

The forward premium in a model with heterogeneous prior beliefs

Eric O’N. Fisher^{a,b,c}

^a *University of California at Santa Barbara, USA*

^b *Federal Reserve Bank of Cleveland, USA*

^c *The Ohio State University, USA*

Abstract

This paper explores a model of bond prices where agents have diverse prior beliefs about domestic and foreign inflation. In the long run, the foreign exchange forward premium reflects expected differences in inflation, but in the short run, it depends upon the diversity of prior beliefs. If some people have diffuse priors about a country’s inflation process, then its currency commands a forward premium that is eventually dissipated. Using data on the dollar–mark premium from the 1980s, it shows that this kind of diversity really matters. Thus models with a single representative agent give an inadequate description of the data.

Keywords: International finance; Learning, information, and knowledge; Asset pricing

1. Introduction

This paper takes the idea of heterogeneity in financial markets seriously. It develops a theory of the foreign exchange forward premium based upon the notion that people in the world economy have diverse prior beliefs about inflation. For most plausible specifications of prior beliefs, agents eventually have completely accurate knowledge about each country’s inflation

processes. Indeed, in the long run, yields reflect the common inflation forecasts, and the forward premium predicts expected depreciation of the spot rate accurately. Thus the asymptotic behavior of the world economy can be modeled using the artifice of a single “representative agent” having “rational expectations” about all the “fundamentals” in the world economy.

But in the short run a fascinating theory of asset prices emerges. This theory has two important elements. First, the learning matters. Interest differentials depend upon the stochastic inflation history in the world economy, and the model provides a simple explanation for the “forward discount anomaly.”¹ Second, the heterogeneity of beliefs matters. In this paper, I will show that it is not enough to ask traders what their inflation forecasts are; it is actually necessary to ask them *how sure they are* of their own forecasts. Because different classes of agents can hold more or less precise forecasts, bond yields typically have an option value inherent in them, even if everyone agrees on expected inflation. An asset is worth the sum of its expected real stream of income *and* the option value of reselling it at a later date. This option value can never be negative, and it is typically positive. Thus yields are lower than they would be if the world economy consisted of a single representative agent.

Whether this effect is stronger for domestic or foreign assets is at the heart of the theory. One implication is that diverse precision of beliefs about a country’s inflation process will raise the price of its bonds and thus lower their yields. Hence there will be a forward premium for that currency. An important insight is that all the moments characterizing agents’ beliefs matter. Thus it is not appropriate to consider only each person’s point forecasts of expected inflation; it matters how precise these forecasts are.

How does a theory of asset prices emerge in a model where people have diverse prior beliefs? In particular, one question arises immediately: Why is this model not plagued by Milgrom and Stokey’s (1982) No Trade Theorem? The answer lies in the subtle distinction between an environment in which agents have common priors but diverse ex post signals and one in which everyone has different priors but observes the same signals.

Consider, for example, the 500th digit in the decimal representation of e . A speaker walks into the seminar room and offers a contract that pays \$1 if that digit is 5. I may believe that it is likely to be an even number, and you may have more diffuse beliefs. We could easily announce our priors (thus establishing common knowledge), and we would both agree that you would pay more for the contract than I would.² Then the speaker opens a laptop and begins to read off numbers from the Taylor series expansion: 1.0, 2.0, 2.5, 2.67, and so on. After each new number, the speaker allows us to trade. It is quite possible that you and I would be willing to do so, with perhaps especially active trading once we get near the n -th term, where $n! \approx 10^{501}$. Thus the existence of a market (with limited short selling) where agents have heterogeneous priors is completely consistent with equilibrium.

Now think of a similar situation, but the speaker announces that he is willing to sell for \$1 a contract that pays \$1000 if the 500th digit in the decimal expansion of e is *not* 5. He then takes out his laptop and boots up. No matter what your priors were, you would be unwilling to buy that asset precisely because the speaker has shown he has received a superior signal about its

¹ See Engel (1996) for a good discussion.

² Of course, I might want to short that contract to the greatest extent possible, but we would both agree that some limit on my position is warranted because I have limited wealth. Likewise, you might like to go arbitrarily long, but your position too will be finite in practice. It will become apparent below that a limited short-selling assumption is necessary for equilibrium to exist in this kind of market.

value. Now differential information destroys the very existence of equilibrium. What is the essence of the difference between the two examples? In the first case, all the priors are common knowledge and so are all the public signals. But in the second case, even if the prior beliefs are common knowledge, the signals are not.

The model developed in Section 3 is akin to the first example. The analysis builds upon the work of Harrison and Kreps (1978), who showed that the asset's price typically exceeded the valuation of the most bullish trader. They stated that this was a formalization of Keynes's notion of a beauty contest. Morris (1996) extended this work to incorporate learning in a Bayesian framework, and Fisher (2003) extended his model to explain asset bubbles that arose in the foreign exchange experiments reported in Fisher and Kelly (2000).

This paper makes four contributions. First, it applies Morris's (1996) work by building a model of bonds and extends it by incorporating more general stochastic processes. Second, it is a completely novel analysis of the foreign exchange forward premium. To the best of my knowledge, no one has built or calibrated a model like this in international finance. Third, the model's calibration and estimation shows that plausible priors can explain some of the forward premium for the German mark during the first half of the 1980s. That period and that currency were chosen to complement Lewis (1989) impressive empirical analysis using Bayesian techniques of a reduced form model of the exchange rate. Fourth, I actually estimate the precision with which different classes of agents hold plausible prior beliefs; then I show that there is strong evidence in favor of a model with diversity of beliefs. The typical homogeneous agent model in macroeconomics is just not supported by the data.

What are this paper's main results? First, it shows that diverse prior beliefs about a country's inflation process induce a forward premium for its currency at horizons greater than one month. Second, it shows that the "peso problem" is not as simple as has been assumed; indeed, the typical interpretation of this phenomenon imposes very severe restrictions on agents' beliefs. Third, it gives a simple explanation for a strong version of the forward discount anomaly. When there is diverse prior information about a country's inflation process, its one-month forward rate will be negatively correlated with realized depreciations. Fourth, the model is calibrated and then estimated using actual data from the United States and Germany during the 1980s. The model performs well enough, although the effect of informational heterogeneity on the forward premium is not large. The calibrations of the model outperform a simple benchmark based upon covered interest parity, and they show that diversity of prior beliefs improves the model's fit. Fifth, I use a non-linear regression to test for homogeneity of beliefs in the data, and the Wald test overwhelmingly rejects the workhorse model in international finance. Diversity of beliefs really matters in these data.

The rest of this paper is organized as follows. Section 2 gives a simple but extended example because the model is strikingly different from the norm in international finance. Section 3 contains a formal description of the model, and Section 4 discusses the forward premium both when there is one representative agent and when there are several agents in the world economy. Section 5 calibrates and then estimates the model for plausible specification of the agents' prior beliefs. It also shows that the models' predictions give rise to the forward premium anomaly. Section 6 gives some brief conclusions.

2. A simple example

Consider two zero-coupon bonds maturing in two years, one denominated in dollars, the other in euros, and each with a face value of 100. These two bonds are identical in every

Table 1
Conjugate priors

	Prior beliefs about American inflation	Prior beliefs about European inflation
Type 1 agents	$\alpha = 100, \beta = 100$	$\alpha = 100, \beta = 100$
Type 2 agents	$\alpha = 100, \beta = 100$	$\alpha = 0.01, \beta = 0.01$

way—with respect to risk, liquidity, and other relevant factors—but differ solely in their currency of denomination. American inflation can take on one of the two values: 0% or 8%. European inflation can assume the same two values. Thus the inflation rate in either country is a binomial random variable. To make things very simple, we will slow down economic time and assume that information relevant to inflation forecasts arrives only once a year.

There are two classes of agents in the world economy, and each has unbiased beliefs about domestic and foreign inflation rates. Everyone is risk neutral and has very precise beliefs about American inflation. But one class of agents has much less precise beliefs about European inflation. Table 1 summarizes the relevant priors.³

A Type 2 agent is essentially a frequentist about European monetary policy: his posterior beliefs will reflect the history of European inflation almost exactly.

Agents discount felicity with the common factor $\delta = 0.97 \approx 1/1.03$. Since they are risk neutral and may hold different beliefs, it is natural to impose that there is limited short selling and to assume that there is sufficient liquidity in the market to price the current stock of assets. Let $b(s, t)$ be the price of a dollar bond when there have been s years of high American inflation during the first t years; the notation $b^*(s, t)$ is analogous for the instrument denominated in euros. These bonds can be priced using backward induction.

