NBER WORKING PAPER SERIES

EXCHANGE RATE DYNAMICS AND LEARNING

Pierre-Olivier Gourinchas
Aaron Tornell

Working Paper 5530

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 1996

We thank Daron Acemoglu, Olivier Blanchard, Ricardo Caballero, John Campbell, Rudi Dornbusch, Ben Friedman, Graciela Kaminski, John Leahy, Paul O'Connell, Julio Rotemberg, Jiang Wang, and seminar participants at MIT, Harvard and the NBER for useful discussions and comments. All errors are our own. This paper is part of NBER's research program in International Finance and Macroeconomics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper 5530
April 1996

EXCHANGE RATE DYNAMICS AND LEARNING

ABSTRACT

Interest rate expectations are essential for exchange rate determination. Using a unique Survey data set on interest rate forecasts from 1986 to 1995 for G7 countries, we find that interest rate shocks were significantly more persistent in sample than expected by the market. This is consistent with ff3's finding that changes in the forward rate reflect changes in exchange rate expectations. We then present a model of nominal exchange rate determination that rationalizes the forward discount puzzle and exhibits the delayed overshooting pattern found by ee: following a monetary expansion that reduces the domestic interest rate, there is a gradual depreciation of the exchange rate followed by a gradual appreciation several months later. Delayed overshooting results from (a) the interaction of *learning* about the current state of affairs, and the intrinsic dynamic response of interest rates to monetary shocks and (b) the discrepancy between the actual distribution of shocks in sample and its expectation by market participants. This discrepancy is consistent with rational expectations if either (a) there is a small sample or Peso problem or (b) the true structure of the economy evolves over time and agents are learning with some delay.

Pierre-Olivier Gourinchas
Department of Economics
MIT
50 Memorial Drive
Cambridge, MA 02139

Aaron Tornell
Department of Economics
Harvard University
Littauer M-6
Cambridge, MA 02138
and NBER

1. Introduction

This paper presents a nominal exchange rate model with optimizing agents and learning that rationalizes two anomalies of the foreign exchange market: the forward discount puzzle and the "delayed overshooting" of exchange rates in response to monetary shocks.

The forward discount puzzle has been the subject of a large body of empirical literature (see Hodrick (1988) and Lewis (1994) for surveys). Typically a persistent pattern of exchange rate appreciation (resp. depreciation) coexists with a positive (resp. negative) differential of the domestic over the foreign interest rates. This implies that excess returns in the foreign exchange market are partially predictable. Moreover, these predictable excess returns are time-varying.

The delayed overshooting puzzle has been uncovered by Eichenbaum and Evans (1995). These authors find that unanticipated expansionary shocks to US monetary policy are followed by (a) *persistent* reductions in US interest rates, and (b) a *gradual depreciation* of the dollar, followed by a *gradual appreciation* several months later. Figures 1a-d replicate the findings of Eichenbaum and Evans.[1] This hump-shaped exchange rate response is a violation of the rational expectations *overshooting principle* (Dornbusch (1976)) whereby the exchange rate should depreciate instantly, and then appreciate gradually towards its long run equilibrium value (see Figure 5). This *delayed overshooting* path is potentially consistent with predictable excess returns as there exists a time interval during which a positive interest differential coexists with an appreciating currency.[2]

Two classes of explanations have been put forth to rationalize the forward discount puzzle: time-varying risk premia and expectational errors (see Lewis (1994) and Frankel and Rose (1994) for surveys). Under the first explanation, fluctuations in the forward rate reflect changes in the risk premium. As Fama (1984) points out, under this interpretation, the risk premium must be more volatile than predictable excess returns. In equilibrium, the risk premium will fluctuate with relative asset supplies, conditional variances, and the intertemporal elasticity of substitution of consumption. However, as with the equity premium puzzle (see Mehra and Prescott (1985)), one has to invoke unrealistically high risk aversion coefficients in order to make the risk premium fluctuations implied by the data compatible with the low volatility of the above-mentioned variables.

Using survey data on exchange rate expectations, Frankel and Froot (1989) have decomposed predictable excess returns into their risk premium and expectational error components. Their results indicate that (a) almost none of the bias can be attributed to risk premium fluctuations and (b) changes in the forward premium reflect one for one changes in expected appreciation. Thus, expectational errors are responsible for most of the bias. Expectational errors may arise either when agents

[1] Appendix A describes the methodology. The results of Clarida and Gali (1994), Grilli and Roubini (1994) also point in the same direction.

[2] The time dimension of this phenomenon is worth emphasizing: the nominal exchange rate peaks 10 to 36 months after the initial shock.

have to learn about some unobservable shift in the economic environment.[3] Learning about a one-time unobservable shock to fundamentals in the foreign exchange market has been analyzed by Lewis (1989a) and (1989b). If a change in regime occurs, agents will gradually update their beliefs about the probability that the new regime is in place, generating systematic forecast errors during the transition. These learning models explain a significant part of the exchange rate mispredictions implied by the forward discount bias. However, they do not account for the fact that predictable excess returns do not appear to die out over time between regime switches.[4]

Since learning generates forecast errors that die out over time, models based on learning about a *one-time* change in regime cannot be expected to deliver a hump-shaped impulse response of the exchange rate to monetary shocks. Another force needs to be added in order to generate delayed overshooting. In this paper, we introduce such a force: in a stationary economy in which the equilibrium exchange rate is a function of current interest rates and expected future exchange and interest rates, hump-shaped dynamics may result from (a) the interaction of *learning about the duration of monetary shocks* (transitory vs. persistent) and the *gradual response of interest rates* to monetary shocks and (b) the *discrepancy* between the actual distribution of interest rate shocks in sample and its perception by market participants.

As a first step towards solving these anomalies, we analyze interest rate expectations. In order to do so, we assume that the economy is constantly hit by monetary shocks which can be temporary or permanent. Temporary monetary

[3] Another possibility is the presence of irrational traders as in Frankel and Froot (1986). In their model money managers make their decisions based on the advice from "fundamentalists", who use rational expectations, and from "chartists" that use *ad-hoc* forecasting rules. The expected exchange rate is then a weighted average of the expectations of chartists and fundamentalists. In their model fundamentalists do not take into account the presence of chartists. Assuming that the fundamentalists do recognize the existence of chartists implies strong restrictions on the exchange rate behavior. In this case the exchange rate must move exactly so as to induce fundamentalist to take the opposite side of chartists at any point in time. Suppose for instance that chartists are trend chasers: they buy more when the currency appreciates. Following a monetary *contraction*, the exchange rate must *depreciate*. Otherwise, both chartists and fundamentalists would want to increase their holdings of domestic currency and the excess demand would increase in the exchange rate. However, empirical results suggest that the exchange rate *appreciates* following monetary contractions.

[4] Another class of models in which agents have to learn about a regime shift is the so-called "Peso problem," whereby if an expected shift in regime does not materialize in sample, expectations will appear systematically biased to the econometrician. Kaminski (1993) shows that Peso problems can account for part of the forward discount premium in a model in which regime switches follow a Markov process. This class of models does not deliver a hump-shaped impulse response of the exchange rate to a monetary shock. Moreover, using option prices on Dollar-Deutschemarks between 1984 and 1993 to extract jump-expectations associated with shifts in regimes or bursting bubbles, Baily and Kropywiansky (1994) conclude that although there were significant jump-expectations, enough jumps were present in sample. Therefore peso problems are unlikely to have induced a major bias in exchange rate pricing.

shocks generate transitory changes in the interest rate, while permanent monetary shocks generate persistent interest rate changes that die out gradually. Agents only observe the realization of interest rates. This specification is flexible enough and can accommodate purely transitory processes as well as persistent ones. Key parameters of this process are the speed of convergence of the persistent component and the relative variance of transitory and persistent components. Our initial task consists in estimating the forecasting rule applied by market participants. We use a unique survey data set on interest rate expectations, published by *Currency Forecasters' Digest*, with monthly observations from 1986 to 1995 for G-7 countries. We find that (a) there is no evidence of a transitory component in the *sample* distribution, while (b) market participants implicitly assume a substantial share of the shocks to be transitory. This contrast is striking: while the estimated relative variance of transitory shocks in sample is close to 0, the relative share implicitly assumed by the market is often significantly larger than one, indicating that most shocks are believed to be transitory. We emphasize that this result need not be inconsistent with rational expectations. First, the sample realization of shocks may exhibit more persistent shocks than the ergodic distribution, equivalently, there might be a Peso problem. Second, the true data generating process may shift over time, and agents may be in the continual process of "catching up". According to this interpretation, at any point in time, agents do not know the correct parameters.

Using our framework, we then develop a model of nominal exchange rate pricing that rationalizes both predictable excess returns and the delayed overshooting puzzle. We consider a typical optimizing model with a representative agent, imperfect information and risk aversion, and derive the equilibrium exchange rate in *closed form.* The interest rate process is as described above. During each period agents estimate the probability distribution of future interest rates and exchange rates using the current realization of interest rates. Based on these assessments agents make their consumption and portfolio decisions in order to maximize expected utility. Lastly, the exchange rate adjusts every period in order to ensure that demand and supply of domestic bonds are equalized.

To illustrate how our model generates the delayed overshooting found by Eichenbaum and Evans consider a *persistent* increase in the domestic interest rate at time 0 followed by no other shock (conditional delayed overshooting). The interest rate differential jumps up at impact and returns gradually to its long run value. If agents know the exact composition of the interest rate shock, the exchange rate appreciates on impact, overshooting its long run level, and then depreciates gradually ensuring that uncovered interest parity holds. This is the standard interest rate effect characterized by Dornbusch (1976). However, if agents do not know the nature of the shock, a delayed overshooting path, as in Figure 5, may occur. This path reflects an additional learning effect. Since the actual sequence of interest rate differentials is greater than what agents expect, their assessment of the mean of future differentials goes up. This generates an increasing demand for domestic assets over time. This learning effect may dominate the interest rate effect, implying an appreciation over time in order to clear the increased demand for domestic assets. However, this appreciation cannot go on forever. As soon as market participants have gathered enough information from interest rate realizations, learning slows down, and the interest rate effect eventually dominates the learning effect. Since domestic interest rates decline toward their steady state level, the exchange rate must revert to a depreciating path at some point in time.

Uncertainty regarding the duration of the shock does not automatically generate conditional delayed overshooting. The key factors that determine whether it occurs are the speed of interest rate adjustment and the learning speed – that is, the amount of information conveyed by current interest rates. An economy converging more rapidly to its long run equilibrium is less likely to exhibit conditional delayed overshooting. Since convergence happens very quickly, permanent monetary shocks look like transitory ones. Changes in the informativeness of the interest rate signal have more complex effects. Conditional delayed overshooting only occurs for intermediate learning speeds. With high learning speed, beliefs converge almost instantly to the true value of the persistent component of interest rates. Conversely, if learning is too slow, the market belief increases very little at the time of the shock. Afterwards, although the market belief is updated upwards, the increase in demand for domestic assets is too small to dominate the interest rate effect and there is no delayed overshooting.

Although the experiment presented above captures the basic intuition of the paper, it is misleading along one important dimension. We described the response to a *persistent* shock. However, the empirical evidence presented in Eichenbaum and Evans (1995), Grilli and Roubini (1994) and Clarida and Gali (1994) does not control for the persistence of the shocks. As long as market participants use the correct distribution of shocks, there is no *unconditional* delayed overshooting, nor predictable excess returns. However, if persistent shocks are more frequent in sample than expected by agents, we find that *unconditional* delayed overshooting and predictable excess returns may occur. This case coincides with our findings using survey data on interest rates. We characterize unconditional delayed overshooting and show that empirical parameters for G-7 countries belong to the "delayed overshooting region". Given the lack of constraints imposed on the coefficients and the simplicity of our economy, these results suggest that learning about the current state of affairs and interest rate expectations are essential components of exchange rate determination.

Finally, it is important to stress that predictable excess returns can exist without delayed overshooting. In the last section, we extend the basic model by assuming that the interest rate process includes an ARCH (autoregressive conditional heteroscedasticity) component. This assumption captures the tendency in financial data for volatility clustering: large (small) shocks are followed by large (small) shocks. In this setup the larger the interest rate shock, the smaller the precision of the Bayesian updating. Thus, upon realization of a shock, agents' risk premium increases, as the perceived variance of the asset is larger. Over time, as the variance and the learning speed converge back to their equilibrium values, the extra risk premium disappears. This time varying risk premium generates predictable excess returns that vary over time. However, as long as there are no expectational errors, this does not generate unconditional delayed overshooting. We then use a two step procedure to estimate our model for G-7 countries from 1974 to 1992. In a first step, we estimate the interest rate differential process by *Maximum Likelihood.* The results strongly support the assumption of conditional heteroscedasticity. In a second step, using a Simulated Method of Moments, we estimate the model's remaining parameter. Having estimated our model, we use it to perform a Monte-Carlo study of deviations from Uncovered Interest Parity. We show that risk premium fluctuations related to learning account for a substantial fraction but not the entirety of the forward discount puzzle.

The structure of the paper is as follows. In section 2 we analyze the interest rates forecasts data. In section 3 we present the basic model and specify the conditions under which delayed overshooting occurs. Section 4 assumes a conditionally heteroscedastic process for the exchange rate and shows that this implies a time-varying risk premium. We present our conclusions in section 5. Lastly, appendix A replicates the methodology behind the impulse responses of Eichenbaum and Evans (1995) for a larger time horizon, appendix B describes in more details the empirical results of section 2, while appendix C contains all algebraic derivations.

