollary, $\text{plim } \hat{\beta}$ becomes a larger negative number when the volatility of the forward premium decreases. We assess this prediction using the cross-country relation between the magnitude of $\hat{\beta}$ and the estimated volatility of the forward premium. Both columns 3 and 5 of Table 1 and Figure 2 show a tight connection between the magnitude of $\hat{\beta}$ and the volatility of the forward premium in the direction predicted by our model.

6. Conclusion

In this paper we present a model in which adverse-selection problems between market makers and traders rationalizes a negative covariance between the forward premium and changes in exchange rates. The key feature of our model is that the adverse-selection problem facing market makers is worse when, based on public information, a currency is expected to appreciate.

Macroeconomists generally assume that asset markets are Walrasian in nature. This assumption is questionable on empirical grounds. Other assets, such as treasury bills are traded in over-the-counter markets in which market makers and traders interact (see e.g. Fleming (1997) and Massa and Simonov (2003)). Our results suggest that adverse selection problems in these markets are a promising avenue of research for understanding asset pricing

puzzles that have been difficult to resolve purely on the basis of risk considerations.

We conclude by noting an important question that we do not address in this paper: at what horizon does private information matter for exchange rates? Albuquerque, de Francisco, and Marques (2007) provide empirical evidence on the answer to this question. These authors argue that private information has significant effects on various asset returns, including exchange rates, at least at a monthly frequency. Further work on this issue would be extremely useful for assessing the plausibility of our proposed resolution of the forward premium puzzle.

References

- Albuquerque, Rui, Greg Bauer, and Martin Schneider (2007) “International Equity Flows and Returns: A Quantitative Equilibrium Approach,” *The Review of Economic Studies*, 74: 1-30.
- Albuquerque, Rui, Eva de Francisco, and Luis Marques (2007) “Marketwide Private Information in Stocks: Forecasting Currency Returns,” forthcoming *Journal of Finance*.
- Bacchetta, Philippe and Eric van Wincoop (2006) “Incomplete Information Processing: A Solution to the Forward Discount Puzzle,” mimeo, Study Center Gerzensee.
- Bansal, Ravi and Magnus Dahlquist (2000) “The Forward Premium Puzzle: Different Tales from Developed and Emerging Economies,” *Journal of International Economics* 51 (1, June): 115–44.
- Burnside, Craig “Empirical Asset Pricing Models: A Critical Review,” mimeo, Duke University, 2007.
- Engel, Charles (1996) “The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence,” *Journal of Empirical Finance* 3 (2, June), 123–92.
- Fleming, Michael “The Round-the-Clock Market for U.S. Treasury Securities,” *Federal Reserve Bank of New York Economic Policy Review*, July 1997.
- Corsetti, Giancarlo, Amil Dasgupta, Stephen Morris and Hyun Son Shin (2004) “Does One Soros Make a Difference? A Theory of Currency Crises with Large and Small Traders,” *Review of Economic Studies* 71 (1, January) 2004, 87-113.
- Glosten, Lawrence and Paul Milgrom (1985) “Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders,” *Journal of Financial Economics* 14 (1, March), 71–100.
- Han, Bing, David Hirshleifer, and Tracy Wang (2006) “Investor Overconfidence and the Forward Discount Puzzle,” University of Texas at Austin.
- Hansen, Lars Peter and Kenneth J. Singleton (1982) “Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models,” *Econometrica*, 50, (5, September), 1269-1286.
- Lyons, Richard K. (2001) *The Microstructure Approach to Exchange Rates*. Cambridge, Mass.: MIT Press.
- Massa, Massimo and Andrei Simonov (2003) “Reputation and Interdealer Trading: a Microstructure Analysis of the Treasury Bond Market,” *Journal of Financial Markets*, 6 (2, April), 99-141.
- Sarno, Lucio and Mark Taylor (2001) *The Microstructure of the Foreign Exchange Market, A Selective Survey of the Literature*, Princeton Studies in International Economics, (89, May).

A. Appendix

We work with the most general model presented in the main text. In this version we assume that, with probability $v > 1/2$, uninformed traders sell (buy) the pound forward when $\phi_t > 0$ ($\phi_t < 0$).