Consider the dollar-denominated asset. In the second year, if inflation in the United States was low in the first year, then everyone's posterior beliefs are such that:

$$b(0, 1) = \delta[(101/201)100 + (100/201)(100/1.08)] \approx 93.42.$$

On the other hand, if there has been high inflation in the United States, then:

$$b(1, 1) = \delta[(100/201)100 + (101/201)(100/1.08)] \approx 93.39.$$

This slightly lower price and higher yield reflect higher expected American inflation. Thus both classes of agents will agree at the null history that the dollar bond will cost:

$$b(0, 0) = \delta[(1/2)b(0, 1) + (1/2)b(1, 1)/1.08] \approx 87.25.$$

Pricing the euro bond is not quite as simple. If there has been low European inflation in period 1, then Type 2 agents will be bullish about European inflation and will hold all the euro-denominated assets. Thus:

$$b^*(0, 1) = \delta[(1.01/1.02)100 + (0.01/1.02)(100/1.08)] \approx 96.93.$$

³ DeGroot (1970, p. 40) shows that the natural family of conjugate priors is the beta distribution. This distribution has two parameters $\alpha > 0$ and $\beta > 0$, and its density function is $f(x) = [\Gamma(\alpha + \beta)/\Gamma(\alpha)\Gamma(\beta)]x^{\alpha-1}(1-x)^{\beta-1}$ if $0 < x < 1$ and $f(x) = 0$ otherwise, where $\Gamma(\alpha)$ is the gamma function. The mean of this random variable is $\alpha/(\alpha + \beta)$, and its variance is $\alpha\beta/(\alpha + \beta)^2(\alpha + \beta + 1)$.

But if there has been high inflation in Europe, then only Type 1 of agents will hold the European bonds, and they will pay:

$$b^*(1, 1) = \delta[(100/201)100 + (101/201)(100/1.08)] \approx 93.39.$$

Thus at the null history everyone will agree that the initial price of euro bonds is:

$$b^*(0, 0) = \delta^2[(1/2)b^*(0, 1) + (1/2)b^*(1, 1)/1.08] \approx 88.95.$$

The initial yield on European bonds is lower than that in American, *even though at the null history everyone in the world economy expects that the American and European inflation rates will be identical. The euro trades at a premium because some agents have less precise beliefs about European inflation than others.*

Of course, the difference between American and European yields is the two-year forward premium on the euro. Now let $f(s, s^*; T - t)$ be the $T - t$ year forward premium when there have been s years of high American inflation and s^* years of high European inflation during the first t years of a pair of bonds that mature in year T . A simple approximation shows that the two-year euro forward premium at the null history is:

$$f(0, 0; 2) \approx (\ln 0.9454 - \ln 0.9273)/2 \approx 0.97\%$$

per annum. The euro trades two years forward at a premium simply because the market has more diverse beliefs about European inflation. Still, the difference between the yields on one-year dollar and euro bonds at the null history would be:

$$f(0, 0; 1) = 0.0\%,$$

since everyone expects the same inflation rate for America and Europe.

For simplicity, impose that purchasing power parity holds after any history, and thus the nominal depreciation of the dollar reflects the realized inflation differential. Assume further that the realizations of American and European inflation are described by independent binomial random variables with equi-probable outcomes.⁴ Then, at the null history, the one-year forward rate is an unbiased predictor of the expected depreciation. *Still, at long horizons, the forward discount is a biased predictor of exchange rate changes even though all agents are risk neutral.* The remarkable aspect of this example is that everyone has unbiased inflation expectations for every country and there is no expected inflation differential. But there is a premium on euro-denominated assets simply because some agents have less precise beliefs about the future path of European inflation. Thus the long euro trades at a premium since the option value of reselling European bonds is not negligible.

Let me conclude this extended example by mentioning an important fact that is true in a much more general framework. The forward premium arises in a financial model with “cliente effects.” In particular, some assets are held by only a subset of the agents after some histories. There would be no option value to holding a bond denominated in any currency if everyone always held the same portfolio all the time.

⁴ If every agent has unbiased beliefs, we are implicitly imposing the equilibrium condition that $\alpha = \beta$.

3. The model

This section shows how this example generalizes. Assume again that all the agents in world financial markets are risk neutral. It is possible to allow each person to discount streams of future utility at the idiosyncratic rate δ_i . A bond dominated in domestic currency has par value V , matures in $T \geq 1$ month, and pays semi-annual coupons $c/2$, all denominated in units of the domestic currency. Even though these are “risk-free” assets, they are claims to nominal streams of income and thus are subject to inflation risk. Foreign bonds of the same maturity have par value V^* and pay an analogous coupon of $c^*/2$, both denominated in foreign currency.

The domestic inflation rate is an independent and identically distributed stochastic process with finite support. Let $\tilde{\pi}$ be a random variable with support $\Pi = \{\pi_1, \dots, \pi_k\}$, θ_j be the probability that the j -th realization of $\tilde{\pi}$ occurs, and $\theta = (\theta_1, \dots, \theta_k)$. It is appropriate to think of each period as a month, the highest frequency at which relevant price data are generally available. The random variable for foreign inflation $\tilde{\pi}^*$, its support $\Pi^* = \{\pi_1^*, \dots, \pi_{k^*}^*\}$, and the concomitant probabilities $\theta^* = (\theta_1^*, \dots, \theta_{k^*}^*)$ are all analogous.

World financial markets consist of several different classes of agents. The representative agent from each class has idiosyncratic priors about the inflation processes in each of the two countries. Let Δ^k be the relevant simplex; then the measurable function $\rho_i : \Delta^k \rightarrow [0, 1]$ denotes the i -th agent’s prior beliefs about the inflation rates of the domestic country and $\rho_i^* : \Delta^{k^*} \rightarrow [0, 1]$ represents that person’s prior beliefs about inflation in the foreign country.

The assumption that there are different classes of agents allows one to analyze price-taking equilibria, but since the agents may have different priors about world inflation processes, there may be no equilibrium unless a limited short-selling constraint is imposed. This is perhaps not an unrealistic assumption about actual financial markets, but it has profound implications for the nature of the equilibrium in this model. The descriptions of inflation and beliefs are quite general.

The inflation profiles in each country are the relevant history. Thus $h^t = ((\tilde{\pi}_1, \tilde{\pi}_1^*), \dots, (\tilde{\pi}_t, \tilde{\pi}_t^*))$ is a history of length $0 \leq t \leq T$. Now let $n(h^t) = (n_1(h^t), \dots, n_k(h^t))'$ be a vector denoting the number of times each different level of domestic inflation has been observed in history h^t . Of course, $n(h^t)$ is the sufficient statistic for estimating the home inflation process and $n^*(h^t) = (n_1^*(h^t), \dots, n_{k^*}^*(h^t))'$ is analogous for the foreign inflation.⁵ The null histories are such that $n(h^0) = (0, \dots, 0)'$ and $n^*(h^0) = (0, \dots, 0)'$.

An agent’s posteriors about the process driving domestic inflation are:

$$\varphi_i(\theta|h^t) = \frac{(\theta_1)^{n_1} \dots (\theta_k)^{n_k} \rho_i(\theta)}{\int (x_1)^{n_1} \dots (x_k)^{n_k} \rho_i(x) dx}$$

Likewise, the i -th agent’s posterior beliefs about foreign inflation are:

$$\varphi_i^*(\theta^*|h^t) = \frac{(\theta_1^*)^{n_1^*} \dots (\theta_{k^*}^*)^{n_{k^*}^*} \rho_i^*(\theta^*)}{\int (x_1)^{n_1^*} \dots (x_{k^*})^{n_{k^*}^*} \rho_i^*(x) dx}$$

Both integrals are taken over the relevant simplexes. In essence, these posteriors represent the agent’s model of the inflation process, given the common history that everyone has observed. If

⁵ These statistics are sufficient only because I have assumed that the inflation processes are i.i.d. In reality, inflation is obviously quite persistent. This will be a major foible in the model’s empirical implementation.

the priors are well behaved, these posteriors converge to the true processes, but the speed of convergence depends upon the precision of initial beliefs.