2. Stylized Facts About Interest Rates Differentials

We start by analyzing interest rates expectation formation. We first propose a flexible multivariate representation of the interest rates differential process. Interest rates shocks can be temporary or persistent. Temporary shocks last one period. Persistent shocks decay slowly over time. This representation allows for any pattern of positive autocorrelation of interest rate differentials, depending of the relative variance of the two primitive shocks and the speed of convergence. We estimate the parameters of the data generating process for G-7 countries against the US interest rate on monthly data. We find that (a) there is strong persistence and (b) the transitory component is negligible. This indicates that the interest rate process is well approximated by an AR process, without significant moving average part. Using survey data on interest rate forecasts at 3, 6 and 12 months horizons, we then estimate the parameters of the market filter that generate those forecasts. The market filter that best replicates the interest rates forecasts exhibits (a) a higher degree of persistence of persistent shocks and (b) a large and significant transitory component.

2.1. Modelling the Interest Rate Differential

We adopt a state-space representation for the interest rate differential. The interest rate differential between any two countries, di_t is the sum of a persistent and a transitory component:

$$di_t = \mu + di_t^p + \nu_t \tag{2.1}$$

where μ represents a constant, di_t^p the unobservable persistent component and ν_t is the transitory innovation. The persistent component follows an $AR(q)$ process:

$$\lambda(L)\, di_t^p = \epsilon_t \tag{2.2}$$

with $\lambda(L) = 1 - \sum_{i=1}^{q} \lambda_i L^i$. We assume that the transitory and persistent innovations are independent and normally distributed with mean 0 and variance σ_ν^2 and σ_ϵ^2 respectively. This representation is flexible enough and can accommodate an $AR(q)$ as well as a white noise. Indeed, as we will show briefly, this representation is equivalent to a restricted ARMA process when agents do not observe individual shocks.

Following Dornbusch (1976) model we can give a monetary interpretation to this interest rate process when $q = 1$ and $\lambda = \lambda_1 < 1$. We can think of the transitory shock ν_t as a relative velocity shock, and of the persistent shock ϵ_t as caused by a permanent relative money supply shock. In the presence of sticky prices in the short run, a permanent reduction (increase) in the nominal money stock leads to an increase (reduction) in the domestic interest rate. As prices adjust over time, *real* money supply increases and the interest rate declines gradually until it reaches its steady state value. The speed of adjustment λ corresponds to the root of the Dornbusch model which is less than one. A *smaller* λ means *faster* convergence to long run equilibrium.[5] This response is consistent with the empirical finding of Eichenbaum and Evans (1995): a shock to the US money supply induces a persistent change in the US interest rate in the opposite direction.

A setup with temporary and permanent monetary shocks seems appropriate to capture investors' reactions to Federal Open Market Committee (FOMC) meetings.[6] One can think of each meeting as a monetary shock. Since investors only have access to the minutes of the meetings after a six weeks delay, they must conjecture whether the last decision -or lack of decision- of the FOMC will be persistent or transitory, and whether it reflects a response to inflationary pressures or concern for a weakening economy. A recent example may illustrate this point. On July 6 1995, after a two day FOMC meeting, the Fed announced a cut of 25 basis points in the Federal Funds Rate, to 5.75%, ending an eighteen months long period of rising interest rates. The market was somewhat caught by surprise, as some market participants expected interest rates to remain unchanged and the bond market reacted strongly with a 1 5/8 points increase in bond prices following the announcement.[7] This prompted a widespread revision of short term interest

[5] In Dornbusch (1976), the economy converges faster the lower the semi elasticity of money demand to interest rates, and the higher the semi elasticity of output to interest rates.

[6] See Batten, Blackwell, Kim, Nocera and Ozeki (1990) for a description of the operating procedures of major Central Banks.

[7] "Just a few weeks ago, financial markets were convinced that the Fed was on the verge of cutting rates. But recent economic data have led some Fed watchers to change their mind" Wessel (1995, July 5).

rates expectations, as traders expected further cuts, or in our terminology a persistent shock.[8] In the following days and weeks, however, it appeared increasingly apparent that the economy was still moving forward. Thus, expectations of further cuts decreased, and some even feared an interest rate increase, i.e. a transitory shock.[9] Furthermore, the minutes of the July 6 meeting, released on August 25, reveal that the Fed itself was surprised by the strength of the US economy. Despite being divided about the initial cut, it had planned further interest rates reduction in the weeks following the meeting.[10] Thus, transitory shocks may arise when the Fed acts on inaccurate forecasts or to reflect balance of power adjustments among the Open Market Committee members. Both elements are not observed by market participants who have then to infer the motivation behind recent policy decisions.[11]

Agents only observe the realization of the interest rates differential. They do not know whether a change in the interest rates differential is caused by a transitory (v_t) or persistent (ϵ_t) innovation. Similar processes have often been used in the learning literature, starting with Muth (1960) in his exploration of the link between rational and adaptive expectations. The best an agent can do is to recognize that the interest rate differential follows an $ARMA(q,q)$. Indeed, applying $\lambda(L)$ to (2.1), we obtain:

$$\lambda(L)\; di_t = \epsilon_t + \lambda(L)\; \nu_t = \phi(L)\, u_t \tag{2.3}$$

The right hand side is a moving average of order q: $\phi(L) = 1 + \sum_{j=1}^{q} \phi_j L^j$. The coefficients of the ARMA representation are restricted by the following system of equations, obtained by equating moments:

$$\sigma_\epsilon^2 + \left(1 + \sum_{i=1}^{q} \lambda_i^2\right) \sigma_\nu^2 = \left(1 + \sum_{i=1}^{q} \phi_i^2\right) \sigma_u^2 \tag{2.4}$$

$$\phi_i = -\lambda_i \frac{\sigma_\nu^2}{\sigma_u^2};\;\; 1 \le i \le q \tag{2.5}$$

[8] "Now speculation turns to the future.[...] Recent history suggests there is more to come; the Fed rarely changes direction without seeing a series of moves ahead. Every 'monetary-policy decision must be thought of as a first step along a path,' Vice Chairman Alan Blinder said in a speech in May." Wessel (1995, July 7)

[9] "After Mr Greenspan delivered his semiannual address to the House Banking Committee, the stock and bond market concluded the Fed hadn't any plans to cut interest rates futher, traders said. Disappointed, stock and bond investors responded by selling heavily." Rebello (1995, July 20)

[10] "Federal Reserve policy makers were deeply divided about their decision to cut short-term interest rates[...] last month, the Fed disclosed. In an indication that Fed officials, at least in July, expected further reductions in rates in the months ahead, the Committee also voted to lean toward lowering rates in the weeks following the meeting" Wessel (1995, August 28).

[11] Furthermore, the FOMC may grant authority to the Chairman to implement interest rates changes between meetings. Fed watchers must then be actively monitoring the New York Federal Reserve Bank's open market operations and reserve flows (see Wessel and Raghavan (1994, March 24)).

As can be checked immediately, the moving average part disappears if and only if there is no transitory component for the interest rate differential ($\sigma_{\nu}^2 = 0$). Since the original state-space representation has $q+2$ unknowns and $ARMA(q,q)$ estimation yields $2q+1$ parameters, this gives $q-1$ overidentification restrictions.[12] Defining the relative variance coefficient $\eta = \sigma_{\nu}^2/\sigma_{\epsilon}^2$, it is easy to check that the coefficients of the moving average part $\{\phi_i\}_{i=1}^{q}$depend only on η and the λ_i' s: $\phi_i = \phi_i\left(\{\lambda_j\}_{j=1}^{q}, \eta\right)$. The previous decomposition suggests that one way to estimate the process (2.1)-(2.2) consists in running ARMA processes of various orders, estimating $\{\phi_i, \lambda_i\}_{i=1}^{q}$, and inverting (2.4)-(2.5), imposing the $q-1$ overidentifying restrictions. There are three problems with this approach. First, this is not efficient if the model is well specified. Second, it is unclear how one should choose the order of the autoregressive part. This problem is particularly relevant when there is no transitory component so that the process follows an $AR(q)$. In practice, we picked the ARMA process minimizing Akaike's criterion. The results are reported in appendix B. In most cases, this identification procedure failed to detect a moving average component.[13] Lastly, the state-space representation is more appropriate to make exact finite sample forecasts. We will need such forecast to estimate the parameters of the distribution implicitly assumed by market participants.

The results from the identification procedure are reported in appendix B, Table B.1, B.2, and B.3. The results are consistent with our state-space representation and indicate that transitory components are small or absent.

2.2. Survey Data

While survey data on monthly money market rates are available for the US, and have been used in previous studies, we were unable to find similar survey forecasts for foreign countries.[14] Instead, we obtained interest rate forecasts from the *Currency Forecasters' Digest*, now called *Financial Times Currency Forecaster*. Contributors include multinational companies as well as forecasting services from major investment banks, i.e. the most active player on the fixed income and foreign exchange markets.[15] This monthly publication collects interest rates and their

[12]When $q=1$, the model is exactly identified. However, λ and ϕ must still be of opposite signs.

[13]Two exceptions are UK-US and Germany-US for the Euro-3 months rate (Table B.1). Both cases violate the restriction imposed by our assumptions. Moving average components are detected for France-Us and Italy-US using the 1 month money market rates.

[14]Froot (1989) uses quarterly data on the three months T-bill from 1969 to 1986 from the Nagan Bond and Money Market Letter. This dataset has also been used in Friedman (1980).

[15]The Forecasting services that contribute to the Currency Forecaster's Digest are: Predex, Merril Lynch, Mellon Bank, Harris Trust, Bank of America, Morgan Grenfell, Chase Manhattan, Royal Bank of Canada, Midland Montagu, Generale de Banque, MMS International, Chemical Bank, Union Bank of Switzerland, Multinational Computer Models, Goldman Sachs International, Business International, M. Murenbeeld, and Westpac Bank. The multinational companies that contribute are: General Electric, Du Pont, WR Grace, Allied Signal, Monsanto, Ingersoll-Rand, General-Motors, Data General, Eli Lilly, Aetna, American Express, Johnson and Johnson, Sterling Drug, Firestone, 3M, Union Carbide, Texaco, United Brands, SmithKline Beckman,

forecasts 3, 6 and 12 months hence for the prime rate, three months and one year Eurodollar-rates and ten year government bonds. It then reports a "market average" weighting individual respondents according to their relative importance.[16] The countries covered are Australia, Canada, France, Germany, Italy, Japan, Switzerland, UK and US. We will restrict ourselves to the G-7 countries Canada, France, Germany, Italy, Japan, UK and the US. The period for which interest rate forecasts are provided is 1986-1995. This dataset is unique in its coverage and consistency. We have not found any other source of interest rate forecasts prior to 1986 covering all G-7 countries.

We emphasize that one should be cautious when using survey data. First, there is probably no such thing as a "market expectation". Since Currency Forecaster's Digest aggregates individual forecasts according to some proprietary and non-disclosed rule, one should exert extra caution. We do not believe, however, that individual respondents have an incentive to misreport their forecasts. Each participant, in exchange for her own contribution which remains secret, gets back the market average, incorporating in an unspecified way this same contribution. This may induce more truthful revelation and mitigate incentives either to herd and discard one's private information or to report "extreme" forecasts. As an alternative to survey data, we could have instead used forward interest rates as implied by the term structure of interest rates. Under the rational expectation theory of the term structure of interest rates, the forward rates is equal to the expected future short-term interest rate. Froot (1989), using survey data on US interest rates finds that the expectation hypothesis fails at short maturities (less than 12 months), an indication that forward rates incorporate time-varying risk premium. Therefore, as a first attempt, it seems more appropriate to use directly survey data.

Prime rates are short term lending rates to preferred customers and are available from September 1987 to September 1995. They are at the bottom of the lending structure and reflect monetary policy as well as market structure in the banking industry. In particular, the prime is not a market rate. Thus, one would suspect that shocks to the prime rate might originate from other sources than monetary innovations. Similarly, banks may react with some delay to innovations in money market rates or T-bill rates. Since the difference between the money market rate and the prime determines the banking margin, banks might try to keep the prime constant for a while after a cut in the discount rate, so as to increase their margin. Market structure determines the extent of the discrepancy. For all these reasons, the prime is likely to be a noisy indicator of monetary policy. In addition, the prime is changed only infrequently. Figure $2a - g$ show the prime rate together with the money market rate.[17] It is clear that the prime is close enough to the market rate for the US, the UK, Canada and Japan. For France, Germany and Italy however,

American National Can, RJ Reynolds, Colgate-Palmolive, Warner-Lambert, Schering-Plough, Quaker Oats, Beatrice Foods, Hercules, Baxter Travenol, and Interpublic Group.

[16]Currency Forecaster's Digest does not disclose its aggregation rule nor individual forecasts.

[17]The money market rate is line 160b from the International Financial Statistics tape.

the prime exhibits substantial inertia and appears disconnected from the money market rates, at least for part of the sample, so that our modelling assumptions seem inappropriate. We run the estimation using the prime rate only for the US, the UK, Canada and Japan.

Instead of using prime rates, we can use 3-months rates on Euro deposits. These rates, and their forecasts at 3, 6 and 12 months are available from Currency Forecaster's Digest from August 1986 to October 1995.[18]. Euro rates are not likely to reflect immediately the stance of domestic monetary policy. First they have a longer maturity than money market rates. Second, these are offshore rates that can be shielded from country risk and may also reflect less accurately domestic monetary conditions. Figure $3a-g$ reports the euro rates (annualized) against the money market rates. It appears that in most cases the correlation is quite good. Therefore, we keep all G-7 countries in our sample.