A.1. Computing Forward Exchange Rates

The Ask Rate when the Public Signal is Positive When $\phi_t = \phi$, v of the uninformed traders buy the pound forward, and if $\varepsilon_{t+1} = \varepsilon$, then with probability q informed traders buy the pound forward, so

$$\Pr(\text{buy}|\varepsilon_{t+1} = \varepsilon, \phi) = (1 - \alpha)v + \alpha q. \quad (\text{A.1})$$

If $\varepsilon_{t+1} = -\varepsilon$, then, with probability $1 - q$ informed traders buy the pound forward, so

$$\Pr(\text{buy}|\varepsilon_{t+1} = -\varepsilon, \phi) = (1 - \alpha)v + \alpha(1 - q). \quad (\text{A.2})$$

Hence,

$$\Pr(\text{buy}|\phi) = (1 - \alpha)v + \frac{\alpha}{2}. \quad (\text{A.3})$$

Substituting (A.1) and (A.3) into (2.4) we obtain:

$$\Pr(\varepsilon_{t+1} = \varepsilon|\text{buy}, \phi) = \frac{(1 - \alpha)v + \alpha q}{2v(1 - \alpha) + \alpha}. \quad (\text{A.4})$$

$$\Pr(\varepsilon_{t+1} = -\varepsilon|\text{buy}, \phi) = \frac{(1 - \alpha)v + \alpha - \alpha q}{2v(1 - \alpha) + \alpha}. \quad (\text{A.5})$$

Using (A.4), (A.5), and (2.3) we obtain:

$$E(\varepsilon_{t+1}|\text{buy}, \phi) = \frac{\alpha}{2v(1 - \alpha) + \alpha}(2q - 1)\varepsilon.$$

Substituting this result into equation (2.2) implies:

$$F_t^a(\phi) = S_t \left[1 + \phi + \frac{\alpha}{2v(1 - \alpha) + \alpha}(2q - 1)\varepsilon \right]. \quad (\text{A.6})$$

When $\nu = 1$ this expression reduces to the one found in equation (2.13).

The Ask Rate when the Public Signal is Negative $F_t^a(-\phi)$ is equal to the market maker's expectation of S_{t+1} conditional on having received a buy order and on $\phi_t = -\phi$:

$$F_t^a(-\phi) = E(S_{t+1}|\text{buy}, -\phi) = S_t [1 - \phi + E(\varepsilon_{t+1}|\text{buy}, -\phi)]. \quad (\text{A.7})$$

Now

$$E(\varepsilon_{t+1}|\text{buy}, -\phi) = \Pr(\varepsilon_{t+1} = \varepsilon|\text{buy}, -\phi)\varepsilon + \Pr(\varepsilon_{t+1} = -\varepsilon|\text{buy}, -\phi)(-\varepsilon). \quad (\text{A.8})$$

To calculate the probabilities on the right-hand-side of (A.8):

$$\Pr(\varepsilon_{t+1} = \varepsilon|\text{buy}, -\phi) = \frac{\Pr(\text{buy}|\varepsilon_{t+1} = \varepsilon, -\phi) \Pr(\varepsilon_{t+1} = \varepsilon)}{\Pr(\text{buy}|-\phi)}. \quad (\text{A.9})$$

When $\phi_t = -\phi$, $1 - v$ of the uninformed traders buy the pound forward, and if $\varepsilon_{t+1} = \varepsilon$, then with probability q informed traders buy the pound forward, so

$$\Pr(\text{buy}|\varepsilon_{t+1} = \varepsilon, -\phi) = (1 - \alpha)(1 - v) + \alpha q. \quad (\text{A.10})$$

If $\varepsilon_{t+1} = -\varepsilon$, then with probability $1 - q$, informed traders buy the pound forward, so

$$\Pr(\text{buy}|\varepsilon_{t+1} = -\varepsilon, -\phi) = (1 - \alpha)(1 - v) + \alpha(1 - q). \quad (\text{A.11})$$

Hence,

$$\Pr(\text{buy}|-\phi) = (1 - \alpha)(1 - v) + \frac{\alpha}{2}. \quad (\text{A.12})$$