Define an indicator function

$$z(x) = \begin{cases} \pi_j & \text{if } x = (0, \dots, \theta_j, \dots, 0) \\ 0 & \text{otherwise} \end{cases}.$$

Then a representative agent's forecast of expected domestic inflation is:

$$\bar{\pi}_i(h^t) = \iint \varphi_i(\theta|h^t)z(x)d\theta dx.$$

Likewise, a typical forecast of foreign inflation is:

$$\bar{\pi}_i^*(h^t) = \iint \varphi_i^*(\theta^*|h^t)z^*(x)d\theta^* dx,$$

where all the variables and the indicator function are analogous. In each of these formulas, the inner integral is taken with respect to a person's prior beliefs and the outer integral is taken with respect to realized inflation rates. Thus this model allows for a natural generalization of expected inflation where agents have heterogeneous priors about the mechanics of monetary policy.

Let the pricing kernel $g : \Delta^k \rightarrow R_+$ have the rule

$$g(x) = \begin{cases} b(n_1, \dots, n_j + 1, \dots, n_k) & \text{if } x = (0, \dots, \theta_j, \dots, 0) \\ 0 & \text{otherwise} \end{cases},$$

where $b(n_1, \dots, n_j + 1, \dots, n_k)$ is the price of a domestic bond when the successor to history h^t has $n(h^{t+1}) = (n_1(h^t), \dots, n_j(h^t) + 1, \dots, n_k(h^t))$. Then the price of a domestic bond satisfies the recursion:

$$b(h^t) = \max_i \left\{ \delta_i \iint \frac{\varphi_i(\theta|h^t)[c/12 + g(x)]}{1 + z(x)} d\theta dx \right\},$$

where $c/12$ captures the fact that a pro-rated share of the semi-annual coupon is paid implicitly each month. The price of a foreign bond likewise satisfies:

$$b^*(h^t) = \max_i \left\{ \delta_i \iint \frac{\varphi_i^*(\theta^*|h^t)[c^*/12 + g^*(x)]}{1 + z^*(x)} d\theta^* dx \right\}.$$

These recursive formulae are at the crux of the model of the forward premium. They state a bond sells for what that most bullish class of agents will pay for it. This price is the expected present value of the pro-rated coupon and capital gains. But each agent's expectations depend upon prior beliefs about the relevant country's inflation process.

After history h^t , the $T - t$ forward discount on domestic currency is given by the difference between home and foreign yields. Thus

$$f(h^t, T - t) = (b^*(h^t)/b(h^t))^{1/(T-t)}$$

is the forward discount in percent per annum for a contract maturing at T .

A simple description of the spot exchange rate closes the model. Let the domestic and foreign price levels at the null history be $p(h^0) = 1$ and $p^*(h^0) = 1$; for any other history, the price levels are $p(h^t) = \prod_{s=1}^t (1 + \tilde{\pi}_s)$ and $p^*(h^t) = \prod_{s=1}^t (1 + \tilde{\pi}_s^*)$. Imposing heroically that the real exchange is constant,⁶ we see that the spot exchange rate is:

$$e(h^t) = p(h^t)/p^*(h^t).$$

This definition follows the American convention: the exchange rate is denominated in units of domestic currency per unit of foreign exchange. Since domestic and foreign inflation processes are independent multinomial random variables, the log of the spot exchange rate has a unit root. This model of the spot rate is unrealistic in the extreme, having little to recommend it other than analytical simplicity. Still, the actual rate of depreciation reflects the historical inflation differential, and purchasing power parity holds identically in every period.

4. The behavior of the forward premium

The forward premium will depend in general upon all the prior beliefs and the stochastic realizations of the inflation profiles in the world economy. If $\rho_i(\cdot)$ and $\rho_i^*(\cdot)$ are well behaved for every agent, then the posteriors will converge to the true inflation processes. Thus, after a sufficiently long history, the forward premium at any horizon will eventually reflect the expected inflation differential.⁷ Since all the posterior beliefs converge to the true inflation profiles, it is appropriate to speak of “the inflation differential,” and the most patient classes of agents will set the prices of domestic and foreign bonds. Then covered interest parity will insure that the expected depreciation of the domestic currency reflects the domestic inflation differential.

During the early periods, the forward premium is determined by the configuration of priors and by the (stochastic) initial realizations of the inflation. In general, the forward rate is not a Martingale. The easiest way to see this is to note that each agent’s valuation of any bond depends upon both his priors and the history of inflation in both countries. The expected value of any stream of income—and thus the forward rate at all horizons—is not independent of history. Hence the spot exchange rate and the forward premium will be correlated in the early periods. A general description of the forward premium is quite involved, and it is appropriate to consider two separate cases. I will first analyze the case with one representative agent in the world economy. Then I will discuss the general case with several agents.

4.1. One representative agent

Consider the situation in which one agent has prior beliefs $\rho(\theta)$ and $\rho^*(\theta^*)$ about the processes driving domestic and foreign inflation. There is no reason to require these beliefs to be unbiased, and there are no simple restrictions that might be imposed upon their correlation.

Assume that the representative agent has the same subjective expected inflation for the domestic and foreign central banks. Then, abstracting from Siegel’s paradox, expected depreciation will be zero since both domestic and foreign bonds will be discounted identically. Still, the early history of inflation will have a strong effect on the forward premium, especially if either $\rho(\theta)$ or

⁶ This assumption is not idle; it is the only way of ensuring that a risk-neutral agent is indifferent between holding domestic and foreign bonds after any history.

⁷ Still, the bond pricing formulae exhibit Siegel’s (1972) paradox.

$\rho^*(\theta^*)$ are diffuse. For example, if there is an initial (random) spell of low inflation in the home country, the domestic yields will drop and foreign exchange will trade forward at a discount. Since yields reflect expected inflation, early realizations of the inflation process are doubly important for bonds with a long horizon. First, they occur when the agent's beliefs have the least precision. Second, the movements in the price of long bonds are amplified since their time to maturity is distant. Thus the forward premium at long horizons will be quite volatile initially.

What if the subjective probabilities are biased? Then the country with the higher subjective expected inflation has a bond that trades at a steep discount and a correspondingly high yield. Hence that country's currency will trade at a forward discount. In the long run, the actual realizations of the inflation rate will reflect the true underlying monetary process. The econometrician will observe a secular change in the forward premium that is not justified by the actual historical inflation differential. If the original priors are quite diffuse, then the period of learning will be fairly rapid, and the forward premium will forecast the actual rate of depreciation of the spot exchange rate after only a short time. But if the (incorrect) prior beliefs about either country have a high degree of precision, then the econometrician would see a very long period during which the forward premium was a biased predictor of changes in the spot rate.

4.2. Several types of agents

If there are several heterogeneous agents in the world economy, then a fascinating theory of asset prices emerges. First, all of the elements of learning are still present. Second, the heterogeneity of beliefs also matters. In particular, the price of a bond now reflects both the subjective expected present value of its cash flow and the option value of reselling it after some future history. Hence a currency will trade at a forward premium if there is a wide diversity of opinion about that country's inflation process. Since foreign exchange is traded forward at one month, three months, six months, and a year, this option value is highest for forward rates at longer horizons. Also, the one-month forward rate will not include a component having to do with the heterogeneity of beliefs.

It is easiest to illustrate these ideas by imposing in the rest of this subsection that agents have conjugate prior beliefs about the actual inflation and have identical subjective discount factors. Assume that agent i has conjugate priors described by Dirichlet distributions with parameters $\alpha_i = (\alpha_{i,1}, \dots, \alpha_{i,k})$ and $\alpha_i^* = (\alpha_{i,1}^*, \dots, \alpha_{i,k}^*)$.⁸ Again, Π is the vector of possible domestic inflation rates and Π^* be analogous. Then this person's expectations about domestic and foreign inflation at the null history are:

$$\bar{\pi}_i(h^0) = (\alpha_i)' \Pi / \alpha_{i,0} \quad \text{and} \quad \bar{\pi}_i^*(h^0) = (\alpha_i^*)' \Pi^* / \alpha_{i,0}^*,$$

where $\alpha_{i,0} = \sum_{j=1}^k \alpha_{i,j}$ measures the precision of his beliefs about domestic inflation and $\alpha_{i,0}^* = \sum_{j=1}^{k^*} \alpha_{i,j}^*$ is analogous. This person's posteriors induce these inflation forecasts:

$$\bar{\pi}_i(h^t) = (\alpha_i + n(h^t))' \Pi / (\alpha_{i,0} + t) \quad \text{and} \quad \bar{\pi}_i^*(h^t) = (\alpha_i^* + n^*(h^t))' \Pi^* / (\alpha_{i,0}^* + t).$$

⁸ See DeGroot (1970), p. 174. This distribution is the natural conjugate for the multinomial, and it is a generalization of the beta distribution. The i -th agent believes that the prior probability of the k -th event is $\alpha_{i,k} / \alpha_{i,0}$, and he believes that the variance of this outcome is $\alpha_{i,k}(\alpha_{i,0} - \alpha_{i,k}) / (\alpha_{i,0})^2 (\alpha_{i,0} + 1)$. If $\alpha_{i,0} \approx 0$, then this agent has non-informative priors.