2.3. Maximum Likelihood Estimation of the Interest Rate Process

We develop in this section the estimation procedure for our state-space representation, directly imposing the restrictions (2.4)-(2.5). This will prove useful when estimating the market parameters. We rewrite (2.1)-(2.2) as:

$$di_t = \mu + H'\xi_t + \nu_t \quad (2.6)$$
$$\xi_t = F\xi_{t-1} + \epsilon_t \quad (2.7)$$

where $\xi_t = \left(di_t^p, ..., di_{t-q+1}^p\right)'$, $H' = (1, 0, ..., 0)'$ is a qx1 vector. ξ_t is the state vector for the process, (2.6) the measurement equation and (2.7) the space equation. Defining the informations set $I_t = \{di_{t-j}, j \geq 0\}$ and $\hat{\xi}_{t+1|t} = E\left[\xi_{t+1}|I_t\right]$, we can derive the filter and the smoother.[19] Under the normality assumption, and assuming additionally that $\hat{\xi}_{1|0}$ is normally distributed, ξ_{t+1} is normally distributed conditionally on I_t, with mean $\hat{\xi}_{t+1|t}$ and variance $\hat{P}_{t+1|t}$. We can then write the conditional likelihood of i_{t+1} as:

$$\log f_{di_{t+1}|I_t}\left(i_{t+1}|I_t\right) \propto \log\left|H'\hat{P}_{t+1|t}H + \sigma_\nu^2\right| + \left(\frac{\left(di_{t+1} - \mu - H'\hat{\xi}_{t+1|t}\right)^2}{H'\hat{P}_{t+1|t}H + \sigma_\nu^2}\right) \quad (2.8)$$

We maximize the sample log likelihood $\sum_{t=0}^{T-1} \log f_{di_{t+1}|I_t}\left(i_{t+1}|I_t\right)$ with respect to the vector of parameters $\theta = \left(\{\lambda_i\}_{i=1}^p, \eta, \sigma_\epsilon^2, \mu\right)'$.[20] To initiate the estimation procedure, we need an estimate of the space variable $\hat{i}_0^{p0}$ and its conditional mean square error. Maximum likelihood estimation over the vector θ is then performed.

[18]All rates except on the Euro Pound are London interbank offer rates. The Euro Pound rate is a Paris interbank offer rate.

[19]See appendix B and Hamilton (1994, chapter 13).

[20]Eurorates estimation was modified to take into account a maturity larger than the sampling frequency.

Once an estimate $\hat{\theta}^0$ is found, we run the smoother in order to revise the initial state vector. That is, the smoother gives us the initial value of the persistent component, conditional on the entire sample information and the filter parameters, $\hat{i}_0^{p1} = E\left[i_0^p | I_T, \hat{\theta}^0\right]$, and its mean square error. In general, this revised estimate does not correspond to the initial one. We can then iterate the maximum likelihood estimation with this new initial state variable until convergence to $\hat{\theta}^1$. Iterating this procedure will give ultimately a parameter vector consistent with the initial state vector.[21]

We report in this section maximum likelihood estimation for the prime rate for the UK, Canada, Japan versus the US and the Euro 3-months rates for all G-7 countries against the US.

Results for the prime rate, reported in Tables 2.1, 2.2 and 2.3, indicate that persistent shocks disappear extremely slowly. The long run autocorrelation is always above 0.94. For UK-US and Japan-US, the long run autocorrelation is not statistically different from 1, indicating a possible unit root in the persistent component of the interest rate differential. Although the standard errors reported are then incorrect, small sample forecasts based on the point estimates are still correct. These extremely high values for the long run autocorrelation may reflect the relative inertia of the prime rate.[22] The implied annual serial correlation ranges from 0.54 to 0.94. The relative variance parameter is never statistically different from 0.[23]. This confirms the ARMA results and suggests that transitory components are not present. Since these results might be the consequence of the relative inertia of the prime rates (see Figure $2a - g$), we present next the results for the Euro 3-months interest rates estimated over the entire floating period in Tables 2.4, 2.5, 2.6, 2.7, 2.8, and 2.9.[24]

These tables confirm earlier results: there is no significant transitory component in the interest rate differential. The long run persistence is smaller, as we guessed, ranging from 0.86 (France-US) to 0.97 (Italy-US). The short run autocorrelation is frequently higher than one, indicating further deviations from equilibrium 3 months after the initial shock. The relative variance is never significantly different from zero. In summary, interest rate differentials exhibit strong persistence and no significant transitory component. This result appears robust to the order of the autoregressive process.

[21]The asymptotic properties are the same whether we iterate or not.

[22]Our framework assumes a shock every period. If the interest rate differential is changed only infrequently, we are likely to spuriously estimate higher persistence.

[23]The estimation procedure imposes strictly positive standard errors. When the transitory component is strictly 0, the filter may have difficulties working correctly. When this happens, we directly estimate an $AR(q)$ process on the data. A value of 0 is then reported for η.

[24]The euro 3-months reported in the IFS tape (lines 60ldd, and 60ea) is identical to the Currency Forecaster's Digest over the 1986-1995 period. The prime rate, on the other hand, is defined differently in the IFS tape and in the survey dataset.

Table 2.1: UK-US. PRIME RATE DIFFERENTIAL

	AR1		AR2		AR3		AR4	
λ_1	0.9957	(0.1026)	1.3397	(0.0050)	1.5217	(0.1037)	1.0871	(0.1043)
λ_2			-0.3505	(0.0034)	-0.7466	(0.1054)	-0.0885	(0.1045)
λ_3					0.2237	(0.0737)	0.0628	(0.0842)
λ_4							-0.0617	(0.1043)
η	0.0000	(0.0211)	0.0022	(0.1916)	0.0056	(0.1092)	0.0000	(0.3227)
$\lambda(1)$	0.9957	(0.1026)	0.9891	(0.0039)	0.9987	(0.1796)	0.9997	(0.2085)
Log Lik.	-0.0773		-0.0689		-0.0608		-0.0594	

Note: Standard errors in parentheses. Sample period: september 1987 to october 1995. 98 monthly observations.

Table 2.2: CANADA-US. PRIME RATE DIFFERENTIAL

	AR1		AR2		AR3		AR4	
λ_1	0.9635	(0.0451)	1.0248	(0.0379)	1.0439	(0.0343)	1.2702	(0.0268)
λ_2			-0.0788	(0.0347)	-0.2118	(0.0484)	-0.5662	(0.0262)
λ_3					0.1225	(0.0602)	0.3876	(0.0247)
λ_4							-0.1336	(0.0184)
η	0.0000	(0.2382)	0.0000	(0.1077)	0.0000	(0.5501)	0.0034	(0.0011)
$\lambda(1)$	0.9635	(0.0451)	0.9460	(0.0277)	0.9546	(0.0331)	0.9579	(0.0196)
Log Lik.	-0.0759		-0.0758		-0.0748		-0.0740	

Note: Standard errors in parentheses. Sample period: september 1987 to october 1995. 98 monthly observations.

Table 2.3: JAPAN-US. PRIME RATE DIFFERENTIAL

	AR1		AR2		AR3		AR4	
λ_1	0.9956	(0.0126)	1.6867	(0.1031)	1.5517	(0.0030)	1.2926	(6.8990)
λ_2			-0.6940	(0.1031)	-0.4308	(0.0050)	-0.1164	(6.8391)
λ_3					-0.1316	(0.0060)	-0.1030	(6.7745)
λ_4							-0.0818	(6.6674)
η	0.0000	(0.1576)	0.0115	(0.0599)	0.0110	(0.0189)	0.0066	(0.1520)
$\lambda(1)$	0.9956	(0.0126)	0.9927	(0.1458)	0.9893	(0.0034)	0.9914	(0.2699)
Log Lik.	-0.0143		0.0044		0.0056		0.0071	

Note: Standard errors in parentheses. Sample period: september 1987 to october 1995. 98 monthly observations.

Table 2.4: UK-US. Euro 3 months Differential

	AR1		AR2		AR3		AR4	
λ_1	0.9314	(0.0395)	1.1029	(0.1006)	1.1404	(0.0263)	1.1578	(0.0282)
λ_2			-0.1819	(0.1017)	-0.4372	(0.0262)	-0.4683	(0.0261)
λ_3					0.2360	(0.0198)	0.3197	(0.0312)
λ_4							-0.0746	(0.0197)
η	0.0000	(0.4379)	0.0002	(0.1483)	0.0001	(0.1525)	0.0001	(0.1996)
$\lambda(1)$	0.9314	(0.0395)	0.9210	(0.0401)	0.9392	(0.0203)	0.9346	(0.0232)
Log Lik.	-0.3994		-0.3905		-0.3776		-0.3768	

Note: Standard errors in parentheses. Sample period: january 1974 to september 1995. 261 monthly observations.

Table 2.5: France-US. Euro 3 months Differential

	AR1		AR2		AR3		AR4	
λ_1	0.8616	(0.0743)	0.8602	(0.2269)	0.8629	(0.0680)	1.2417	(0.0662)
λ_2			-0.0003	(195.1985)	-0.0332	(0.0922)	-0.4694	(0.0685)
λ_3					0.0349	(0.0532)	0.1011	(0.1692)
λ_4							0.0362	(0.1061)
η	0.0006	(0.3834)	0.0002	(0.1309)	0.0000	(0.3501)	0.4677	(0.1113)
$\lambda(1)$	0.8616	(0.0743)	0.8600	(0.3176)	0.8646	(0.0428)	0.9096	(0.0457)
Log Lik.	-0.6413		-0.6409		-0.6416		-0.6417	

Note: Standard errors in parentheses. Sample period: january 1974 to september 1995. 261 monthly observations.

Table 2.6: Germany-US. Euro 3 months Differential

	AR1		AR2		AR3	
λ_1	0.9654	(0.0271)	1.2350	(0.1186)	1.3135	(0.0693)
λ_2			-0.2790	(0.1307)	-0.6337	(0.0693)
λ_3					0.2876	(0.0693)
η	0.0000	(0.7172)	0.0000	(0.8371)	0.0000	(1.6420)
$\lambda(1)$	0.9654	(0.0271)	0.9561	(0.0609)	0.9674	(0.1201)
Log Lik.	-0.3722		-0.3555		-0.3373	

Note: Standard errors in parentheses. Sample period: january 1974 to september 1995. 261 monthly observations.

Table 2.7: ITALY-US. EURO 3 MONTHS DIFFERENTIAL

	AR1		AR2		AR3		AR4	
λ_1	0.9595	(0.1043)	1.6604	(0.0061)	1.0356	(0.0243)	0.6966	(0.0332)
λ_2			-0.6879	(0.0053)	0.3131	(0.0161)	0.4573	(0.0325)
λ_3					-0.3955	(0.0114)	0.0551	(0.0334)
λ_4							-0.2820	(0.0252)
η	0.2785	(0.3920)	2.6235	(1.3933)	0.7134		0.0000	(0.2251)
$\lambda(1)$	0.9595	(0.1043)	0.9725	(0.0068)	0.9532		0.9270	(0.0170)
Log Lik.	-0.3076		-0.2932		-0.2935		-0.2835	

Note: Standard errors in parentheses. Sample period: january 1974 to september 1995. 261 monthly observations.

Table 2.8: CANADA-US. EURO 3 MONTHS DIFFERENTIAL

	AR1		AR2		AR3		AR4	
λ_1	0.9019	(0.0993)	1.1407	(0.0483)	0.9502	(0.1011)	1.6924	(0.0119)
λ_2			-0.2351	(0.0355)	0.0387	(0.0579)	-0.9306	(0.0120)
λ_3					-0.1051	(0.0388)	-0.0251	(0.0153)
λ_4							0.2400	(0.0095)
η	0.0001	(0.5474)	0.1785	(0.8295)	0.0002	(0.6863)	1.2431	(2.4798)
$\lambda(1)$	0.9019	(0.0993)	0.9056	(0.0378)	0.8837	(0.0577)	0.9767	(0.0360)
Log Lik.	-0.1443		-0.1413		-0.1395		-0.1196	

Note: Standard errors in parentheses. Sample period: january 1974 to september 1995. 261 monthly observations.

Table 2.9: JAPAN-US. EURO 3 MONTHS DIFFERENTIAL

	AR1		AR2		AR3		AR4	
λ_1	0.9347	(0.0167)	1.3062	(0.0191)	1.1159	(0.0707)	1.4934	(0.7733)
λ_2			-0.3711	(0.0121)	-0.1063	(0.0604)	-0.6696	(1.3582)
λ_3					-0.0825	(0.0707)	0.1170	(0.8772)
λ_4							0.0088	(0.2811)
η	0.0000	(0.9909)	0.1438	(0.9975)	0.0003	(5.9217)	0.3024	(0.6574)
$\lambda(1)$	0.9347	(0.0167)	0.9350	(0.0153)	0.9271	(0.1225)	0.9496	(0.0234)
Log Lik.	-0.3845		-0.3723		-0.3725		-0.3722	

Note: Standard errors in parentheses. Sample period: january 1974 to september 1995. 261 monthly observations.

2.4. Estimating the Market filter

We now turn to the estimation of the market filter. We assume that agents use linear forecasting formulas, as summarized by the Kalman filter equations. However, it is unlikely that agents know the exact parameters driving the interest rate differential process. According to one interpretation, these parameters may be time-varying and agents may be in the permanent process of revising their estimates. More generically, we adopt an agnostic view and will estimate the parameters of the filter implicitly used, which we denote the "market filter".[25] Denote $\tilde{\theta} = \left(\left\{\tilde{\lambda}_i\right\}_{i=1}^{p}, \tilde{\eta}, \tilde{\sigma}_\epsilon^2\right)$ the parameters of the market filter. For a given market filter we can generate the associated forecasts at any horizon. Suppose that the current estimate for the state variable at time t is $\hat{\xi}_{t|t}$. According to (2.6)-(2.7), the market forecast for the interest rate differential τ periods from now is:

$$di_t^\tau\left(\tilde{\theta}\right) \equiv E\left[i_{t+\tau}|I_t\right] = \mu + H'\hat{\xi}_{t+\tau|t} = \mu + H'F^\tau\hat{\xi}_{t|t} \qquad (2.9)$$

The forecast constructed in such a way uses only information up to time t. Under the assumption that the reported forecast is measured with error, i.e.:

$$\hat{di}_t^\tau = di_t^\tau + v_t^\tau$$

and that the measurement error is uncorrelated with the true forecast, we can estimate $\tilde{\theta}$ by minimizing:

$$S\left(\tilde{\theta}\right) = \sum_\tau \sum_{t=1}^{T}\left(\hat{di}_t^\tau - di_t^\tau\left(\tilde{\theta}\right)\right)^2$$

where the summation is over both observations and forecasts horizons.[26] We report results for the prime rate first in Tables 2.10, 2.11 and 2.12, then for the euro 3-months rate in Tables 2.13, 2.14, 2.15, 2.16, 2.17 and 2.18.