Substituting (A.10) and (A.12) into (A.9) we obtain:

$$\Pr(\varepsilon_{t+1} = \varepsilon|\text{buy}, -\phi) = \frac{(1 - \alpha)(1 - v) + \alpha q}{2(1 - \alpha)(1 - v) + \alpha}. \quad (\text{A.13})$$

$$\Pr(\varepsilon_{t+1} = -\varepsilon|\text{buy}, -\phi) = \frac{(1 - \alpha)(1 - v) + \alpha(1 - q)}{2(1 - \alpha)(1 - v) + \alpha}. \quad (\text{A.14})$$

Using (A.13), (A.14), and (A.8) we obtain,

$$E(\varepsilon_{t+1}|\text{buy}, -\phi) = \frac{\alpha}{2(1 - \alpha)(1 - v) + \alpha}(2q - 1)\varepsilon.$$

Equation (A.7) implies:

$$F_t^a(-\phi) = S_t \left[1 - \phi + \frac{\alpha}{2(1 - \alpha)(1 - v) + \alpha}(2q - 1)\varepsilon \right]. \quad (\text{A.15})$$

The Bid Rate when the Public Signal is Positive $F_t^b(\phi)$ is equal to the market maker's expectation of S_{t+1} conditional on having received a sell order and on $\phi_t = \phi$:

$$F_t^b(\phi) = E(S_{t+1}|\text{sell}, \phi) = S_t [1 + \phi + E(\varepsilon_{t+1}|\text{sell}, \phi)]. \quad (\text{A.16})$$

Now

$$E(\varepsilon_{t+1}|\text{sell}, \phi) = \Pr(\varepsilon_{t+1} = \varepsilon|\text{sell}, \phi)\varepsilon + \Pr(\varepsilon_{t+1} = -\varepsilon|\text{sell}, \phi)(-\varepsilon). \quad (\text{A.17})$$

To calculate this we need

$$\Pr(\varepsilon_{t+1} = \varepsilon|\text{sell}, \phi) = \frac{\Pr(\text{sell}|\varepsilon_{t+1} = \varepsilon, \phi) \Pr(\varepsilon_{t+1} = \varepsilon)}{\Pr(\text{sell}|\phi)}. \quad (\text{A.18})$$

Since agents either buy or sell, we can use our calculations above, i.e. (A.1), (A.2), and (A.3) to get the probabilities:

$$\Pr(\text{sell}|\varepsilon_{t+1} = \varepsilon, \phi) = 1 - (1 - \alpha)v - \alpha q, \quad (\text{A.19})$$

$$\Pr(\text{sell}|\varepsilon_{t+1} = -\varepsilon, \phi) = 1 - (1 - \alpha)v - \alpha(1 - q), \quad (\text{A.20})$$

$$\Pr(\text{sell}|\phi) = 1 - (1 - \alpha)v - \frac{\alpha}{2}. \quad (\text{A.21})$$

Substituting (A.19) and (A.21) into (A.18) we obtain:

$$\Pr(\varepsilon_{t+1} = \varepsilon|\text{sell}, \phi) = \frac{1 - (1 - \alpha)v - \alpha q}{2 - 2(1 - \alpha)v - \alpha}, \quad (\text{A.22})$$

and

$$\Pr(\varepsilon_{t+1} = -\varepsilon|\text{sell}, \phi) = \frac{1 - \alpha - (1 - \alpha)v + \alpha q}{2 - 2(1 - \alpha)v - \alpha}. \quad (\text{A.23})$$

Using (A.22), (A.23), and (A.17) we obtain:

$$E(\varepsilon_{t+1}|\text{sell}, \phi) = -\frac{\alpha}{2 - 2(1 - \alpha)v - \alpha}(2q - 1)\varepsilon.$$

Equation (A.16) implies:

$$F_t^b(\phi) = S_t \left[1 + \phi - \frac{\alpha}{2 - 2(1 - \alpha)v - \alpha}(2q - 1)\varepsilon \right]. \quad (\text{A.24})$$

The Bid Rate when the Public Signal is Negative $F_t^b(-\phi)$ is equal to the market maker's expectation of S_{t+1} conditional on having received a sell order and on $\phi_t = -\phi$:

$$F_t^b(-\phi) = E(S_{t+1}|\text{sell}, -\phi) = S_t [1 - \phi + E(\varepsilon_{t+1}|\text{sell}, -\phi)]. \quad (\text{A.25})$$