The first important fact is that if all the agents' priors have the same precision, then the agent who is initially most bullish about a country's inflation prospects will always be so. This agent will always hold that country's bonds, and there will be no option value inherent in bonds denominated in its currency at long horizons. Here's why. Let i be such that $\bar{\pi}_i(h^0) = \min_j \{\alpha'_j \Pi / \alpha_{j,0}\}$. Since all the priors about domestic inflation have the same precision, we may put $\alpha_{i,0} = \alpha_0$. But then

$$\bar{\pi}_i(h^t) = (\alpha_i + n(h^t))' \Pi / (\alpha_0 + t).$$

Hence the ranking of the agents' expected inflation forecasts does not change since they all observe the same history. Then a simple argument using backward induction from any terminal history shows that this agent will pay the most for the bond denominated in the domestic currency. Of course, the same is true for the class of agents that is most bullish about foreign inflation, even though the relevant precision about that process may be different. Thus no bond price will have any option value inherent in it, and the forward rate at any horizon will reflect a simple learning process.

This observation has important implications for the forward premium. Only when agents' prior beliefs are of different precision will a bond have a lower yield than that forecast by the most bullish group in the world economy.⁹ In other words, the higher moments of the priors matter in a model with limited short selling, an illustration of Morris's (1996) switching condition for this model. Thus, if people in the world economy have unbiased but heterogeneous beliefs about a country's inflation prospects, then its bonds will have a relatively low yield, and its currency will trade forward at a premium.

Now consider the polar case where at least two agents' prior beliefs have different precision but all the agents have the same initial forecasts for domestic inflation. Thus $\bar{\pi}_i(h^0) = \bar{\pi}$ and there is a class of agents with the least precise beliefs; let these priors have precision $\alpha_{i,0}$. Likewise, the agents with the most precise beliefs have priors with precision $\alpha_{j,0} > \alpha_{i,0}$. A second important fact is that only these two classes of agents will ever hold the domestic bond, and there is a simple way of describing who holds these bonds when. The agents with the most precise priors hold domestic bonds if and only if the history has been such that $\Pi' n(h^t) / t > \bar{\pi}$. In other words, the agents with the most precise prior beliefs hold domestic bonds when average domestic inflation has been high, and those with the least precise beliefs hold them when it has been low.

Here's why this fact is true. One can always write the posteriors as:

$$\bar{\pi}_i(h^t) = \frac{\alpha_{i,0} \bar{\pi}}{\alpha_{i,0} + t} + \frac{t(\Pi' n(h^t) / t)}{\alpha_{i,0} + t}.$$

Thus anyone's posterior beliefs are a weighted average of the common prior expected inflation and the commonly observed average history of domestic inflation. Now consider the partial derivative:

$$\partial \bar{\pi}_i(h^t) / \partial \alpha_{i,0} = t \frac{\bar{\pi} - n(h^t)' \Pi / t}{(\alpha_{i,0} + t)^2}.$$

⁹ A colleague's comments helped me hammer this point home more forcefully. Imagine testing my ideas using survey data on inflation expectations. It is *not enough* to calculate the dispersion of point forecasts among different classes of traders. Instead, it is necessary to get data on how sure each agent was about his or her forecast. Since inflation surveys typically do not collect this kind of information, one must make a strong statistical assumption linking dispersion of forecasts among agents with the degree to which some of them felt that their subjective beliefs were imprecise.

Thus this posterior is increasing if and only if $\bar{\pi} - n(h^t)'II/t > 0$. In other words, when inflation has been low, the agents with the least precise beliefs are most bullish about domestic bonds; otherwise the agents with the most precise beliefs are most bullish. An implication is that even if average inflation converges in probability toward the common prior $\bar{\pi}_i(h^0) = \bar{\pi}$, there will almost surely be trade in domestic bonds after any finite horizon.

This discussion provides a simple explanation of the “forward premium anomaly.” If agents have heterogeneous beliefs, then the forward premium will depend upon the stochastic history of the world economy—even when everyone has unbiased priors about the inflation differential and expected movements in the spot exchange rate. Consider the simple case when the foreign inflation process is known and thus everyone has perfectly precise and unbiased prior beliefs about foreign inflation. Assume also that there are two types of agents with unbiased beliefs about domestic inflation; one group has very precise priors and the other has imprecise ones. At the null history, everyone expects the spot exchange rate to depreciate according to the common expected inflation differential. After a history of high home inflation, only people with precise priors would hold home bonds. Thus short-term interest rates—based upon asset prices for which the option value of eventual resale is negligible—would still reflect the actual expected inflation differential, and the one-period forward premium would not be correlated with future depreciations. But after a history of low home inflation, the agents with imprecise priors hold domestic bonds and the interest differential would be lower than the actual expected inflation differential. Thus the one-period forward premium would indicate an expected appreciation that would not happen on average. Hence there is a negative correlation between actual depreciations and the one-period forward premium.

Thus there is an ineluctable interaction between each agent’s expectations about the inflation processes and the degrees of precision that characterize his beliefs. Also, the volatility of the forward premium depends upon the precision of these prior beliefs. When there is less precision, there is greater volatility of bond prices and a larger reaction of the forward premium at long horizons. This premium will exhibit conditional heteroscedasticity since periods of high volatility are bunched together as agents are learning about the true inflation processes in the initial periods of the world economy.

The volatility of the forward premium will also be higher if there are diffuse priors about a bond that has low coupon payments. Since such a bond has a relatively long duration, small changes in expected inflation have large effects on yields. Thus the forward premium will move significantly with the advent of news relevant to forecasting inflation.

This model of the forward premium also has important implications for the “peso problem,” first described in an analytical framework by Krasker (1980). When will a country’s currency trade at a forward discount for a sustained period, even if the econometrician has observed no large depreciation of the spot rate? One obvious possibility is that agents anticipate a large devaluation that occurs with a small probability; in this model, that notion corresponds to an element in the support of a country’s inflation process that is very large but which may have small weight. But it is obvious that the “peso problem” endures only when all classes of agents have high precision about a small probability event, an unlikely situation indeed. Instead, it is quite plausible that people have diffuse priors about the monetary policy of a central bank undertaking a new regime of price stabilization. This diversity of beliefs would tend to keep the forward currency strong, and learning would tend to undercut any initial fear of hyperinflation. Thus the forward discount reflects both the lack of confidence in the inflation reduction scheme and the degree of conformity in traders’ beliefs.

5. The model and the data

This section accomplishes three goals. First, it describes the data from an important period in recent monetary history during which the dollar traded forward at a discount, even though it continued to appreciate on the spot market for almost five years. Second, it shows in detail how the model was calibrated. To the best of my knowledge, this is one of the first attempts at calibrating a model in which the diversity of prior beliefs has substantial empirical bite. Third, it actually *estimates* the model in two different ways and shows that diversity of prior beliefs may well characterize the data. Again, I believe this is the first time that an economic model in macroeconomics or finance with heterogeneous beliefs has been estimated and then tested.¹⁰

5.1. Data on the dollar–mark forward premium and the relevant inflation rates

Fig. 1 presents monthly data on the mark forward premium at different horizons during the first half of the 1980s. This was a period during which it was common for the dollar to trade forward at a discount, although there was a continued appreciation of the dollar until after the Plaza Accord in September 1985. This episode of international monetary history gave rise to a large literature on the foreign exchange risk premium. Two facts are salient. First, the dollar was trading forward at a discount during this entire period. Second, the actual change in the spot exchange rate was an order of magnitude larger than the forward premium, and the forward premium was the wrong sign during most of this period. Two other facts are worth emphasizing. First, the forward premia at different horizons are highly correlated. Second, all the premia were much more volatile at the beginning of this period, when Paul Volcker became chair of the Federal Reserve Board, than at the end, when a successful disinflation had been undertaken.