The results indicate that the market overestimates the speed of convergence of persistent shocks. While the long run autocorrelation is close to 0.9 in sample, the market estimates much higher long run autocorrelation, sometimes significantly in excess of 1.[27] This extra persistence is compensated by a large -and often significant- estimate for the relative variance. Market participants always attributes a sizeable share of the shocks to transitory components, even though we have failed to find such components in sample. Without the excess transitoryness, interest rate shocks would be expected to die out at a much slower rate than observed in the data. However, the presence of the transitory component introduces a dampening effect: a share of the shock disappears extremely rapidly, while the rest slowly decays. We

[25]Note that the market filter refers to the parameters of the filter used implicitly by market participants, not to the actual data-generating parameters in sample.

[26]Eurorates estimation reports robust standard errors since the horizon is larger than the sampling frequency.

[27]While this indicates that the process might not be perceived as second order stationary, finite sample forecasts are still well defined.

Table 2.10: UK-US. PRIME DIFFERENTIAL. MARKET FILTER

	AR1		AR2		AR3	
λ_1	0.9798	(0.0027)	1.7227	(0.0290)	0.9019	(0.1370)
λ_2			-0.7301	(0.0288)	0.6831	(0.2180)
λ_3					-0.5984	(0.0919)
$\hat{\eta}$	0.2222	(0.5703)	9.8255	(2.3310)	2.6759	(1.2620)
$\lambda(1)$	0.9798	(0.0027)	0.9926	(0.0006)	0.9865	(0.0033)

Note: Standard errors in parentheses. Sample period: september 1987 to october 1995. Forecasts at 3, 6 and 12 months. 294 monthly observations.

Table 2.11: CANADA-US. PRIME DIFFERENTIAL. MARKET FILTER

	AR1		AR3	
λ_1	0.9938	(0.0022)	0.7948	(0.0617)
λ_2			0.8876	(0.0646)
λ_3			-0.6873	(0.0284)
$\hat{\eta}$	2.6622	(1.4979)	15.0602	(9.3543)
$\lambda(1)$	0.9938	(0.0022)	0.9951	(0.0019)

Note: Standard errors in parentheses. Sample period: september 1987 to october 1995. Forecasts at 3, 6 and 12 months. 294 monthly observations.

Table 2.12: JAPAN-US. PRIME DIFFERENTIAL. MARKET FILTER

	AR1		AR2		AR3	
λ_1	0.9901	(0.0027)	1.6523	(0.1734)	0.7544	(0.1360)
λ_2			-0.6564	(0.1710)	0.7930	(0.1391)
λ_3					-0.5554	(0.0128)
$\hat{\eta}$	0.6768	(0.1092)	12.2833	(0.7228)	2.5929	(0.1440)
$\lambda(1)$	0.9901	(0.0027)	0.9959	(0.0024)	0.9921	(0.0032)

Note: Standard errors in parentheses. Sample period: september 1987 to october 1995. Forecasts at 3, 6 and 12 months. 294 monthly observations.

report on Figure 4 the actual forecasts and the fitted values for the dollar-pound interest rate differential. This figure indicates that our estimated market filter accurately reproduces the dynamics of the market forecasts.

The results are also strongly supportive of the model. For almost all countries and specification, the noise to signal ratio is large and significant. This implies a substantial share allocated to the transitory component.

The overall results are striking: (a) there is no significant or systematic transitory component of interest rates. This is true for all countries in the sample, and for different time periods, while (b) the parameters of the filter that best replicates interest rate expectations exhibit a substantial transitory components. This is true also across countries, and for different interest rates measures.

3. A Representative Agent Model

In this section we build on our empirical findings and develop a model of exchange rate determination that rationalizes both predictable excess returns and delayed overshooting. We present a model of a world economy with two countries, one consumption good, a domestic bond, and a foreign bond. The economy is populated by atomistic agents who live two periods and derive utility from consumption. Total population is constant across time. Each agent of generation t is born with an endowment W_t of the consumption good. It uses this endowment in consumption, purchasing x_t units of the domestic bond which has a price f_t in terms of the consumption good, and $W_t - c_t - f_t x_t$ units of the foreign bond which has a price of one in terms of the consumption good. The domestic bond is risky. At time $t+1$ it pays i_t units of the consumption good, and is sold at a price f_{t+1} (to the next generation). At time t, i_t is known, while f_{t+1} is not. The foreign bond is riskless. At time $t+1$ its price is 1 and it pays i^* units of the consumption good. When old, at time $t+1$, the representative agent consumes all its wealth. It follows that

$$c_{t+1} = [W_t - c_t - f_t x_t]\,[1 + i^*] + [i_t + f_{t+1}]\, x_t \tag{3.1}$$

In this setup, the exchange rate is the price of the foreign bond in terms of the domestic bond. Since the price of the foreign bond in terms of the consumption good is one, $1/f_t$ is the exchange rate.

A peculiarity of budget constraint (3.1) is that f_t and f_{t+1} enter separately. This specification combined with a constant absolute risk aversion utility function will allow us to derive the equilibrium exchange rate in closed form. This will make it possible to determine whether delayed overshooting occurs.

Following the decomposition adopted in the previous section we assume that the domestic interest rate has three components: its steady state level i^*, a persistent component i_t^p, and a transitory shock ν_t:

$$i_t = i^* + i_t^p + \nu_t \tag{3.2}$$

Table 2.13: UK-US. Euro 3 months Differential. Market Filter

	AR1		AR2		AR3		AR4	
λ_1	0.9774	(0.0024)	1.7154	(0.0369)	2.0440	(0.0586)	1.2901	(0.0878)
λ_2			-0.7229	(0.0361)	-1.2918	(0.1119)	-0.8923	(0.1444)
λ_3					0.2427	(0.0547)	1.2420	(0.1042)
λ_4							-0.6570	(0.0503)
$\hat{\eta}$	0.5970	(0.2077)	16.9190	(1.0496)	37.5876	(14.2598)	2.4625	(0.6937)
$\lambda(1)$	0.9774	(0.0024)	0.9925	(0.0010)	0.9949	(0.0006)	0.9829	(0.0016)

Note: Robust standard errors in parentheses. Sample period: august 1986 to october 1995. Forecasts at 3, 6 and 12 months. 333 monthly observations.

Table 2.14: France-US. Euro 3 months Differential. Market Filter

	AR1		AR2		AR3	
λ_1	1.0108	(0.0025)	0.0958	(0.2238)	-0.8612	(0.0025)
λ_2			0.9248	(0.2274)	0.8454	(0.3426)
λ_3					1.0122	(0.2519)
$\hat{\eta}$	211.9981	(26.9606)	57.6443	(13.8397)	8,3256	(0.9336)
$\lambda(1)$	1.0108	(0.0025)	1.0207	(0.035)	0.9964	(0.044)

Note: Robust standard errors in parentheses. Sample period: august 1986 to october 1995. Forecasts at 3, 6 and 12 months. 333 monthly observations.

Table 2.15: Germany-US. Euro 3 months Differential. Market Filter

	AR1		AR2		AR3		AR4	
λ_1	0.9794	(0.0017)	1.7322	(0.0276)	0.8082	(0.0017)	0.3568	(0.0942)
λ_2			-0.7393	(0.0272)	0.8287	(0.1236)	0.8103	(0.1252)
λ_3					-0.6508	(0.0453)	0.4081	(0.1290)
λ_4							-0.5970	(0.0818)
$\hat{\eta}$	0.0813	(0.1236)	11.0333	(3.8241)	2,2680	(1.6464)	0.6299	(0.2793)
$\lambda(1)$	0.9794	(0.0017)	0.9929	(0.0038)	0.9861	(0.0034)	0.9782	(0.0020)

Note: Robust standard errors in parentheses. Sample period: august 1986 to october 1995. Forecasts at 3, 6 and 12 months. 333 monthly observations.

Table 2.16: ITALY-US. EURO 3 MONTHS DIFFERENTIAL. MARKET FILTER

	AR1		AR2		AR3	
λ_1	1.0221	(0.0039)	0.1443	(1.0651)	1.7184	(0.0471)
λ_2			0.8971	(1.0759)	-1.8203	(0.0578)
λ_3					1.1144	(0.0262)
$\hat{\eta}$	28.2254	(3.3153)	79.0868	(8.3929)	265.1512	(19.1073)
$\lambda(1)$	1.0221	(0.0039)	1.0414	(0.0139)	1.0124	(0.0068)

Note: Robust standard errors in parentheses. Sample period: august 1986 to october 1995. Forecasts at 3, 6 and 12 months. 333 monthly observations.

Table 2.17: CANADA-US. EURO 3 MONTHS DIFFERENTIAL. MARKET FILTER

	AR1		AR2		AR3		AR4	
λ_1	0.9721	(0.0049)	0.2758	(0.1578)	0.9339	(0.0049)	0.1774	(0.1503)
λ_2			0.6756	(0.1531)	0.6285	(0.9276)	1.2014	(0.1755)
λ_3					-0.5774	(0.045)	0.1024	(0.1596)
λ_4							-0.5090	(0.1275)
$\hat{\eta}$	2.9905	(0.9276)	0.7360	(0.3445)	20.6790	(0.5783)	5.2123	(2.8278)
$\lambda(1)$	0.9721	(0.0049)	0.9514	(0.0001)	0.9850	(0.0056)	0.9721	(0.0048)

Note: Robust standard errors in parentheses. Sample period: august 1986 to october 1995. Forecasts at 3, 6 and 12 months. 333 monthly observations.

Table 2.18: JAPAN-US. EURO 3 MONTHS DIFFERENTIAL. MARKET FILTER

	AR1		AR2		AR3		AR4	
λ_1	0.9819	(0.0023)	1.5692	(0.0350)	0.7313	(0.0023)	0.3124	(0.1140)
λ_2			-0.5787	(0.0342)	0.7970	(0.1800)	0.9836	(0.1206)
λ_3					-0.5449	(0.0078)	0.0104	(0.1305)
λ_4							-0.3307	(0.0895)
$\hat{\eta}$	0.2903	(0.1800)	3.9028	(2.4011)	2.9789	(2.0294)	0.3508	(0.1854)
$\lambda(1)$	0.9819	(0.0023)	0.9905	(0.0012)	0.9834	(0.3574)	0.9757	(0.0028)

Note: Robust standard errors in parentheses. Sample period: august 1986 to october 1995. Forecasts at 3, 6 and 12 months. 333 monthly observations.

The persistent component has two elements: the effects on i_t of *past* persistent shocks which die out at rate λ, and the effects of contemporaneous persistent shocks ϵ_t:

$$i_t^p = \lambda i_{t-1}^p + \epsilon_t \tag{3.3}$$

We assume that the shocks ν_t and ϵ_t are independent normal variables with 0 mean and respective variance σ_ν^2 and σ_ϵ^2. Agents in the market only observe the realization of interest rates (i_t and i^*). They do not know whether a change in the domestic interest rate is caused by a transitory (ν_t) or a persistent shock (ϵ_t). All agents receive the same information and draw the same conclusions from it.[28]

We postulate that the objective function of the representative young investor at time t is CARA:

$$U(c_t, c_{t+1}) = -\exp(-\gamma c_t) - \beta E\left[-\exp(-\gamma c_{t+1}) \mid I_t\right] \tag{3.4}$$

where γ represents the coefficient of absolute risk aversion. The information set at time t, I_t, includes all past and current interest rates and prices. That is, $I_t = \{I_{t-1}, i_t, f_t\}$. This information set is common to all agents. This is an appropriate assumption since macroeconomic variables are the relevant ones in the foreign exchange market. It is unlikely that market participants would possess asymmetric information regarding these variables for extended periods of time.

We choose a constant absolute risk aversion specification in order to solve independently the portfolio decision and the savings decision. This will allow us to obtain a linear closed form solution for the equilibrium price of the domestic bond. The specification is standard in the finance literature, and all its inconveniences are well known.

The timing is as follows. At time t, after the realization of the domestic interest rate i_t is revealed, the representative young investor chooses x_t and c_t in order to maximize her expected utility (3.4). In so doing she takes as given the current and expected future price of the domestic bond f_t and $E(f_{t+1}|I_t)$. In forming her expectation about f_{t+1}, each young investor uses her estimate of the probability distribution of i_{t+1}.[29] Lastly, at time $t+1$ the old representative investor sells her holdings of the domestic bond to the "$t+1$ young", consumes all her wealth and transmits her information to her offspring.

3.1. The Learning Problem

Using survey data we found in section 2 that shocks to the interest rate differential are considered by market participants to be more transitory than what they

[28] This way of modelling the learning problem has a long standing in the economics literature. It was used by Mussa (1975) to show that it is rational to form expectations of future inflation using an adaptive mechanism, and by Lucas (1973) in his analysis of the Phillips curve.

[29] As we will show, in this representative agent model, there is nothing to be learned from prices. In more general finance models with information heterogeneity, each investor would extract information from current prices. Prices might then be fully revealing if markets are complete. To prevent this, finance models usually assume that supply varies stochastically (Grossman (1976), Hellwig (1980)). This assumption is not necessary in our set-up as prices are uninformative.

actually are in sample. In all the derivations that follow, we assume that agents perceive $\tilde{\eta} = \frac{\tilde{\sigma}_\nu^2}{\tilde{\sigma}_\epsilon^2}$ while the actual relative variance is η.[30] The empirical findings of section 2 imply $\tilde{\eta} >> \eta$. In the following subsections we will derive the equilibrium exchange rate and compute the impulse responses to interest rate shocks.