Now

$$E(\varepsilon_{t+1}|\text{sell}, -\phi) = \Pr(\varepsilon_{t+1} = \varepsilon|\text{sell}, -\phi)\varepsilon + \Pr(\varepsilon_{t+1} = -\varepsilon|\text{sell}, -\phi)(-\varepsilon). \quad (\text{A.26})$$

To calculate this we need:

$$\Pr(\varepsilon_{t+1} = \varepsilon|\text{sell}, -\phi) = \frac{\Pr(\text{sell}|\varepsilon_{t+1} = \varepsilon, -\phi) \Pr(\varepsilon_{t+1} = \varepsilon)}{\Pr(\text{sell}|-\phi)}. \quad (\text{A.27})$$

Since agents either buy or sell, we can use our calculations above, i.e. (A.10), (A.11), and (A.12) to get the probabilities:

$$\Pr(\text{sell}|\varepsilon_{t+1} = \varepsilon, -\phi) = v + \alpha - q\alpha - v\alpha, \quad (\text{A.28})$$

$$\Pr(\text{sell}|\varepsilon_{t+1} = -\varepsilon, -\phi) = v + q\alpha - v\alpha, \quad (\text{A.29})$$

$$\Pr(\text{sell}|-\phi) = v + \frac{1}{2}\alpha - v\alpha. \quad (\text{A.30})$$

Substituting (A.28) and (A.30) into (A.27) we obtain:

$$\Pr(\varepsilon_{t+1} = \varepsilon|\text{sell}, -\phi) = \frac{v + \alpha - q\alpha - v\alpha}{2v + \alpha - 2v\alpha}, \quad (\text{A.31})$$

$$\Pr(\varepsilon_{t+1} = -\varepsilon|\text{sell}, -\phi) = \frac{v + q\alpha - v\alpha}{2v + \alpha - 2v\alpha}. \quad (\text{A.32})$$

Using (A.31), (A.32), and (A.26) we obtain:

$$E(\varepsilon_{t+1}|\text{sell}, -\phi) = -\frac{\alpha}{2v - (2v - 1)\alpha} (2q - 1) \varepsilon.$$

Equation (A.25) implies:

$$F_t^b(-\phi) = S_t \left[1 - \phi - \frac{\alpha}{2v - (2v - 1)\alpha} (2q - 1) \varepsilon \right]. \quad (\text{A.33})$$

Summarizing To summarize we have:

$$\begin{aligned} F_t^a(\phi_t) &= \begin{cases} S_t [1 + \phi + (2q - 1) \varepsilon \alpha / z] & \text{if } \phi_t = \phi, \\ S_t [1 - \phi + (2q - 1) \varepsilon \alpha / (2 - z)] & \text{if } \phi_t = -\phi, \end{cases} \\ F_t^b(\phi_t) &= \begin{cases} S_t [1 + \phi - (2q - 1) \varepsilon \alpha / (2 - z)] & \text{if } \phi_t = \phi, \\ S_t [1 - \phi - (2q - 1) \varepsilon \alpha / z] & \text{if } \phi_t = -\phi. \end{cases} \end{aligned} \quad (\text{A.34})$$

where $z = 2v(1 - \alpha) + \alpha$. In the main text we consider the case where $v = 1$, in which case $z = 2 - \alpha$ and the forward exchange rates are given by (2.13).

A.2. Is the Behavior of the Informed Optimal?

We now verify that it is optimal for an informed agent to buy the pound forward when he receives a signal $\zeta_t = \varepsilon$. Given the signal, his expectation of S_{t+1} is:

$$E(S_{t+1}|\zeta_t = \varepsilon) = S_t [1 + \phi_t + E(\varepsilon_{t+1}|\zeta_t = \varepsilon)].$$