The next step is to gather data on the actual inflation rates of Germany and the United States during those five years.¹¹ The median rates of consumer price inflation, calculated from the 60 annualized monthly changes, were 3.2% in Germany and 3.8% in the United States. The data generating processes for these time series presumably have continuous support, but the model is analytical and computationally tractable only for a multinomial distribution. Indeed, the model implies that all the moments of the processes describing beliefs about inflation matter, and I had to make a choice about how best to model the inflation processes using multinomial

¹⁰ Using proprietary survey data, Elliott and Ito (1999) demonstrate that there is a diversity of beliefs in these markets. But they do not estimate an *economic* model where diversity matters. Chavas (2000) studies the U.S. beef market and shows that a significant fraction of suppliers have naïve expectations about price formation, but he did not develop a full-fledged model of informational heterogeneity. In his own vocabulary, traders “within each information group” are rational, but they are implicitly blissfully unaware of the forecasters of other types of suppliers. Baak (1999) studies the same market and uses a different econometric technique to estimate the fraction of “boundedly rational” agents. Again, this model of a market is not really fully specified because it lacks an explicit description of what constitutes common knowledge among all agents.

¹¹ I used the BLS series for all urban consumers (all items) not seasonally adjusted for American prices. The data on German prices are completely analogous, and they come from the Statistisches Bundesamt Deutschland. The spot exchange rate and forward premium were given to me in private correspondence by Nelson Mark. They were originally weekly data, and I chose the first week of each month to constitute the relevant monthly data. The interested reader will find all the data used in this paper at http://economics.sbs.ohio-state.edu/efisher/Dollar_Mark_Historical.xls. They span the period from March 1973 (the beginning of the modern era of floating exchange rates) through December 1985 (three months after the Plaza Accord).

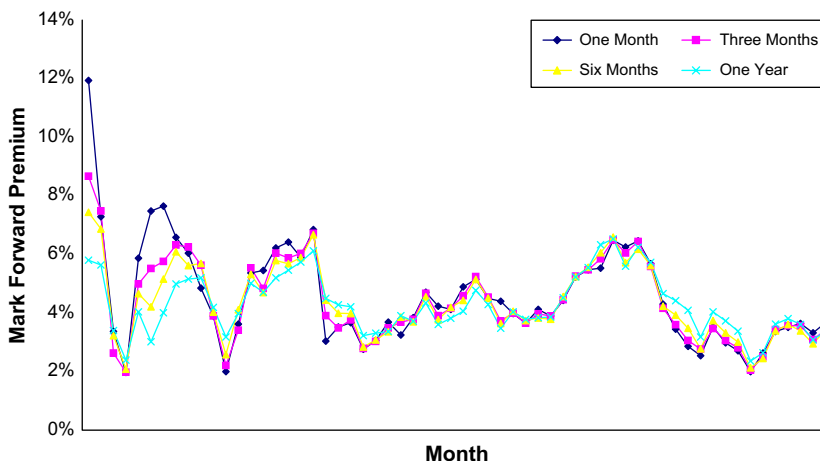


Fig. 1. The data (January 1981–December 1985).

distributions with discrete supports. Truth in advertising dictates that I should emphasize again that the assumption of independently and identically distributed data generating processes is probably at least as problematic as assuming discrete supports. The correlation between current inflation and inflation in near months has serious implications for the efficiency of the Method of Simulated Moments estimator.

Table 2 shows a tractable histogram describing the distributions of the actual inflation rates. These data were used in the simulations.

Why did I choose histograms with five bins? One needs a model with at least three bins to have independent measures of the first two moments of the inflation process, and I thought that higher moments also might matter. In brief, I classified the actual inflation histories for the two countries using these discrete data.¹² Even with this simple structure for domestic and foreign inflation, there are 4368 possible histories during the 12 months before a one-year forward contract matures.¹³ The simulations are computationally complex, although feasible on a personal computer with a Pentium chip.

It is worth emphasizing that nowhere did I use the data on the spot exchange rate in calibrating the model. Indeed, I am trying to build and evaluate a model of the forward premium, not one of the spot market where purchasing power parity is imposed after every possible history. I am thankful for this small mercy!

¹² There is a practical problem lurking in these discrete supports. The support for the American inflation process is more dispersed than that for German inflation. The model indicates that forward premia will depend upon inflation expectations along all possible histories, even those not realized in the data. Thus Siegel's paradox will tend to make dollar-denominated bonds more valuable, especially at long horizons. The implied American yields will be lower than would if the two inflation processes had identical discrete supports.

¹³ The degree of computational complexity becomes especially daunting for the Method of Simulated Moments estimators. In that case, I used 50 random histories and a numerical algorithm that searches through a two-dimensional parameter space; an iteration takes around 10 min on a fast laptop, and the algorithm converges in around 6 h from an initial starting value for the parameters.

Table 2
Discrete distributions of consumer price inflation, 1981 through 1985

Germany		United States	
Support (%)	Frequency (%)	Support (%)	Frequency (%)
-4	2	-5	2
0	10	0	8
4	70	5	67
9	15	11	18
12	3	14	5

5.2. How to simulate the model

Since the model has predictions for the forward premium at each horizon, I have 240 pieces of data that I am trying to fit with a parsimonious parameterization. The theory in Section 4 indicates that a model with only two classes of agents is already quite interesting, and considerations of scientific elegance and computational tractability led me to impose that restriction. When both agents have identical priors, the simulation captures the typical representative agent assumption in finance, albeit with an element of Bayesian learning. When the agents have different priors, the heterogeneity of beliefs is important, and measuring this effect is my primary empirical contribution.

Here are the broad strokes of how to simulate the model. Each simulation begins by specifying both agents' prior beliefs about the American and German inflation processes. In keeping with the spirit of Bayesian analysis, these beliefs need not even be proper distributions; indeed, any five non-negative numbers has an interpretation as a Dirichlet distribution with a certain precision. Then, sequentially for each of the 60 periods in the data, dollar-denominated bonds and mark-denominated bonds are priced, feeding the program the actual realizations of the (discrete) inflation processes as the five-year history after December 1980 unfolds. Of course, all four prior distributions are updated appropriately using the actual history. Each agent has his own subjective valuation of American and German yields, and the model prices American and German bonds after every history, using the relevant subjective evaluations about all possible future histories until the bonds mature. Finally, each simulation spits out the dollar premium at the relevant four horizons in every period as the actual history unfolds. The simulations all impose that the discount factor for each class of agents is 0.97 and that domestic and foreign assets are both zero-coupon bonds.

5.3. Method of Simulated Moments estimation

The model with two representative agents has 22 parameters: each agent has five parameters that describe prior beliefs about the American inflation process, five others that characterize the German inflation processes, and a subjective discount factor. The model is a complicated mapping from these parameters to predictions about the data.

It is appropriate to keep the estimation of the model as simple as possible, for reasons of both analytical elegance and empirical tractability. First, I imposed that the two agents had identical discount factors $\delta_1 = \delta_2 = 0.97$; it is notoriously difficult to estimate these parameters accurately, and I am already conducting an unorthodox empirical analysis. Second, I examined only two kinds of prior beliefs. Table 3 summarizes the non-sample restrictions that I have imposed.

Table 3

Non-sample information and estimated parameters

	First class of agents	Second class of agents
Discount factor	$\delta_1 = 0.97$	$\delta_2 = 0.97$
<i>Case of strongly rational expectations</i>		
Beliefs about American inflation	$\alpha_1 = \beta_1 \times (0.02, 0.08, 0.67, 0.18, 0.05)'$	$\alpha_2 = \beta_2 \times (0.02, 0.08, 0.67, 0.18, 0.05)'$
Beliefs about German inflation	$\alpha_1^* = \beta_3 \times (0.02, 0.1, 0.7, 0.15, 0.03)'$	$\alpha_2^* = \beta_4 \times (0.02, 0.1, 0.7, 0.15, 0.03)'$
<i>Case of Bayesian adaptive expectations</i>		
Beliefs about American inflation	$\alpha_1 = \beta_1 \times (0, 0.05, 0.365, 0.365, 0.22)'$	$\alpha_2 = \beta_2 \times (0, 0.05, 0.365, 0.365, 0.22)'$
Beliefs about German inflation	$\alpha_1^* = \beta_3 \times (0.02, 0.23, 0.47, 0.2, 0.08)'$	$\alpha_2^* = \beta_4 \times (0.02, 0.23, 0.47, 0.2, 0.08)'$

Two separate cases were estimated by the Method of Simulated Moments. First, I imposed strongly rational expectations about the actual German and American inflation processes; this is a very strong form of rational expectations because it assumes that the agents know the entire histograms of these processes, except for a multiplicative constant that is interpreted as the precision of the prior belief. Thus one class of agents was given this prior for American inflation $\alpha_1 = (0.02, 0.08, 0.67, 0.18, 0.05)'$ and this prior for German inflation $\alpha_1^* = (0.02, 0.1, 0.7, 0.15, 0.03)'$. These priors correspond to the actual historical inflation processes that occurred from January 1981 through December 1985, and both have unitary precision.¹⁴ Imposing that the other class of agents had the same priors up to a multiplicative constant, I used the Method of Simulated Moments to estimate the two free precision parameters. In sum, I am imposing that $\beta_1 = \beta_3 = 1$ and then estimating β_2 and β_4 in the top half of Table 3. In the second case, agents had “adaptive expectations”; their prior beliefs were based upon inflation histories during the 93 previous months from March 1973 to December 1980, the entire modern era of floating exchange rates. Hence in this second case, one class of agents was given these priors for American inflation $\alpha_1 = 93 \times (0, 0.05, 0.365, 0.365, 0.22)'$ and these priors about German inflation $\alpha_1^* = 93 \times (0.02, 0.23, 0.47, 0.2, 0.08)'$. The second class of agents had the same beliefs, but I again estimated two free precision parameters. Now I am imposing that $\beta_1 = \beta_3 = 93$ and then estimating β_2 and β_4 in the bottom half of Table 3.