We will denote by $\tilde{\alpha}_t$ the representative agent's estimator of i_t^p conditional on the information set $I_t = \{I_{t-1}, i_t\}$, and $\tilde{\eta}$ and by $\tilde{\sigma}_t^2$ the mean square error of this estimate. That is

$$\tilde{\alpha}_t = E\left[i_t^p \,|I_t, \tilde{\eta}\right], \qquad \tilde{\sigma}_t^2 = E\left[(\alpha_t - i_t^p)^2 \,|I_t, \tilde{\eta}\right] \tag{3.5}$$

The corresponding estimates with the correct relative variance are denoted α_t and σ_t^2 respectively. During each period t, young investors learn from old investors that i_{t-1}^p is normally distributed with mean $\tilde{\alpha}_{t-1}$ and variance $\tilde{\sigma}_{t-1}^2$. This induces the prior belief that i_t^p is normally distributed with mean $\lambda\tilde{\alpha}_{t-1}$ and variance $\lambda^2\tilde{\sigma}_{t-1}^2 + \tilde{\sigma}_\epsilon^2$. After observing i_t the young use Bayesian inference to update their belief about the mean of i_t^p. The posterior distribution of i_t^p is given in the following Lemma:

Lemma 3.1. *If at time t the young learn from the old that $i_{t-1}^p \,|I_{t-1} \sim N(\tilde{\alpha}_{t-1}, \tilde{\sigma}_{t-1}^2)$, then after observing i_t they conclude that i_t^p is normally distributed with mean $\tilde{\alpha}_t$ and variance $\tilde{\sigma}_t^2$, where:*

$$\tilde{\alpha}_t = [1 - \tilde{k}_t]\,\lambda\,\tilde{\alpha}_{t-1} + \tilde{k}_t\,[i_t - i^*] \tag{3.6}$$

$$\tilde{\sigma}_t^2 = [1 - \tilde{k}_t]\,[\lambda^2\tilde{\sigma}_{t-1}^2 + \tilde{\sigma}_\epsilon^2] \tag{3.7}$$

$$\tilde{k}_t = \frac{\lambda^2\tilde{\sigma}_{t-1}^2 + \tilde{\sigma}_\epsilon^2}{\lambda^2\tilde{\sigma}_{t-1}^2 + \tilde{\sigma}_\epsilon^2 + \tilde{\sigma}_\nu^2} \leq 1$$

See the appendix for a proof. The t-young will pass on to the $t+1$-young the knowledge that $i_t^p \,|I_t \sim N(\tilde{\alpha}_t, \tilde{\sigma}_t^2)$. After observing i_{t+1} the $t+1$-young will in turn conclude that $i_{t+1}^p \,|I_{t+1} \sim N(\tilde{\alpha}_{t+1}, \tilde{\sigma}_{t+1}^2)$, and pass on this information to the $t+2$-young. Therefore, the system (3.7) can be used as a recursive updating formula under the assumption that at the beginning of history the prior was that i_t^p was normally distributed.

The gain $\tilde{k}_t$ measures how much weight is given to new observations, which depends on the perceived quality of the public signal ($\tilde{\eta}$). No weight is given to past beliefs when there are no transitory disturbances ($\tilde{\sigma}_\nu^2 = 0$), as the interest rate change is perceived to be permanent. As is standard in normal updating, the gain $\tilde{k}_t$ and the variance do not depend on the particular realization of the interest rate:

[30] We also found that the market assumed a higher persistence. It is easy to show that this has no effect on the unconditional delayed overshooting. To save on notational complexity, we assume here that the market accurately perceives the speed of convergence of persistent shocks.

the system (3.7) is independent of i_t and deterministic. In particular, it converges to a steady state $\left(\tilde{k}, \tilde{\sigma}^2\right)$.[31] In what follows, we assume that the process has been going on long enough to have reached its steady state. Solving for the gain and the steady state variance of the belief, we get:

Lemma 3.2. *The steady state gain and variance are given by:*

$$\tilde{k} = \tilde{k}(\lambda, \tilde{\eta}) = \frac{1+\tilde{\Delta}-\tilde{\eta}\,(1-\lambda^2)}{1+\tilde{\Delta}+\tilde{\eta}\,(1+\lambda^2)}, \tag{3.8}$$

$$\tilde{\sigma}^2 = \frac{\left(1-\tilde{k}\right)\tilde{\sigma}_\epsilon^2}{1-\left(1-\tilde{k}\right)\lambda^2}$$

$$\frac{\partial \tilde{k}(\lambda, \tilde{\eta})}{\partial \lambda} \geq 0; \quad \frac{\partial \tilde{k}(\lambda, \tilde{\eta})}{\partial \tilde{\eta}} \leq 0$$

$$\lim_{\tilde{\eta}\to\infty} \tilde{k}(\lambda, \tilde{\eta}) = 0; \quad \tilde{k}(\lambda, 0) = 1;$$

where $\tilde{\Delta}^2 = \left[\tilde{\eta}\,(1-\lambda^2)+1\right]^2 + 4\,\tilde{\eta}\,\lambda^2$;

See the appendix. It follows from (3.8) that the gain depends only on the perceived relative variances of the noise and signal components ($\tilde{\eta}$) and the speed of convergence (λ). The gain is zero and no learning occurs when the noise is infinite while learning is immediate when the signal is perfect. The gain decreases monotonically with the noise to signal ratio. Intuitively, with a higher λ (slower convergence), today's interest rates contain more information about the persistent component of interest rates. As a consequence of the lemma, there is a unique $\tilde{\eta}$ associated with any couple $\left(\lambda, \tilde{k}\right) : \tilde{\eta} = \tilde{\eta}\left(\lambda, \tilde{k}\right)$. We can thus indifferently analyze the properties of the system in terms of $(\lambda, \tilde{\eta})$ or in terms of $\left(\lambda, \tilde{k}\right)$. In the rest of this section, we parameterize our economy in terms of λ and $\tilde{k}$.

In the special case where agents use the correct parameters, $\tilde{\eta} = \eta$, we can derive the same formula, without the "tilde". That is $\tilde{\alpha}_t = \alpha_t$, $\tilde{\sigma}^2 = \sigma^2$, $\tilde{k} = k$ where

$$\begin{aligned} \alpha_t &= [1-k]\,\lambda\,\alpha_{t-1} + k\,[i_t - i^*] \\ k &= \frac{1+\Delta-\eta\,(1-\lambda^2)}{1+\Delta+\eta\,(1+\lambda^2)} \end{aligned} \tag{3.9}$$

3.2. Market Equilibrium

During each period f_t adjusts so that the market for domestic bonds clears. We assume that the population of investors has measure one. Thus, the market clearing condition is

$$\bar{X} = x_t\left(i_t,\ f_t,\ E(f_{t+1}\,|I_t, \tilde{\eta})\right) \tag{3.10}$$

[31] See Hamilton (1994, chapter 13).

where $\bar{X}$ is the aggregate supply of the domestic bond, which is fixed.

In principle, solving this problem could be quite complex since the current price f_t depends on the future price of the bond expected by each investor. To find an equilibrium we use the standard method in Finance[32]. We conjecture a *linear* price rule which does not convey any information beyond what is already contained in interest rates. We then solve the learning and the portfolio problems, and obtain the price function that equilibrates the domestic bond market. Lastly, we validate the initial conjecture. Using this method, we are able to exhibit *one* equilibrium. No claim can be made regarding uniqueness.

We make the following conjecture:

Conjecture 3.1. *The price of the domestic bond is linear in the current interest rate and in the market belief:*

$$f_t = a + b\,\tilde{\alpha}_t + c\,i_t \tag{3.11}$$

The constants a, b and c will be endogenously determined. Since at time t all agents observe i_t, the price does not reveal any new information. Therefore, under the price rule (3.11), the information set is $I_t = \{I_{t-1}, i_t, f_t\} = \{I_{t-1}, i_t\}$.

We will use the conjectured price function to derive the optimal demand for domestic bonds by the representative young investor. We will then use the market clearing condition (3.10) to determine the market clearing price f_t. As a third step, we will confirm that the market clearing f_t has the form conjectured in (3.11), and we will determine a, b and c.

Lemma 3.3. *If the price function has the form conjectured in (3.11), the demand for the domestic bond is given by:*

$$x_t(f_t, i_t, \tilde{\alpha}_t) = \frac{-f_t[1+i^*] + i_t + [b+c]\lambda\tilde{\alpha}_t + ci^* + a}{\gamma\pi^2} \tag{3.12}$$

where $\pi^2 = \left(b\tilde{k} + c\right)^2 (\lambda^2\bar{\sigma}^2 + \tilde{\sigma}_\epsilon^2 + \tilde{\sigma}_\nu^2)$

The proof is in the appendix. The term π^2 is the conditional variance of next period's price. One can rewrite the demand function in a more familiar form:

$$x_t(f_t, i_t, \alpha_t) = \frac{i_t + E\left[f_{t+1} | I_t, \tilde{\eta}\right] - f_t[1+i^*]}{\gamma\,\pi^2}$$

The demand is increasing in the expected future price and interest rate, and decreasing in the current price and variance. Substituting (3.12) in the market clearing condition (3.10) it follows that:

$$f_t = \left[-\gamma\,\pi^2\bar{X} + ci^* + a + (b+c)\lambda\tilde{\alpha}_t + i_t\right][1+i^*]^{-1} \tag{3.13}$$

Lastly, equalizing the coefficients of the conjectured price function (3.11) with those of (3.13) we obtain the equilibrium price function:

[32] see Wang (1994) for an application to trading volume.

Proposition 3.1 (Equilibrium Exchange Rate). *The domestic bond's equilibrium price function is linear in the state variables and is given by:*

$$
\begin{aligned}
f(i_t, \tilde{\alpha}_t) &= \frac{1+i_t}{1+i^*} + \frac{\lambda \tilde{\alpha}_t}{[1+i^*][1+i^*-\lambda]} - \rho \qquad (3.14) \\
\rho &= \frac{\gamma \pi^2 \bar{X}}{i^*}
\end{aligned}
$$

This validates Conjecture 3.1 and characterizes the equilibrium. The equilibrium exchange rate $(1/f_t)$ appreciates if the current interest rate i_t rises, or if the market belief about the persistent component of the interest rate, $\tilde{\alpha}_t$, increases. It depreciates if the *risk premium* ρ goes up; ρ increases with the total conditional volatility of next period price π^2, the supply of the risky asset $\bar{X}$, and the coefficient of risk aversion γ. Note that when the two shocks are purely transitory ($\lambda = 0$), the market belief does not enter the pricing equation, since it does not provide any information regarding future realizations of the interest rate. It should be clear that in the case where agents use η the formula for the exchange rate is obtained by simply replacing $(\tilde{\alpha}_t, \tilde{k})$ by (α_t, k).

3.3. Conditional Delayed Overshooting

Eichenbaum and Evans (1995) find that expansionary shocks to US monetary policy are followed by sharp and persistent reductions in US interest rates. Furthermore, they find that the exchange rate systematically follows a "delayed overshooting" path: after an initial depreciation, the US dollar continues to depreciate for several months before it starts appreciating. A similar humped pattern is found by Clarida and Gali (1994) and Grilli and Roubini (1994). We replicate these findings in Figure 1a-d.

The delayed overshooting path is characterized by Eichenbaum and Evans (1995) as the impulse response of the exchange rate to an unanticipated monetary shock. In the context of our model, we compute the path that the exchange rate would follow if a shock to the domestic interest rate were to take place at time 0, followed by no other shocks. In order to derive the intuition, in this subsection we consider a path *conditional on a persistent shock*. In the next subsection we consider a path which is not conditioned on whether the shock is persistent or transitory. That is, in this subsection we assume that $\epsilon_0 = \kappa$, $\nu_0 = 0$ and $\epsilon_t = \nu_t = 0$ for $t > 0$. It follows from (3.2) and (3.3) that the domestic interest rate follows the *conditional* path:

$$
i_t = i^* + \lambda^t \kappa \qquad \text{for all } t \geq 0 \qquad (3.15)
$$

We also assume for notational simplicity that in the period before the shock took place the interest rate and the expected value of the persistent component of interest rates were: $i_{-1} = i^*$ and $\alpha_{-1}=0$ respectively. Thus, the exchange rate $1/f(i_{-1}, \alpha_{-1})$ was equal to $1/\bar{f} = 1/[1-\rho]$.

3.3.1. Full-Information Dynamics.

It is illustrative to consider first the *Full Information* exchange rate path, which corresponds to the Dornbusch (1976) experiment. Full information in our set-up refers to a situation in which agents observe the current realization of the persistent component of interest rates i_t^p but ignore future shocks. It is straightforward to check, replacing α_t by i_t^p and adopting the same methodology as before, that the full information price is given by:

$$f^{FI}(i_t, i_t^p) = \frac{1+i_t}{1+i^*} + \frac{\lambda i_t^p}{[1+i^*][1+i^*-\lambda]} - \rho^{FI} \tag{3.16}$$

where the superscript FI stands for full information and $\rho^{FI} = \gamma \bar{X}\left\{\sigma_\epsilon^2\ (b+c)^2 + \sigma_\nu^2 c^2\right\} i^{*-1}$. It follows from (3.16) and (3.15) that the price dynamics generated by a persistent shock under full information are given by:

$$f_t^{FI,p} - \bar{f}^{FI} = \frac{\kappa\lambda^t}{1+i^*-\lambda} \tag{3.17}$$

which exhibits the standard overshooting result as expected. The exchange rate appreciates initially, and from then on depreciates gradually (see ??).