Since

$$\begin{aligned} E(\varepsilon_{t+1}|\zeta_t = \varepsilon) &= \Pr(\varepsilon_{t+1} = \varepsilon|\zeta_t = \varepsilon)\varepsilon + \Pr(\varepsilon_{t+1} = -\varepsilon|\zeta_t = \varepsilon)(-\varepsilon), \\ &= q\varepsilon + (1 - q)(-\varepsilon) = (2q - 1)\varepsilon. \end{aligned}$$

we have

$$E(S_{t+1}|\zeta_t = \varepsilon) = S_t [1 + \phi_t + (2q - 1)\varepsilon].$$

The expected payoff associated with buying the pound forward is:

$$\pi_i^e = E(S_{t+1}|\zeta_t = \varepsilon) - F_t^a(\phi_t).$$

Hence when $\phi_t = \phi$ the agent's expected profit is:

$$\begin{aligned} \pi_i^e &= S_t [1 + \phi + (2q - 1)\varepsilon] - S_t [1 + \phi + (2q - 1)\varepsilon\alpha/z], \\ &= (1 - \alpha/z)(2q - 1)\varepsilon S_t. \end{aligned}$$

When $\phi_t = -\phi$ the agent's expected profit is:

$$\begin{aligned} \pi_i^e &= S_t [1 - \phi + (2q - 1)\varepsilon] - S_t [1 - \phi + (2q - 1)\varepsilon\alpha/(2 - z)], \\ &= [1 - \alpha/(2 - z)](2q - 1)\varepsilon S_t. \end{aligned}$$

Since $z = 2v(1 - \alpha) + \alpha$, it follows that $\alpha < z$ and that $\alpha \leq 2 - z$ (with equality only in the case where $v = 1$). Thus $\pi_i^e > 0$ when $\phi_t = \phi$ and $\pi_i^e \geq 0$ when $\phi_t = -\phi$ (with equality only if $v = 1$). Hence it is optimal for agents to buy the pound forward when $\zeta_t = \varepsilon$.

We also verify that it is optimal for an informed agent to sell the pound forward when he receives a signal $\zeta_t = -\varepsilon$. Given the signal his expectation of S_{t+1} is:

$$E(S_{t+1}|\zeta_t = -\varepsilon) = S_t [1 + \phi_t + E(\varepsilon_{t+1}|\zeta_t = -\varepsilon)].$$

Since

$$\begin{aligned} E(\varepsilon_{t+1}|\zeta_t = -\varepsilon) &= \Pr(\varepsilon_{t+1} = \varepsilon|\zeta_t = -\varepsilon)\varepsilon + \Pr(\varepsilon_{t+1} = -\varepsilon|\zeta_t = -\varepsilon)(-\varepsilon), \\ &= (1 - q)\varepsilon + q(-\varepsilon) = -(2q - 1)\varepsilon, \end{aligned}$$

we have

$$E(S_{t+1}|\zeta_t = -\varepsilon) = S_t [1 + \phi_t - (2q - 1)\varepsilon].$$

The expected payoff associated with selling the pound forward is:

$$\pi_i^e = F_t^b(\phi_t) - E(S_{t+1}|\zeta_t = -\varepsilon).$$

Hence when $\phi_t = \phi$ the agent's expected profit is:

$$\begin{aligned} \pi_i^e &= S_t [1 + \phi - (2q - 1)\varepsilon\alpha/(2 - z)] - S_t [1 + \phi - (2q - 1)\varepsilon], \\ &= [1 - \alpha/(2 - z)] (2q - 1)\varepsilon S_t. \end{aligned}$$

When $\phi_t = -\phi$ the agent's expected profit is:

$$\begin{aligned} \pi_i^e &= S_t [1 - \phi - (2q - 1)\varepsilon\alpha/z] - S_t [1 - \phi - (2q - 1)\varepsilon], \\ &= (1 - \alpha/z) (2q - 1)\varepsilon S_t. \end{aligned}$$

Using the same argument as above, $\pi_i^e \geq 0$ when $\phi_t = \phi$ (with equality only if $v = 1$) and $\pi_i^e > 0$ when $\phi_t = -\phi$. Hence it is optimal for agents to sell the pound forward when $\zeta_t = -\varepsilon$.