What is a good benchmark against which to judge the model? The simplest model of the forward rate imposes covered interest parity. Then the Fisher equation and the assumption of real interest rates equalization imply that:

$$f(h^t, T - t) = E[(\pi_T - \pi_T^*) | h^t],$$

where $E[\pi_T | h^t]$ is the expected (annualized) domestic inflation rate between t and T . This expectation is taken with respect to the history h^t , and the expectation $E[\pi_T^* | h^t]$ is analogous for the foreign country. Imposing that realized ex post inflation is the proper proxy for expected inflation, Fig. 2 graphs these simple predictions. The benchmark prediction at a one-month horizon is more volatile than that at 12 months because monthly price changes can be quite variable.

¹⁴ The reader is reminded that beliefs with a precision of 1 are very diffuse. Indeed the first monthly observation is given as much weight as the prior beliefs. On the other hand, beliefs of precision 100 imply that the entire history of 60 months is not more important than the priors brought to the market.

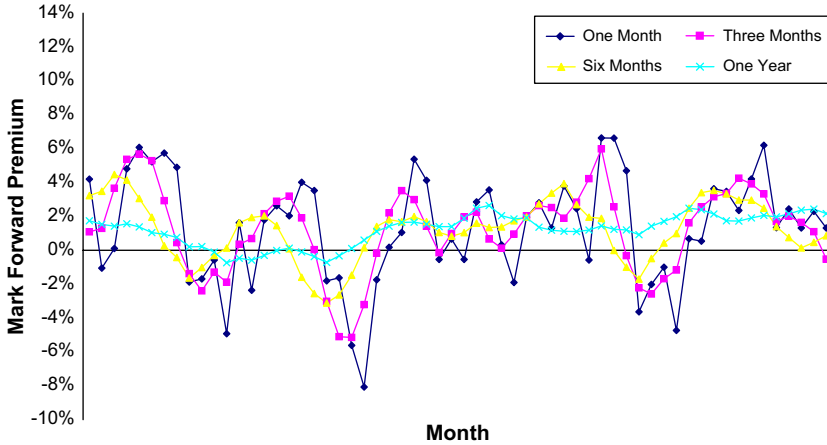


Fig. 2. Simple predictions (January 1981–December 1985).

What is a simple measure of a model’s goodness of fit? Let $y = (y'_1, y'_2, y'_3, y'_4)'$ be the stacked vector of the 240 observations on the forward premium at the four different horizons for the five years under consideration, and let $\hat{y} = (\hat{y}'_1, \hat{y}'_2, \hat{y}'_3, \hat{y}'_4)'$ be the corresponding predicted values. If one estimates k parameters, then the adjusted R^2 is an adequate summary measure. This statistic gives equal weights to forecasts at each horizon. Its value for the benchmark model—in which no parameter is estimated—is 0.22; thus actual ex post inflation explains very little of the variation in the actual forward premium, a fact that should be of no surprise to the empirically oriented reader.

Table 4 gives the estimates and model-fit statistics from the Method of Simulated Moments. The Appendix gives my exact technique. Since the computation is very time consuming, I could only estimate two parameters. Because the Quasi-Newton numerical algorithm used for minimization searches over all of two-dimensional real space, I estimated the model in logarithms. That transformation makes sense of negative estimates of the precision parameters.

Four facts in Table 4 are worth emphasizing. First, even with only two free parameters, the model performs much better than the benchmark; this is not much of a surprise since covered interest parity is a real straw man. Second, imposing adaptive expectations—those based upon recent history—fits the data better than imposing strongly rational expectations. Third, there is some very weak evidence that the heterogeneity of prior beliefs does seem to matter. Fourth, the model’s two parameters are estimated very imprecisely, probably because there is much correlation in the inflation data and the numerical techniques for estimating the relevant gradients are very imprecise in a model that is so highly non-linear. I conclude this subsection by reiterating

Table 4
Method of Simulated Moments estimates, exponential transformation (standard errors are in parentheses)

	Strongly rational expectations	Bayesian adaptive expectations
$\hat{\beta}_1$	$\beta_1 \equiv 0$	$\beta_1 \equiv \ln(93)$
$\hat{\beta}_2$	$-18.6 (5.6 \times 10^7)$	$24.2 (1.2 \times 10^8)$
$\hat{\beta}_3$	$\beta_3 \equiv 0$	$\beta_3 \equiv \ln(93)$
$\hat{\beta}_4$	$14.7 (4.3 \times 10^6)$	$-10.1 (2.1 \times 10^4)$
Adjusted R^2	0.59	0.81

Table 5

Non-linear least squares estimates, exponential transformation (corrected asymptotic standard errors are in parentheses)

	Strongly rational expectations	Bayesian adaptive expectations
$\hat{\beta}_0$	-3.33 (0.0017)	-3.40 (0.0018)
$\hat{\beta}_1$	-9.93 (19.0)	0.678 (1.43)
$\hat{\beta}_2$	-9.42 (0.0027)	-5.964 (0.0405)
$\hat{\beta}_3$	-7.27 (0.00011)	-1.83 (0.055)
$\hat{\beta}_4$	-23.7 (0.00081)	0.885 (0.587)
Sum of squared errors	0.053	0.041
Wald test for homogeneity of beliefs (the critical value is 9.21 for a test of size 1%.)	4.41×10^8	39.44

that a plausible model of the forward premium can be calibrated and then estimated, and it fits the data at least as well as the usual benchmark. There is weak preliminary evidence that the diversity of prior beliefs does indeed matter.

5.4. Non-linear regressions

Since the model is a complicated mapping from the parameters to predictions about the data, it is entirely appropriate to consider a non-linear regression. Let

$$y_t = f(x_t, \beta) + u_t$$

be such a specification. Here y_t is the forward premium, x_t includes the history of inflation that is in the information set for the relevant forward premium, and β is a vector of parameters to be estimated. The interpretation of the error term u_t is twofold. First, the discrete inflation processes impose aggregation errors. Second, a parsimonious parameterization will inevitably leave out some important factors that do indeed explain the forward premium.

Thus my second group of regressions estimates the following model

$$y_t = \beta_0 + f(x_t, \beta_1, \beta_2, \beta_3, \beta_4) + u_t,$$

where a constant has been included to make sure that the error terms have the proper location and where the other four parameters are the precision of the prior beliefs of a class of agents about the American and German inflation processes. The Appendix describes the exact method I used for the non-linear estimation.¹⁵

Table 5 reports the estimates, their standard errors, and the Wald statistics based upon the null hypothesis of homogeneous prior beliefs. Several comments are in order. First, these estimates are much better than those based upon the Method of Simulated Moments. This is true in part because the econometric model includes a constant term. This term is estimated quite accurately; the forward premium was roughly $\exp(-3.33) = 3.58\%$ in the rational expectations case and $\exp(-3.25) = 3.88\%$ in the adaptive expectations case. Second, the estimated

¹⁵ Starting from a random initial condition, I used Gauss's Quasi-Newton minimization routine to minimize the sum of squared errors. The objective function is flat in a neighborhood of non-linear least squares estimates, and the model is identified only up to two symmetric classes of beliefs. I ran the program repeatedly until I was satisfied that I was in a neighborhood of a global minimum. Since the precision of anyone's belief has to be a non-negative number, I actually estimate the model in logarithms.