3.3.2. Imperfect Information Dynamics

We assume in this section that agents form correct assessments about the relative share of transitory and persistent components: $\tilde{\eta} = \eta$ and $\tilde{k} = k$. In the case of imperfect information, by substituting (3.15) in (3.14) it follows that the initial jump in f is:

$$f_0^p - \overline{f} = \delta\kappa > 0, \tag{3.18}$$

$$\text{where } \delta = \frac{1+i^*-\lambda(1-k)}{(1+i^*)(1+i^*-\lambda)} > 0$$

Thus, the exchange rate also appreciates upon impact. However, comparing (3.18) and (3.17) we can see that under imperfect information the extent of initial exchange rate appreciation is lower than under full information. Algebraically:

$$\left[f_0^p - \overline{f}\right] / \left[f_0^{FI,p} - \bar{f}^{FI}\right] = 1 - \frac{\lambda(1-k)}{1+i^*} < 1.$$

Investors are less willing to buy up the domestic asset, as the shock might prove transitory. Subsequently the change in the price of the domestic bond under imperfect information is given by the following lemma, which is proven in the appendix:

Lemma 3.4. *The price change following a once and for all persistent shock is:*

$$f_{t+1}^p - f_t^p = \frac{\kappa\lambda^{t+1}}{(1+i^*)(1+i^*-\lambda)} \left[(1-k)^{t+1}(1-\lambda(1-k)) - (1+i^*)\left(\tfrac{1}{\lambda}-1\right)\right] \tag{3.19}$$

Unlike in the full information case, delayed overshooting may occur under imperfect information. That is, the price of the domestic bond may increase until some time $\tau_f > 0$,and decline thereafter. We can determine under which circumstances this occurs by looking at the sign of (3.19). Since λ and k take values on $[0,1]$ and the bracketed term in (3.19) is decreasing in t, we have the following result, which is proven in the appendix:

Proposition 3.2 (Conditional D-O Region). *A necessary and sufficient condition for delayed overshooting after τ periods, conditional on a permanent shock at time* 0, *is:*

$$(1-k)^{\tau+1}(1-\lambda(1-k)) - \left(\tfrac{1}{\lambda}-1\right)(1+i^*) > 0 \tag{3.20}$$

This defines a delayed overshooting region

$$D_\tau = \{(k,\lambda)\,|\textit{such that } (3.20) \textit{ is satisfied}\}$$

The boundary of D_τ is given by:

$$\lambda(k,\tau) = \frac{(1+i^*) + (1-k)^{\tau+1} - \sqrt{\phi}}{2(1-k)^{\tau+2}} \tag{3.21}$$

$$\textit{where } \phi = \left[(1+i^*) + (1-k)^{\tau+1}\right]^2 - 4(1+i^*)(1-k)^{\tau+2}$$

Note that the stylized fact that agents attach more importance to transitory shocks ($\tilde{\eta} > \eta$) is not necessary nor sufficient for conditional delayed overshooting.

Figure 5 shows the path of the *inverse* of the exchange rate in response to a unit standard error increase in the interest rate at $t = 0$ in both the perfect and imperfect information cases.[33] It shows that a positive and persistent interest rate shock induces an initial appreciation of the exchange rate, followed by an appreciating path which lasts about twenty periods before reverting to a depreciating path. If one interprets each period as a week or a month, this graph resembles the impulse response functions estimated by Clarida and Gali (1994), Eichenbaum and Evans (1995) and Grilli and Roubini (1994). The duration of each period should depend on the frequency with which one believes that investors receive "new and relevant" information.

We now describe the intuition behind this result. There are two effects:

- **Interest rate effect.** This is the standard mechanism analyzed by Dornbusch (1976). After an initial upward jump, domestic interest rates follow a declining path. This induces the exchange rate to experience an immediate appreciation followed by a gradual depreciation to ensure that uncovered interest parity holds.

[33]The parameters chosen are: $\lambda = 0.98$, $k = 0.2$, $i^* = 0.05$.

- **Learning effect.** When the shock takes place at time 0 agents only observe an increase in i_0. Since the actual sequence of interest rates that agents observe (i.e. those generated by (3.15)) is greater than i^*, market participants gradually increase their belief about i_t^p (i.e. α_t) using updating equation (3.7). As α_t is revised upwards the demand schedule for domestic bonds shifts upwards over time, generating appreciating pressures on the exchange rate.

The learning effect counteracts the interest rate effect. Therefore, upon impact the exchange rate does not jump as much as it does under perfect information. Afterwards, if the learning effect dominates, the exchange rate will continue appreciating after its initial jump as shown in Figure 5. Since agents know that i_t^p declines at rate λ, at some point in time (call it τ_α) they must start revising α downwards as shown in Figure 6. Thus, the exchange rate cannot continue appreciating forever. It is clear that after τ_α the exchange rate must depreciate because the learning and interest rate effects point in the same direction. Moreover, it can be shown that the exchange rate starts depreciating *before* time τ_α. To see this we compute from (3.7) the revisions in beliefs:

$$\alpha_{t+1} - \alpha_t = \epsilon_0 \lambda^t \left[(1-k)^{t+1} (1 - \lambda(1-k)) - (1-\lambda) \right] \tag{3.22}$$

from which it follows that the switching time for market beliefs is

$$\tau_\alpha = \frac{\ln\left((1-\lambda) / (1 - \lambda(1-k))\right)}{\ln(1-k)}$$

It follows from (3.19) that the switching time for the exchange rate is

$$\tau_f = \frac{\ln\left((1-\lambda)(1+i^*) / (1-\lambda(1-k))\lambda\right)}{\ln(1-k)}$$

It is straightforward to check that $\tau_\alpha > \tau_f$, so that *the exchange rate always peaks before beliefs have started reverting.* This reflects the forward-looking behavior of the exchange rate.

Next we analyze the joint restrictions imposed by Proposition 3.2 on λ and k in order to deliver delayed overshooting. A smaller λ (less persistence) increases the second term in (3.20) proportionally more than the first one, making delayed overshooting less likely. This means that an economy converging more rapidly to its long run equilibrium is less likely to exhibit delayed overshooting. As convergence occurs faster, persistent shocks look like transitory ones. Thus, little weight is given to past observations, weakening the learning effect.

Changes in k (the learning rate) have more complex effects: the first term in (3.20) is concave in k. For a sufficiently large k, the learning process works efficiently and at the time of the shock beliefs almost converge to the true value of the persistent component of the interest rate. As a consequence the subsequent upward revision of beliefs is very small. Therefore, the learning effect is dominated by the interest rate effect and there is no delayed overshooting. In other words, since beliefs have almost converged at time 0, market participants bid the exchange rate down until it is back on the full information rational expectations path. For sufficiently small k, learning occurs very slowly and interest rates convey little information about their persistent component. Thus, the market belief α increases very little at the time of the shock. Afterwards, although α is updated upwards, the learning effect is too small to dominate the interest rate effect.

To summarize, *the simple representative agent model we consider delivers dynamics that rationalize the Eichenbaum and Evans puzzle in the case of a persistent shock. An economy will exhibit conditional delayed overshooting if the interest rate converges slowly (large λ) to its new long run value following a permanent shock to the money supply, and if learning occurs moderately slowly (intermediate values of k).*

We now turn to the issue of the length of time over which the exchange rate moves in the "wrong" direction. As we increase the peak date τ, the conditions on λ and k become more stringent: the frontier of D_τ shifts up, as can be seen by computing $\frac{\partial\lambda(k,\tau)}{\partial\tau}$. As shown in Figure $7a-c$, the delayed overshooting region shrinks as we increase the peak date.

Thus, our analysis has strong cross-sectional implications. Countries should exhibit delayed overshooting if (a) monetary shocks are persistent, resulting, for instance, from a low interest elasticity of money demand, and (b) the learning speed is sufficiently small, but not too small. This in turns implies a variance ratio (transitory over persistent) that is greater than one and bounded. Returning to the estimates from section 2, the first condition is likely to be satisfied, as the long run autocorrelation is usually extremely high, indicating a high persistence, while the second is unlikely to be met: without transitory component ($\eta = 0$), the inference problem is trivial and there should be no delayed overshooting.

3.4. Unconditional Delayed Overshooting

We turn in this subsection to the unconditional response of exchange rates to interest rate shocks. Eichenbaum and Evans (1995), although controlling for the source of the disturbance, do not control for the persistence of shocks. Therefore, the empirical puzzle will be solved only if one is able to generate unconditional delayed overshooting. The unconditional exchange rate path is the path following a change in the observed interest rate $i_0 = \kappa$, not just in the persistent component ϵ_0. We show in the appendix that given the linearity of the exchange rate equation (3.14), the unconditional path is a weighted average of the impulse responses to a persistent and a transitory shock. The weights are given by the respective conditional probabilities of both shocks. Define $f_t^l(\kappa)$ as the exchange rate response at time t to an interest shock κ at time 0, where $l = \{u, p, \tau\}$. (u stands for unconditional, p for permanent and τ for transitory):

$$f_t^u(\kappa) = E\left[f_t \mid i_0 - i^* = \kappa\right] = qf_t^p(\kappa) + [1-q]f_t^\tau(\kappa), \tag{3.23}$$

$$q = E\left[\epsilon_0 \left| i_0 - i^* = 1\right.\right] = \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + \sigma_\nu^2 + \lambda^2\sigma^2}$$

where $q\kappa$ is the expected value of the persistent shock conditional on the realization of i_0. We show in the appendix that the unconditional path satisfies:

$$f_{t+1}^u(\kappa) - f_t^u(\kappa) = \frac{\kappa\lambda^{t+1}}{(1+i^*)(1+i^*-\lambda)} \left[(q-k)(1-k)^t(1-\lambda(1-k)) - q(1+i^*)\left(\frac{1}{\lambda}-1\right)\right] \tag{3.24}$$

We can sign this expression by looking at the brackets. Since $\lambda < 1$, the second term inside the brackets is negative. The first term in brackets can be positive or negative according to the sign of $q - k$. Thus, we have the following result:

Lemma 3.5 (No Unconditional D-O). *When $\tilde{\eta} = \eta$, unconditional delayed overshooting never occurs.*

The proof simply shows that $q < k$. See the appendix. The intuition for this result is straightforward: agents only observe the sum of two shocks. Conditional on this information, the persistent component of interest rates is normally distributed, with mean $q\kappa$. Hence the observed interest rate shock is equally consistent with large and positive persistent shocks $q\kappa + \epsilon$ and small or even negative persistent shocks $q\kappa - \epsilon$. In the former case, agents underreact, as in the previous subsection, while in the latter case agents initially *over-react* to the change in fundamentals. On average, a rational agent will not make mistakes and the standard overshooting result will apply.

In the more general case where $\tilde{\eta} \neq \eta$,there is some scope for unconditional delayed overshooting. All the calculations are the same as in the previous case replacing α_t by $\tilde{\alpha}_t$ and k by $\tilde{k}$. Thus the unconditional path satisfies (3.25)

$$f_{t+1}^u(\kappa) - f_t^u(\kappa) = \frac{\kappa\lambda^{t+1}}{(1+i^*)(1+i^*-\lambda)} \left[\left(q-\tilde{k}\right)\left(1-\tilde{k}\right)^t\left(1-\lambda\left(1-\tilde{k}\right)\right) - q(1+i^*)\left(\frac{1}{\lambda}-1\right)\right] \tag{3.25}$$

As in the previous case the value of q is computed using the *actual* distribution: $q = E\left[\epsilon_0 \left| i_0 - i^* = 1, \eta\right.\right] = \frac{1-(1-k)\lambda^2}{1+\eta(1-(1-k)\lambda^2)}$. Therefore, a necessary condition for unconditional delayed overshooting is $q > \tilde{k}$. Since k (resp. $\tilde{k}$) is decreasing in η (resp. $\tilde{\eta}$), we have the following result.

There are two observationally equivalent interpretations of $\tilde{\eta} > \eta$. One is that market participants make systematic errors regarding the variances of the shocks. The other is that agents know the correct distribution, yet that the sample period under study exhibits an unusually large number of persistent shocks. The latter interpretation indicates a small sample problem over the period studied (86-95) or, equivalently, a the existence of a Peso problem. The former interpretation is consistent with a learning framework where agents learn both about the *size* of persistent and transitory components and their *variance*. However, if the parameters driving the interest rates differential process are fixed and yet unknown, the learning process should converge towards the true estimates, and expectational errors should disappear, as in Lewis (1989a). One needs a framework with *stochastic regimes shifts*. We plan to investigate further such a framework in the future.

Proposition 3.3 (Unconditional D-O Region). *A necessary condition for unconditional delayed overshooting is that shocks are perceived as more transitory i.e.,* $\tilde{\eta} > \eta$.

For a given η *a necessary and sufficient condition for unconditional delayed overshooting after* τ *periods, is:*

$$\left(q-\tilde{k}\right)\left(1-\tilde{k}\right)^{\tau}\left(1-\lambda\left(1-\tilde{k}\right)\right)-\left(\tfrac{1}{\lambda}-1\right)\left(1+i^{*}\right)>0 \qquad (3.26)$$

If it exists, the unconditional delayed overshooting region, given η, *is defined by*

$$D^{u}_{\tau,\eta}=\left\{\left(\tilde{k},\lambda\right)|\text{such that (3.26) is satisfied}\right\}$$

The boundary of $D^{u}_{\tau,\eta}$ *is given by:*

$$\lambda\left(\tilde{k},\tau\right) = \frac{q\left(1+i^{*}\right)+\left(q-\tilde{k}\right)\left(1-\tilde{k}\right)^{\tau}-\sqrt{\phi}}{2\left(1-\tilde{k}\right)^{\tau+1}\left(q-\tilde{k}\right)} \qquad (3.27)$$

$$\text{where } \phi = \left[q\left(1+i^{*}\right)+\left(q-\tilde{k}\right)\left(1-\tilde{k}\right)^{\tau}\right]^{2}-4\left(1+i^{*}\right)\left(1-\tilde{k}\right)^{\tau+1}\left(q-\tilde{k}\right)$$

Figure $8a-c$ reports the boundaries of the unconditional delayed overshooting region when $\eta=0.5$. It appears that the delayed overshooting region is truncated on the right: large values of the gain are not achievable anymore. For a given λ this is equivalent to having a high $\tilde{\eta}$. The larger the true noise to signal ratio, the tighter the constraints on the delayed overshooting region. In the limit, when $\tilde{\eta}$ approaches the true η, the delayed overshooting region vanishes. Conversely, it is immediate to check that the limit of the unconditional delayed overshooting region when $\eta \rightarrow 0$ is simply the conditional delayed overshooting region, since $q \rightarrow 1$. This confirms that the most favorable case for delayed overshooting occurs when the actual shocks are largely persistent (η close to 0) while they are perceived as transitory ($\tilde{\eta}$ large).