A.3. Proof of Propositions 3.2 and 3.3

Let $\delta_{t+1} = (S_{t+1} - S_t)/S_t$ and $f_t = (F_t - S_t)/S_t$. Consider a regression:

$$\delta_{t+1} = a + \beta f_t + \xi_{t+1}. \tag{A.35}$$

In our model

$$\delta_{t+1} = \phi_t + \varepsilon_{t+1} + \omega_{t+1}.$$

It follows from (A.34) that the mid-point forward exchange rate is:

$$F_t(\phi_t) = \begin{cases} S_t (1 + \phi - \theta\varepsilon) & \text{if } \phi_t = \phi, \\ S_t (1 - \phi + \theta\varepsilon) & \text{if } \phi_t = -\phi, \end{cases}$$

where

$$\theta = \alpha \frac{z - 1}{z(2 - z)} (2q - 1) > 0.$$

Hence $f_t = \phi_t - \text{sign}(\phi_t)\theta\varepsilon$. If ask or bid prices are used to define the forward premium then:

$$\begin{aligned} f_t^a(\phi_t) &= \begin{cases} \phi + (2q - 1)\varepsilon\alpha/z & \text{if } \phi_t = \phi, \\ -\phi + (2q - 1)\varepsilon\alpha/(2 - z) & \text{if } \phi_t = -\phi, \end{cases} \\ f_t^b(\phi_t) &= \begin{cases} \phi - (2q - 1)\varepsilon\alpha/(2 - z) & \text{if } \phi_t = \phi, \\ -\phi - (2q - 1)\varepsilon\alpha/z & \text{if } \phi_t = -\phi. \end{cases} \end{aligned}$$

A little algebra shows that these expressions can be rewritten as $f_t^a = f_t + d$ and $f_t^b = f_t - d$ where:

$$d = \frac{\alpha}{z(2-z)}(2q-1).$$

This equation establishes that the regression slope coefficient will be the same for any version of the forward price. Only the regression intercept will be different.

The slope coefficient in the regression (A.35) has the following property:

$$\text{plim } \hat{\beta} = \frac{\text{cov}(\delta_{t+1}, f_t)}{\text{var}(f_t)}.$$

Since the forward premium depends only on the value of ϕ_t , and does not depend on the realized values of ε_{t+1} and ω_{t+1} , it is straightforward to calculate $\text{var}(f_t)$ and $\text{cov}(\delta_{t+1}, f_t)$:

$$\begin{aligned} \text{var}(f_t) &= \frac{1}{2}(\phi - \theta\varepsilon)^2 + \frac{1}{2}(-\phi + \theta\varepsilon)^2 = (\phi - \theta\varepsilon)^2, \\ \text{cov}(\delta_{t+1}, f_t) &= \frac{1}{2}\phi(\phi - \theta\varepsilon) + \frac{1}{2}(-\phi)(-\phi + \theta\varepsilon) = \phi(\phi - \theta\varepsilon). \end{aligned} \quad (\text{A.36})$$

Hence

$$\text{plim } \hat{\beta} = \frac{\phi}{\phi - \theta\varepsilon}. \quad (\text{A.37})$$

So if $\phi < \theta\varepsilon$, then $\text{plim } \hat{\beta} < 0$. This condition is the same as condition (3.10) in proposition 3.3. When $v = 1$ then $z = 2 - \alpha$ and $\theta = (2q - 1)(1 - \alpha)/(2 - \alpha)$ so we obtain condition (3.3) in proposition 3.2.

The corollary to proposition 3.3 follows immediately from equations (A.36) and (A.37).

Prior to observing ϕ_t the expected profits of an informed trader are

$$\begin{aligned} \pi_i^e &= \frac{1}{2}(1 - \alpha/z)(2q - 1)\varepsilon S_t + \frac{1}{2}[1 - \alpha/(2 - z)](2q - 1)\varepsilon S_t, \\ &= \{1 - \alpha/[z(2 - z)]\}(2q - 1)\varepsilon S_t. \end{aligned}$$

When $v = 1$, $\pi_i^e = [(2q - 1)(1 - \alpha)/(2 - \alpha)]\varepsilon S_t = \theta\varepsilon S_t$. Thus, if the model is parameterized such that $\theta\varepsilon$ is large, informed traders will make large expected profits and there will be a forward premium puzzle. Notice that $\theta\varepsilon$ is increasing in q (the quality of the informed agent's information about ε) and ε (the importance of that information in forecasting exchange rates) and is decreasing in α (the number of informed agents).¹¹

¹¹In the more complicated case where $1/2 < v < 1$ it remains true that $\theta\varepsilon$ and an informed trader's profits are both increasing in q and ε , and are both decreasing in α , so choices of q , ε , and α that are consistent with large profits for the informed trader are more likely to be consistent with the forward premium puzzle.