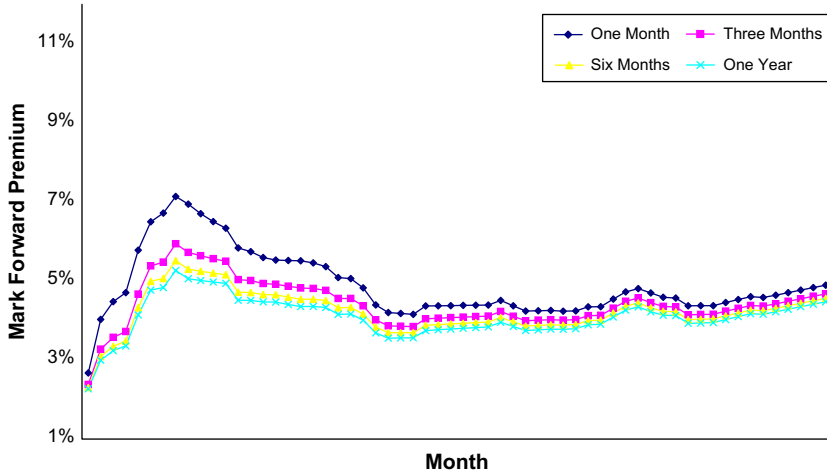


Fig. 3. Model predictions with strongly rational expectations (January 1981–December 1985).

coefficients show that a model of learning for all agents matters in both cases because none of the estimated coefficients is very large. The highest estimate of a precision parameter is for the first class of agents in the adaptive expectations case. One class believed that American inflation of the late 1970s was going to persist, but they did so only with a precision $\exp(0.885) = 2.42$. The meaning of this number is that one class of agents took around two and a half months to start to give significant weight to the in-sample data about inflation realizations. Third, the model based upon adaptive expectations does better than the one based upon strongly rational expectations. Thus the data have strong evidence in favor of the notion that the inflation history of the 1970s influenced beliefs in 1981.¹⁶

Figs. 3 and 4 show the model's predictions. Let me concentrate on Fig. 3 first; this is the model with strongly rational expectations. There are two salient characteristics of these predicted values. First, the initial periods show a very large forward premium. Second, after a while, the model predicts essentially a constant forward premium. The implication is that the diversity of prior beliefs matters the most in the initial periods; these are the periods during which the assets will trade back and forth between the different classes of agents. It is worth exploring the actual inflation history of 1981 to show why the model predicts a change in direction after September 1981. The first nine months of the U.S. inflation history have eight events in the 11% bin and one in the 14% bin. The 10th event falls in the 5% bin, and this causes investors to begin to change beliefs about American inflation radically. The 23rd month of the data also marks a milestone, and this is where there are twice as many events in the 0% bin (four) as in the 14% bin (two). *Thus the model seems to indicate that the market learned that the Volcker disinflation was credible in December 1982.* After this, the diversity of beliefs does not seem to matter.

¹⁶ I performed a Cox test for non-nested models. Using the null that the model with strongly rational expectations is true, I can resoundingly reject the estimates from the model with adaptive expectations. Under the alternative, I equally strongly reject the estimates from the model with strongly rational expectations. These two rejections are not surprising, since it is likely that the true model involves more than two classes of agents.

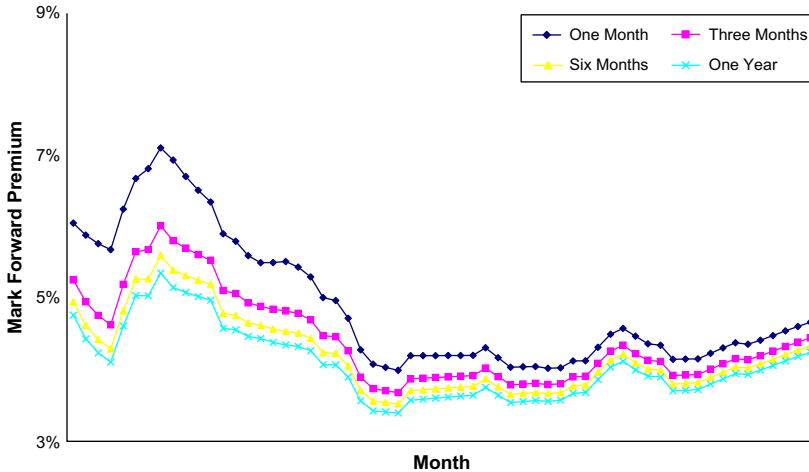


Fig. 4. Model predictions with Bayesian adaptive expectations (January 1981–December 1985).

Fig. 4 tells a similar story; this figure captures the predictions from the model with Bayesian adaptive expectations. The initial months in which the dollar traded forward at a discount reflect the inertial effects of higher inflation in the United States in the 1970s. The change in the forward premium after the summer of 1981 corresponds to the first realization of German inflation that falls in the 4% bin; in the first few months of 1981, there had been fairly significant German inflation. After that event, the heterogeneity of beliefs does not seem to matter much.

Let me conclude this subsection with a brief summary. This is the first time that any economic model with diverse prior beliefs has been brought to the data. There is evidence in favor of two salient facts: (1) learning matters for a proper model of the forward premium; and (2) a model with one homogeneous agent misses an important element of the data. The diversity of prior beliefs matters most in the early periods of the model; this makes a lot of sense because sooner or later the actual data will overwhelm any sensible prior beliefs that are brought into a new inflation regime. This is the fundamental empirical contribution of my work.

5.5. The model and the forward premium anomaly

The empirical discussion so far has avoided the obvious question: Do the predictions also exhibit the forward premium anomaly? In one sense, this query is a straw man because the model imposes the Procrustean requirement that the spot exchange rate satisfy purchasing power parity in every period. Still, there are good theoretical reasons for which the anomaly will arise in this model. Table 6 shows output from the regressions based upon the typical specification:

$$(e_{t+1} - e_t) = \gamma_0 + \gamma_1(f_t - e_t) + u_{t+1},$$

Table 6

Forward premium anomaly in the model's predictions (*t*-statistics values are in parentheses)

	$\hat{\gamma}_0$	$\hat{\gamma}_1$
Strongly rational expectations case	0.001 (0.92)	-0.004 (-31.4)
Bayesian adaptive expectations case	0.0005 (0.25)	0.014 (-25.9)

where all variables are in logarithms, $e_{t+1} - e_t$ is the depreciation of the spot exchange rate $f_t - e_t$ is the model's predicted one-period forward premium, and u_{t+1} is an error term not in the information set at time t . The numbers in parentheses in that table are the t -statistics based upon the null hypotheses that $\gamma_0 = 0$ and $\gamma_1 = 1$. The standard Newey–West correction has been applied to this regression, and I used a lag of 13. There is overwhelming evidence that these calibrations exhibit the forward premium anomaly, although the estimates of γ_1 are not as negative as in some other studies.

The main reason that the model's predictions show the forward premium anomaly is that, after about a year, the predicted forward premium becomes largely stable. Thus a high inflation realization doesn't move the forward premium very much, but it does cause a depreciation of the spot exchange rate because of the model's strong assumption about a constant real exchange rate. Hence there is little correlation between realizations of the spot exchange rate and the forward premium. Since the way the spot exchange rate is modeled is so artificial, I do not attach much importance to the fact that the estimated model exhibits the anomaly.

6. Conclusions

This paper has developed a new theory of the forward premium based upon a model that takes heterogeneity in financial markets seriously. The model has striking predictions for the forward premium; it shows that diverse beliefs about a country's inflation process make its currency trade forward at a premium in contracts whose horizons are greater than one month. The calibration of the model is perhaps plausible, but it predicts perhaps too much learning. Still, the notion that agents had diverse beliefs about monetary policy and doubted that inflation could be abated at the beginning of the 1981 is indeed intuitive.

My primary empirical contribution is that I actually estimate the precision with which different classes of agents in the world economy held plausible prior beliefs. There is evidence in these data that a model with one agent is just not an accurate description of world asset markets. This is the first time that an extension of the elegant models based upon Harrison and Kreps (1978) and Morris (1996) has been taken to the data, and my work shows that these authors were quite right to worry about diversity of beliefs in financial markets.

The model has many weaknesses. First and foremost, it is ludicrous to impose that the real exchange rate is constant. My only defense is that a good model of the spot exchange rate is left to those with superior analytical powers. Another important weakness is that the model assumes that inflation process in each country is independently and identically distributed. Inflation is obviously correlated between countries and across time, but this fact is difficult to incorporate into an analytically tractable model with Bayesian learning. The calibrations and estimates are suggestive but not exhaustive. Again, my defense is that there is no other study in international finance that takes a structural model with heterogeneous priors and Bayesian learning to the data. So this empirical work is just a first step.