3.5. Predictable Excess Returns

In our model the uncovered interest parity condition is $E_t\left[f_{t+1}|I_t\right]+i_t=\left(1+i^{*}\right)f_t$. To see why, note that one unit of the domestic bond has an expected return of

$E_t\left[f_{t+1}\,|I_t\right]+i_t$ in terms of the consumption good. Alternatively the investor can buy f_t units of the foreign bond and obtain a safe return of $1+i^*$. We can therefore define predictable excess returns as:

$$\zeta_t = E_t\left[f_{t+1}\,|I_t\right] + i_t - (1+i^*)\, f_t \tag{3.28}$$

In the case where $\eta = \tilde{\eta}$, substituting the equilibrium price equation (3.14) into (3.28) and taking expectations, we get:

$$\zeta_t = i^*\rho \tag{3.29}$$

Thus, if $\eta = \tilde{\eta}$, predictable excess returns are constant and equal to the risk premium. That is, the forward discount puzzle remains unexplained. This result implies that delayed overshooting *conditional* on a persistent shock does not imply time-varying predictable expected excess returns.

We now turn to the case where $\eta < \tilde{\eta}$. Substituting $\tilde{\alpha}_t$ for α_t in the equilibrium price equation (3.14) and taking the expectation with respect to the actual distribution of shocks in the sample, it follows that (3.28) becomes

$$\tilde{\zeta}_t = i^*\rho + \frac{\lambda\left[1+i^*-\lambda\left(1-\tilde{k}\right)\right]}{(1+i^*)\,(1+i^*-\lambda)}\,(\alpha_t - \tilde{\alpha}_t) \tag{3.30}$$

Predictable excess returns are time varying and are correlated, through the market belief. Under the conditions that ensure unconditional delayed overshooting, agents in the market will have systematically downward biased estimate of the persistent component. This in turn implies that the excess returns are positive and correlated through time. In other words, the predictable excess returns are positively correlated with the bond price forecast error.

These results show that a simple model of exchange rate pricing, can generate delayed overshooting and rationalize the forward discount puzzle if (a) agents are learning about the duration of interest rates shocks, and (b) the ex post distribution of shocks exhibits more persistent shocks than expected by the market. The estimates of the parameters driving the interest rates differential process, and their market perception, as reported in Tables 2.10-2.18 suggest that for most countries against the US, the conditions for delayed overshooting, according to Proposition 3.3, are satisfied.

4. An Extended Model

In the learning model of section 3, the gain and variance of the filter evolve deterministically from their initial values to their steady state values (see (3.7)). As a result, current observations of the interest rate only lead to a revision of the belief regarding the persistent component $\alpha_t = E\left[i_t^p\,|\,i_s;\ s\leq t\right]$. The linearity of the updating formula and the fact that only the first moment depends on realizations of the interest rate are responsible for the absence of predictable excess returns.

In this section we will extend the model by considering a more realistic interest rate process in which the variance of the transitory shock is time-varying. As in section 2 the interest rate is given by (3.2) and (3.3): $i_t = i^* + i_t^p + \nu_t$ and $i_t^p = \lambda\, i_{t-1}^p + \epsilon_t$, where the shocks ϵ_t and ν_t are independent. As before, the persistent shock ϵ_t is normal with mean zero and constant variance σ_ϵ^2. However, for the transitory shock we assume:

$$\begin{aligned} v_t\,|I_{t-1} &\sim N(0, \sigma_v^2(t)) \\ E\left[\nu_t^2\,|\nu_{t-1}\right] &= h_t^2 = \psi_0 + \psi_1 \nu_{t-1}^2; \\ \psi_0 &> 0,\ 0 \le \psi_1 < 1 \end{aligned} \tag{4.1}$$

The autoregressive conditional heteroskedasticity (ARCH(1)) specification in (4.1) captures the tendency in financial data for volatility clustering, i.e., the tendency for large (small) price changes to be followed by other large (small) price changes of unpredictable sign. Thus, following a large transitory shock, there is an increase in variance which leads to a reduction in the *precision* of the investors' belief.[34] In order to verify the validity of our decomposition of the interest rate into a persistent and a transitory component, in subsection 3.4 we estimate the process for interest rate differentials of G7 countries. We find that the interest rate differentials have a strong persistent component and a strong ARCH effect in the transitory component.

The ARCH model and its various extensions have been successfully applied to several financial time series (see Bollerslev, Chou and Kroner (1992) for a survey), including interest rates. In their study of varying risk premia in the term structure Engle, Lilien and Robins (1987) find very strong ARCH effects on the excess holding yield of six-month over three-month T-bills, using quarterly data from 1960-I to 1984-II. They also find ARCH effects using monthly data on one-month and two-month T-bills from 1953.1 to 1971.7. Grier and Perry (1993) look at quarterly interest rate surprises, measured as the difference between the one-month T-bill rate and the three-month forward rate for that period. They also find ARCH effects for the sample 1960-III through 1991-IV. This empirical evidence suggests that by adding an ARCH component to the interest rate process we introduce an important element of realism into the model.[35] In addition, this section also

[34]The results would be equivalent if we assumed instead a time dependence on past variances, i.e. a GARCH process instead of an ARCH. In that case the investor would need to estimate today's variance of the transitory component $E\left[v_{t-1}^2\,|I_{t-1}\right]$, which is equivalent to the expectation of today's conditional variance $E\left[h_{t-1}^2\,|I_{t-1}\right]$. By the law of iterated expectations we have that $E\left[h_{t-1}^2\,|I_{t-1}\right] = E\left[E\left[v_{t-1}^2\,|v_{t-2}, I_{t-1}\right]|I_{t-1}\right] = E\left[v_{t-1}^2\,|I_{t-1}\right]$.

[35]Using weekly and monthly data Brenner, Harjes and Kroner (1994) estimate a model that nests the GARCH model and a model in which volatility is a function of the interest rate level. They find that their model outperforms the others because it does not underpredict volatility when interest rate levels are high as the GARCH model does, nor does it overpredict volatility in stable times as the levels model does.

demonstrate that predictable excess returns can be present without unconditional delayed overshooting. The key insight is that predictable excess returns are generated by a varying risk premium while they were generated in the previous section through expectational errors.

4.1. The Learning Problem

In the extended model, young investors solve basically the same learning problem as in section 3. We assume in all this section that $\tilde{\eta} = \eta$. The only difference introduced by (4.1) is a change in the conditional distribution of $i_t - i^*$. We prove in the appendix the following result:

Lemma 4.1. *If at time t the young learn from the old that $i^p_{t-1} \,|I_{t-1} \sim N(\alpha_{t-1}, \sigma^2_{t-1})$ and if (4.1) holds, then after observing i_t the young conclude that $i^p_t \,|I_t \sim N(\alpha_t, \sigma^2_t)$, where:*

$$
\begin{aligned}
\alpha_t &= [1 - k_t]\, \lambda\, \alpha_{t-1} + k_t\, [i_t - i^*] \\
\sigma^2_t &= [1 - k_t]\, [\lambda^2 \sigma^2_{t-1} + \sigma^2_\epsilon] \\
k_t &= \frac{\lambda^2 \sigma^2_{t-1} + \sigma^2_\epsilon}{(\lambda^2 + \xi_1)\, \sigma^2_{t-1} + \sigma^2_\epsilon + \psi_0 + \psi_1\, (i_{t-1} - i^* - \alpha_{t-1})^2}
\end{aligned}
\tag{4.2}
$$

Since the t-young will in turn pass on to the $t+1$-young the knowledge that $i^p_t \,|I_t \sim N(\alpha_t, \sigma^2_t)$ and so on, the system (4.2) can be used as a recursive updating formula under the assumption that the prior was that i^p_t was normally distributed at the beginning of history.

The updating formulas for α_t and σ^2_t are the same as (3.7) in section 3. The only difference is that the gain k_t depends on last period's interest rate. This will be the key to generating predictable excess returns. The intuition is as follows. The last term in the denominator of the gain is the square of the estimate of the transitory shock $\hat{\nu}_{t-1}$. An increase in i_{t-1} leads to a higher estimate of ν_{t-1}, hence a higher estimate of the variance of today's transitory shock, $\hat{h}^2_t$. This decreases the gain of the filter and increases the variance in the prior distribution of i_t. As a result less weight is given to current observations in updating beliefs. In the absence of shocks, the filter still converges to a steady state. However, one cannot use that steady state when inferring the size of the persistent component.

4.2. The Equilibrium Price Function

As before, we conjecture a linear price rule that does not reveal any information. However, since the filter is time-varying, we allow a_t to be time dependent:

$$f_t = a_t + b\,\alpha_t + c\,i_t \tag{4.3}$$

In equilibrium the time-varying coefficient will be the risk premium.

To solve the portfolio problem, we assume that the risk premium is unpriced risk. Thus, when solving the portfolio problem we replace a_{t+1} by its expectation $\hat{a}_{t+1} = E[a_{t+1} \mid I_t]$ in the objective function (3.4). Note that since the interest rate shock affects the gain with a one-period lag, next period's gain and filter are known as of time t. Thus, agents know at time t the distribution of (i_{t+1}, α_{t+1}). Therefore, we can make the same factorization as in the proof of Lemma 3.3, and obtain a closed form solution for f_t. Following the same steps as in the previous section we find that:

Proposition 4.1 (Equilibrium Exchange Rate). *The domestic bond's equilibrium price function is linear in the state variables and is given by:*

$$\begin{aligned} f_t &= \frac{i_t}{1+i^*} + \frac{\lambda\,\alpha_t}{(1+i^*)\,(1+i^*-\lambda)} + a_t \qquad (4.4) \\ a_t &= \frac{1}{1+i^*} - \gamma\,\bar{X}\,E_t\left[\sum\nolimits_{s=0}^{\infty} \frac{\pi_{t+s}^2}{[1+i^*]^{1+s}}\right] \\ \text{where}\quad \pi_t^2 &= \left(\frac{\lambda k_{t+1}}{(1+i^*-\lambda)} + 1\right)^2 \frac{\sigma_\epsilon^2 + (\lambda^2+\psi_1)\sigma_t^2 + \psi_1(i_t - i^* - \alpha_t)^2 + \psi_0}{(1+i^*)^2} \end{aligned}$$

The coefficients on the interest rate and market belief are the same as in the previous section. The difference is that the independent term and the learning parameter k_t are now time-varying.

Following an interest rate shock the conditional variance of next period's interest rates (the numerator of π_t^2) goes up, leading to an increase in the risk premium. Simultaneously, the learning parameter k_{t+1} goes down, pushing the risk premium down. However, it is straightforward to show that the risk premium unambiguously goes up. Over time, as agents update their beliefs, the risk premium returns to its equilibrium value. We call this mechanism the indirect learning effect. This effect captures an important feature of financial markets: a large shock leads investors to believe that there is more uncertainty than previously expected, leading them to be more cautious. Over time, as interest rates return to normal levels, investors regain confidence.

4.3. Predictable excess returns

The indirect learning effect implies that the risk premium will vary over time. Accordingly, there is the potential for time-varying predictable excess returns. As noted previously, predictable excess returns can be defined as in equation (3.28): $\zeta_t = E\left[f_{t+1}\,|I_t\right] + i_t - (1+i^*)\,f_t$. Replacing f_{t+1} and f_t by their respective formulas, we get:

$$\zeta_t = \gamma\,\bar{X}\left[\pi_t^2 + \sum\nolimits_{s=1}^{\infty} E\left[\pi_{t+s}^2\,|I_t\right](1+i^*)^{-s-1}\right] \qquad (4.5)$$

Thus, time varying predictable excess returns reflect time varying risk premia. We observe time varying risk premia because agents' perceived risk changes over time depending on the sequence of shocks observed.

In order to see whether this time varying risk premium can explain the forward discount bias, we conducted the following Monte Carlo experiment. Using the estimated parameters for the Yen-Dollar exchange rate (see subsections 4.4), we simulated 10000 series of 1000 exchange rate and interest rate observations. For each replication, we performed the following standard "Fama" regression:

$$f_{t+1} - f_t = b_0 + b_1 \left(f_t^w - f_t\right) + u_t$$

where $f_t^w = f_t\,(1 + i^*) - i_t$ is the forward rate in the context of our model. Under Uncovered Interest Parity, $b_1 = 1$.[36] Figure 9 reports the distribution of b_1. It appears that the distribution is centered on $\hat{b}_1 = 0.2997$ and a 95% confidence interval is $[-0.33, 0.93]$. Thus, we can reject the null. We conclude from this exercise that learning about the persistent component of the interest rate differential can explain some of the forward discount bias. It is clear, however, that it cannot explain all of the bias. For the Yen-Dollar exchange rate, Lewis (1994) finds that $b_1 = -2.28$ and is statistically significant. This finding is confirmed by Figure 10, which reports b_1 against its standard error. As can be seen on the graph, most points cluster above the non-significance cone. On average, the simulated results report a positive, significant, but less than 1 coefficient.

Although the fluctuations in the risk premium associated with the precision of the learning process can account for part of the forward discount puzzle, it is clear that there is no unconditional delayed overshooting. The impulse response is now different following a positive and a negative shock. After a positive interest rate shock, the currency appreciates. This appreciation is dampened by the increase in risk premium. Thus a delayed overshooting response may arise. Following a decrease in the interest rates, however, the exchange rate depreciates. this initial depreciation is compounded by the increase in risk premium and leads to an over depreciation. On average, the impulse response does not exhibit any delayed overshooting.