A.4. Mean-Squared Error of Forecasting Rules

Given the result, stated above, that $f_t = \phi_t - \text{sign}(\phi_t)\theta\varepsilon$ and given that $\delta_{t+1} = \phi_t + \varepsilon_{t+1} + \omega_{t+1}$, it follows that:

$$\begin{aligned}MSE_F &= E(\delta_{t+1} - f_t)^2 = E[\varepsilon_{t+1} + \omega_{t+1} + \text{sign}(\phi_t)\theta\varepsilon]^2, \\ &= (1 + \theta^2)\varepsilon^2 + \sigma_\omega^2.\end{aligned}$$

and also:

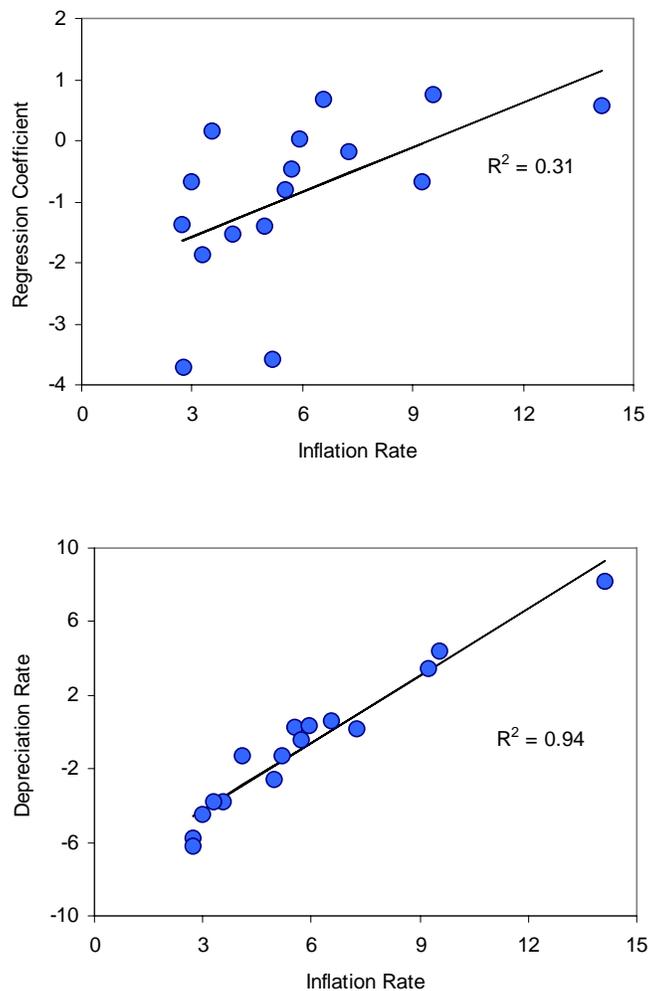
$$MSE_S = E\delta_{t+1}^2 = \phi^2 + \varepsilon^2 + \sigma_\omega^2.$$

This means that the current spot rate is a better predictor of the future spot rate if $\phi < \theta\varepsilon$. This is the same condition that determines when $\beta < 0$ in the model. However, there is a simpler way to verify that the spot is the better predictor when $\beta < 0$. By construction:

$$\begin{aligned}MSE_F - MSE_S &= E(\delta_{t+1} - f_t)^2 - E\delta_{t+1}^2, \\ &= \text{var}(f_t) - 2\text{cov}(\delta_{t+1}, f_t), \\ &= \text{var}(f_t)(1 - 2\beta).\end{aligned}$$