In an important sense, my model is a better description of term structure than it is of the forward premium.¹⁷ Most empirical analyses of the yield curve show that the expectations hypothesis of the term structure is not true; the spreads between certain long rates and short rates do not predict future short rates. Rudebusch (1995) argues that the monetary authority's policy distorts bond prices at the short end of the yield curve. My model offers an alternative avenue

¹⁷ This paragraph was inspired by a referee's comments.

worthy of exploration; future researchers can analyze the effect of heterogeneity of beliefs on the yield curve in a national bond market.

Perhaps this paper will spur other researchers in international finance to investigate models with heterogeneous beliefs. It is remarkable that the “forward discount anomaly” can be explained so easily in a model with risk-neutral agents. It is essential that we economists be careful in our interpretations of the notion of rational expectations as an equilibrium concept. Models with heterogeneous prior beliefs are more general than the usual ones with one representative agent. The equilibria described in this paper all converge to the “rational expectations equilibrium” if agents’ priors are well behaved. And the typical model used in international finance is a special case of the one that has been explored; after all, one can always impose that everyone’s perfectly precise priors agreed with the actual distribution of inflation. But informational heterogeneity and limited short selling surely characterize actual financial markets. So it would be nice to continue building models that incorporate these obvious facts.

Acknowledgement

The author is grateful for the National Science Foundation Grant 9870874. He thanks an anonymous referee, Paul Evans, Eric van Wincoop, Jan Groen, Doug Steigerwald, and participants at the Midwest Macroeconomics Meetings, the Midwest International Economics Group, the Bank of Canada, The Ohio State University, the Federal Reserve Bank of Minneapolis, the University of Victoria, Louisiana State University, and the TAFI conference at San Juan for comments on earlier drafts. The data and the Gauss code used for all the computation and estimation are available upon request.

Appendix

Here is how I constructed the Method of Simulated Moments estimators. I used the fact that the order of the inflation events in a Bayesian model should not matter since I have assumed the natural conjugate priors. Hence, any history that satisfies the final empirical distribution of inflation events in Germany and the United States was equally likely. (Again, I am relying heavily on the assumption of independent inflation events in each month!) So I drew 50 such random five-year histories $\{\tilde{h}^{60}(1), \dots, \tilde{h}^{60}(50)\}$, each element consisting of an entire five-year history of inflation for the two countries.

Then I essentially followed the technique given in Davidson and Mackinnon (2004, chapter 9.6). Let $\beta = (\beta_2, \beta_4)'$ be the vector of precisions to be estimated from Table 3, and let $y_{240 \times 1}$ be the appropriately stacked vector of data on the forward premia. It is convenient to write $z_{240 \times 1}$ as the vector of squared values of these data. For a given vector of parameters $\beta = (\beta_2, \beta_4)'$, each simulated history $s \in \{1, \dots, 50\}$ gives predictions for these data $\hat{y}_{240 \times 1}(\beta, s)$ and their squared values $\hat{z}_{240 \times 1}(\beta, s)$. Let

$$m_y(\beta) = (1/50) \sum_{s=1}^{50} \hat{y}_{240 \times 1}(\beta, s) \text{ and } m_z(\beta) = (1/50) \sum_{s=1}^{50} \hat{z}_{240 \times 1}(\beta, s)$$

be the predicted moments for a given vector of parameters. Now consider the natural weighting matrix $w = I_2 \otimes \hat{i}$ and the quadratic form

$$Q(\beta) = \begin{bmatrix} y - m_y(\beta) \\ z - m_z(\beta) \end{bmatrix}^T w \begin{bmatrix} y - m_y(\beta) \\ z - m_z(\beta) \end{bmatrix}.$$

The Method of Simulated Moments estimator is the vector $\hat{\beta}$ that minimizes $Q(\beta)$.

Now let $\nabla m_y(\hat{\beta})$ and $\nabla m_z(\hat{\beta})$ be the two (row-vector) gradients evaluated at the estimate, and write $\Omega = \Sigma_{yz} \otimes I_{240}$, where Σ_{yz} is the variance–covariance matrix of the data $(y, z)'$. Finally

$$v(\hat{\beta}) = \left[w^T \begin{bmatrix} \nabla m_y(\hat{\beta}) \\ \nabla m_z(\hat{\beta}) \end{bmatrix} \right]^{-1} w^T \Omega w \left[\begin{bmatrix} \nabla m_y(\hat{\beta}) \\ \nabla m_z(\hat{\beta}) \end{bmatrix}^T w \right]^{-1}$$

gives the variance–covariance matrix of the Method of Simulated Moments estimator.

Here is how I did the non-linear estimation. I followed the procedure in Judge et al. (1982, chapter 24). The first step is to minimize the sum of squared errors between the model's predictions and the data; I used the data at all four horizons because the model has predictions for each horizon, given the inflation history. The second step is to evaluate the gradient $\nabla f(x_t, \beta) = (\partial f(x_t, \beta) / \partial \beta_0, \dots, \partial f(x_t, \beta) / \partial \beta_4)$ at the estimated coefficients and each of the 240 data points. The inner product of the resulting 240×5 matrix gives the variance–covariance matrix of the estimated parameters. Because this is a model of Bayesian learning, the error terms in this model are not independent across time, nor are they independent across the different horizons at which foreign exchange is traded forwarded. So the third step corrected the standard errors from the non-linear least squares estimates using the Newey–West (1987) correction. The last step is to test the hypothesis that there is no heterogeneity of prior beliefs in these data. It is natural to use a Wald test based upon the statistic $W = (R\hat{\beta})' (RVR')^{-1} (R\hat{\beta})$, where

$$R\beta = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} \beta = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

captures the two linear restrictions stating that the precisions of beliefs about American inflation are identical and so are the precisions having to do with German inflation. The variance–covariance matrix $V = [\nabla f(x, \hat{\beta})' \nabla f(x, \hat{\beta})]^{-1} G [\nabla f(x, \hat{\beta})' \nabla f(x, \hat{\beta})]^{-1}$, where $\nabla f(x, \hat{\beta})$ is the “stacked” 240×5 gradient evaluated at the estimates and the data and G is the Newey–West correction, constructed from correlations between $\nabla f(x, \hat{\beta})$ and the data at various lags.

References

- Baak, S.J., 1999. Tests for bounded rationality with a linear model distorted by heterogeneous expectations. *Journal of Economic Dynamics and Control* 23, 1517–1543.
- Chavas, J., 2000. On information and market dynamics: the case of the U.S. beef market. *Journal of Economic Dynamics and Control* 24, 833–853.
- Davidson, R., Mackinnon, J.G., 2004. *Econometric Theory and Methods*. Oxford University Press, Oxford.
- DeGroot, M.H., 1970. *Optimal Statistical Decisions*. McGraw-Hill, New York.

- Elliott, G., Ito, T., 1999. Heterogeneous expectations and tests of efficiency in the Yen/Dollar forward exchange rate market. *Journal of Monetary Economics* 43, 435–456.
- Engel, C., 1996. The forward discount anomaly and the risk premium: a survey of recent evidence. *Journal of Empirical Finance* 3, 123–192.
- Fisher, E.O'N., 2003. Explaining bubbles in experimental asset markets. Unpublished manuscript, Ohio State University.
- Fisher, E.O'N., Kelly, F., 2000. Experimental foreign exchange markets. *Pacific Economic Review* 5, 365–388.
- Harrison, J.M., Kreps, D.M., 1978. Speculative investor behavior in a stock market with heterogeneous expectations. *Quarterly Journal of Economics* 92, 323–336.
- Judge, G.C., Hill, R.C., Griffiths, W., Lutkepohl, H., Lee, T.C., 1982. *Introduction to the theory and practice of econometrics*. John Wiley and Sons, New York.
- Krasker, W.S., 1980. The 'Peso Problem' in testing the efficiency of forward exchange markets. *Journal of Monetary Economics* 6, 269–276.
- Lewis, K.K., 1989. Changing beliefs and systematic rational forecast errors with evidence from foreign exchange. *American Economic Review* 79, 621–636.
- Milgrom, P., Stokey, N., 1982. Information, trade and common knowledge. *Journal of Economic Theory* 26, 17–27.
- Morris, S., 1996. Speculative investor behavior and learning. *Quarterly Journal of Economics* 111, 1111–1134.
- Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–708.
- Rudebusch, G.D., 1995. Federal reserve interest rate targeting, rational expectation and the term structure. *Journal of Monetary Economics* 35, 245–274.
- Siegel, J., 1972. Risk, interest rates, and the forward exchange rate. *Quarterly Journal of Economics* 86, 303–339.