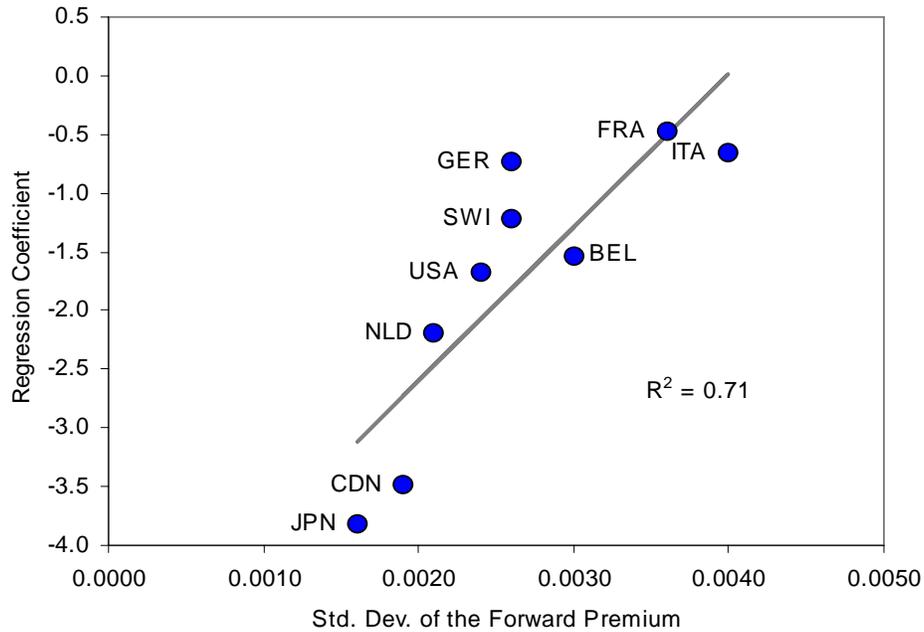
Hence, the spot rate is the better predictor whenever $\beta < 1/2$.

FIGURE 1
INFLATION, DEPRECIATION AND THE SIZE OF $\hat{\beta}$



Note: The x -axis is the average CPI inflation rate during the period 1976–1998. The y -axis in the top panel is the forward premium regression coefficient, $\hat{\beta}$. The y -axis in the bottom panel is the average annual rate of depreciation against the British pound computed over the same period. The countries in the sample are those used in Table 1 plus Austria, Denmark, Ireland, Portugal, Spain, and Sweden.

FIGURE 2
 THE VOLATILITY OF THE FORWARD PREMIUM AND THE SIZE OF $\hat{\beta}$



Note: The x -axis is the standard deviation of the forward premium, measured using average of bid and ask prices, for each of the countries indicated during the period 1976–2005, except for the Euro legacy currencies (for which the sample period ends December 1998) and Japan (for which the sample begins July 1978). The y -axis is the forward premium regression coefficient, $\hat{\beta}$, computed over the same period. The countries in the sample are Belgium, Canada, France, Germany, Italy, Japan, the Netherlands, Switzerland, and the U.S. See Burnside, Eichenbaum, Kleshchelski, and Rebelo (2006) for details of the data.

TABLE 1

Characteristics of the Data, 1976-2005

	Median Bid-ask Spread (percent)	Standard deviations				Forward Premium Regression β	$MSE_F /$ MSE_S
		Bid-ask Spreads	Rate of Depreciation	Forward Premium	Ratio		
Belgium [†]	0.25	0.11	0.027	0.0030	9.1	-1.53	1.049
Canada	0.10	0.05	0.032	0.0019	16.5	-3.49	1.029
France [†]	0.15	0.07	0.027	0.0036	7.5	-0.47	1.035
Germany [†]	0.31	0.15	0.028	0.0026	10.6	-0.73	1.022
Italy [†]	0.17	0.17	0.027	0.0040	6.8	-0.66	1.051
Japan [*]	0.27	0.19	0.035	0.0016	22.1	-3.82	1.018
Netherlands [†]	0.34	0.13	0.027	0.0021	12.8	-2.19	1.033
Switzerland	0.41	0.19	0.029	0.0026	11.0	-1.21	1.028
USA	0.07	0.07	0.031	0.0024	12.9	-1.68	1.026
Average	0.23	0.13	0.029	0.0026	12.1	-1.75	1.032

* Data for Japan begin 7/78

† Data for Euro legacy currencies ends 12/98

Sources: Datastream; see Burnside, Eichenbaum, Kleshchelski and Rebelo (2006) for details.