4.4. Maximum Likelihood Estimation of the Interest Rate Process

In this subsection we confirm the validity of our assumptions regarding the decomposition of the interest rate into a persistent and a transitory component using data for the G7 countries. Figure 11 reports the interest rate differential against the US from 1974 to 1992. If the interest rate satisfies (3.2), (3.3) and (4.1), we can write the conditional likelihood of i_t as:

$$f_{\tilde{i}_t|I_{t-1}}\left(i_t|I_{t-1}\right) = (2\pi)^{-1/2}\,\varpi_t^{-1}\,\exp\left(\frac{\left(i_t - i^* - \lambda\alpha_{t-1}\right)^2}{2\,\varpi_t^2}\right) \quad (4.6)$$

[36] Under strict UIP, we should also observe that $b_0 = 0$. However, this only represents a constant risk premium and we do not impose this restriction.

$$\text{where}\quad \varpi_t^2 = \left(\lambda^2+\psi_1\right)\sigma_{t-1}^2+\sigma_\epsilon^2+\psi_0+\psi_1\left(i_{t-1}-i^*-\alpha_{t-1}\right)^2$$

We can then maximize the sample log likelihood $\sum_{t=1}^{T}\log f_{\tilde{i}_t|I_{t-1}}(i_t|I_{t-1})$ with respect to the vector of parameters $\theta=(\lambda,\psi_0,\psi_1,\sigma_\epsilon,i^*)'$. Note that this estimator is also a Method of Moments estimator corresponding to the first order condition

$$E\left[\frac{\partial\log f(i_t|I_{t-1},\theta)}{\partial\theta}\right]=0$$

as long as the unconditional expectation of the score vector is well defined. We use monthly IFS data for G7 countries against the US, from 1974:1 to 1992:12. Our measure of short term interest rates is the short term money market rate (line 160bc) used in Grilli and Roubini (1994). In order to have a more flexible parameterization, we also allow for an $AR(p)$ in the persistent component and an $ARCH(q)$ in the transitory component. Thus we estimate the following model:

$$\begin{aligned} di_t = i_t - i_t^{us} &= i^* + di_t^p + \nu_t \quad ; \quad \nu_t|I_{t-1}\sim N(0,\sigma_\nu^2(t)) \qquad (4.7)\\ \lambda(L)\,di_t^p &= \epsilon_t \;;\; \epsilon_t\sim N(0,\sigma_\epsilon^2)\\ E\left[\nu_t^2|\nu_{t-1}\right] &= h_t^2=\psi_0+\psi(L)\,\nu_{t-1}^2;\\ \psi_0 &\geq 0,\ \psi_i\geq 0,\ \psi(1)<1 \end{aligned}$$

where λ is a polynomial in the lag operator of order p and ψ is of order q. The constant i^* captures the non zero mean of the interest rate differential, and in the context of the model it can be interpreted as the average constant return.

The estimation procedure is similar to section 2. In practice, increasing the order of the ARCH effect did not improve the results substantially. In contrast, increasing the order of the AR effect was crucial for correctly identifying the model. For low AR orders, the estimation procedure fails to exhibit significant transitory components (both the ARCH coefficients and ψ_0 are extremely small). We therefore report the results for an AR(3), ARCH(1) specification in Table 4.1.

The results indicate strong persistence in the interest differential. Indeed, for most countries, the sum of the AR coefficients is close to 0.95 and strongly significant.[37] In addition, the results indicate a very strong ARCH effect. In all cases, $\psi_1=1$. We plot in Figure 12 the persistent and transitory innovations to the interest rate differentials. One can see that the increase in volatility is associated with the 1979-1985 period and the associated change in the Fed's operating procedure. Interestingly, the volatility of the transitory component of the estimation procedure is strong for the US-UK interest rate differential. Comparatively, shocks to the US-Japan interest rate differential are much weaker. Table 4.1 reflects the comparatively larger ψ_0 estimated for these two countries. The US-Italian interest rate differential is the only one exhibiting a higher volatility of the persistent component.

[37] The only exception is US-Italy.

Table 4.1: MAXIMUM LIKELIHOOD ESTIMATION

U.S.vs	U.K.	France	Germany	Italy	Canada	Japan
λ_1	2.0372	2.3206	2.2374	1.7405	1.5693	2.0295
	(0.4849)	(0.1848)	(0.1150)	(0.3522)	(0.4216)	(0.2376)
λ_2	-1.4913	-1.9897	-1.9817	-1.1336	-1.0777	-1.5128
	(0.8172)	(0.3278)	(0.2018)	(0.5704)	(0.5846)	(0.4191)
λ_3	0.4282	0.6516	0.7388	0.3472	0.4385	0.4638
	(0.3699)	(0.1609)	(0.1059)	(0.2414)	(0.2493)	(0.2035)
ψ_0	0.8082	0.1516	0.1550	0.0552	0.3252	0.0658
	(0.5170)	(0.1002)	(0.1430)	(0.1544)	(0.2725)	(0.0448)
ψ_1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	(0.3983)	(0.2405)	(0.3864)	(0.6129)	(0.2453)	(0.1673)
σ_ϵ^2	0.5967	0.1313	0.2022	0.9878	0.7858	0.4242
	(0.5997)	(0.0813)	(0.0847)	(0.5348)	(0.6240)	(0.2074)
i^*	-1.5698	-1.5262	1.6352	-5.7499	-1.8215	1.9401
	(4.8681)	(2.3156)	(5.6382)	(2.5369)	(1.6417)	(3.3419)
log L	-0.6014	-0.3104	-0.3352	-0.3657	-0.4823	-0.2967
$\Psi(1)$	0.9740	0.9825	0.9945	0.9540	0.9301	0.9806
	(0.3983)	(0.2405)	(0.3864)	(0.6129)	(0.2453)	(0.1673)

Source: IFS.monthly money market rates. Sample: 1974:1-1992:12. Estimates equation 4.7 by maximum likelihood. Arch coefficients are constrained to be positive and $\Phi_1 \leq 1$. Standard errors are reported in parentheses. The updating equations are reported in the appendix for the general case AR(p), ARCH(q). Results for an AR(3) ARCH(3) were very similar and are available upon request.

5. Conclusion

We have presented a model of nominal exchange rate determination that exhibits the delayed overshooting pattern of exchange rates found by Eichenbaum and Evans (1995). Conditional delayed overshooting results from the interaction of learning about the current state of affairs and the intrinsic dynamic response of interest rates to monetary shocks. This interpretation, which is new to our knowledge, has important implications. First, it provides a clear analytical characterization of the factors influencing exchange rate responses to monetary shocks. Countries with rapidly converging interest rates, due to either fast moving prices or a large interest elasticity of money demand, will experience *less* delayed overshooting. Countries with either a very small or a very large variance of transitory shocks will also converge without delayed overshooting: in the former case because learning occurs fast, in the latter case because learning does not have a significant effect on the demand for assets.

Second, we have shown that a simple extension of our model can rationalize unconditional delayed overshooting and predictable excess returns. Our key assumptions is that the sample distribution and its market expectation differ, reflecting either small sample problems or expectational errors. Typically, unconditional delayed overshooting arises when shocks are more permanent than expected by market participants. We have found, using survey data on interest rate forecast that this assumption seem to be strongly supported: while the data fail to exhibit significant transitory components, market participant implicitly assume that a sizeable portion of the shocks is transitory. Moreover, estimating the model on monthly data for G7 countries, we found that our modelling assumptions accurately characterize the interest differential process and that the coefficients are often in the "delayed overshooting region."

Lastly, we have shown that a simple extension of our model can rationalize part of the forward discount puzzle without generating unconditional delayed overshooting. We found that deviations from Uncovered Interest Parity can arise if we make the additional assumption that the quality of the learning process is affected by the size of the shocks – that is, if shocks to the interest rate increase the variance of shocks. This assumption generates time varying risk premia, which in turn induce systematic deviations from uncovered interest parity.

References

Baily, Walter and Leo Kropywiansky, "A Test resolving the 'Peso Problem' in Foreign Exchange Using Option-Implied Jump Expectations," November 1994. M.I.T. mimeo.

Batten, Dallas, Michael Blackwell, In-Su Kim, Simon Nocera, and Yuzuru Ozeki, "The Conduct of Monetary Policy in the Major Industrial Countries," International Monetary Fund, Washington DC 1990. Occasional Paper 70.

Bollerslev, Tim, Ray Chou, and Kenneth Kroner, "ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics*, 1992, *52*, 5–59.

Brenner, Robin, Richard Harjes, and Kenneth Kroner, "Another Look at Alternative Models of the Short-Term Interest Rate," April 1994. mimeo, University of Arizona.

Clarida, Richard and Jordi Gali, "Sources of Real Exchange Rate Fluctuations: How Important Are Nominal Shocks?," 1994. NBER Working Paper 4658.

DeGroot, Morris, *Optimal Statistical Decision*, New York: McGraw-Hill, 1970.

Dornbusch, Rudiger, "Expectations and Exchange Rate Dynamics," *Journal of Political Economy*, 1976, *84*, 1161–1176.

Eichenbaum, Martin and Charles Evans, "Some Empirical Evidence on the Effects of Monetary Policy Shocks on Exchange Rates," *Quarterly Journal of Economics*, Nov. 1995, *110*, 975–1009.

Engle, Robert, David Lilien, and Russell Robins, "Estimating Time Varying Risk Premia in the Term Structure: the ARCH-M Model," *Econometrica*, 1987, *55(2)*, 391–407.

Fama, Eugene, "Forward and Sopt Exchange Rates," *Journal of Monetary Economics*, 1984, *14* (4), 319–338.

Frankel, Jeffrey and Andy Rose, "A Survey of Empirical Research on Nominal Exchange Rates," 1994. NBER Working Paper 4865.

__ **and Kenneth Froot**, "Understanding the US Dollar in the Eighties: the Expectations of Chartists and Fundamentalists," *Economic Record*, 1986, pp. 24–38.

__ **and** __ , "Forward Discount Bias: Is it an Exchange Risk Premium?," *Quarterly Journal of Economics*, 1989, *104*, 139–161.

Friedman, Benjamin, "Interest Rate Expectations versus Forward Rates: Evidence from an Expectations Survey," *Journal of Finance*, 1980, *34*, 965–973.

Froot, Kenneth, "New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates," *Journal of Finance*, 1989, *44* (2), 283–304.

Grier, Kevin and Mark Perry, "The Effect of Money Shocks on Interest Rates in the Presence of Conditional Heteroscedasticity," *Journal of Finance*, 1993, *48(4)*, 1445–55.

Grilli, Vittorio and Nouriel Roubini, "Liquidity and Exchange Rates: Puzzling Evidence From the G-7 Countries," 1994. mimeo Yale University.

Grossman, Sanford, "On the Efficiency of Competitive Stock Markets where Traders have Diverse Information," *Journal of Finance*, 1976, *31*, 573–585.

Hamilton, James, *Time Series Analysis*, Princeton University Press, 1994.

Hellwig, Martin, "On the Aggregation of Information in Competitive Markets," *Journal of Economic Theory*, 1980, *22*, 477–498.

Hodrick, Robert, "The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets," in "Fundamentals of Pure and Applied Economics" Harwood Academic Publishers 1988.

Kaminski, Graciela, "Is There a Peso Problem? Evidence form the Dollar/Pound Exchange Rate, 1976-1987," *American Economic Review*, 1993, *83*, 450–472.

Kim, Soyoung and Nouriel Roubini, "Liquidity and Exchange Rates in the G-7 Countries: Evidence from Identified VAR's," 1995. mimeo Yale University.

Lewis, Karen, "Can Learning Affect Exchange Rate Behavior? The Case of the Dollar in the Early 1980's," *Journal of Monetary Economics*, 1989, *23*, 79–100.

__ , "Changing Beliefs and Systematic Rational Forecast Errors with Evidence from Foreign Exchange," *American Economic Review*, 1989, *79*, 621–636.

__ , "Puzzles in International Finance," 1994. NBER Working Paper 4951.

Lucas, Robert, "Some Evidence on the Output-Inflation Tradeoff," *American Economic Review*, 1973, *63*, 326–334.

Mehra, Rajnish and Edward Prescott, "The Equity Premium Puzzle," *Journal of Monetary Economics*, 1985, *15*, 145–162.

Mussa, Michael, "Adaptive and Regressive Expectations in a Rational Model of the Inflationary Process," *Journal of Monetary Economics*, 1975, *1*, 423–442.

Muth, John, "Optimal properties of exponentially weighted forecasts," *journal of the american statistical association*, 1960, *55*, 299–306.

Rebello, Joseph, "Dollar Plunges as Investors Conclude New Interest Rate Cut Isn't Very Likely," *Wall Street Journal*, 1995, July 20.

Wang, Jiang, "A Model of Competitive Stock Trading Volume," *Journal of Political Economy*, 1994, *102*, 127–168.

Wessel, David, "Fed Was Split Over Rate Cut at July Meeting," *Wall street journal*, 1995, August 28.

__ , "Fed Meeting on Rate Cut May Be Pivotal," *Wall Street Journal*, 1995, July 5.

__ , "Rate Reversal: After Long Tightening, the Fed Loosens Up with a 1/4 Point Cut;," *Wall Street Journal*, 1995, July 7.

__ **and Anita Raghavan**, "Open Market: a Growing Glasnost at the Fed Is Dispelling a Lot of its Mystique," *Wall street journal*, 1994, March 24.

Appendices

A. Delayed Overshooting

We replicated the results of Eichenbaum and Evans (1995) using a slightly longer time period. The data are monthly, from the IFS, and the sample period is 1974:1 to 1992:12. Eichenbaum and Evans covers 1974:1 to 1990:5 while Grilli and Roubini covers 1974:1 to 1991:12. The interest rates are monthly market rates. The exchange rates are quoted as units of foreign currency per dollar. The paper estimates recursive VARS. See Kim and Roubini (1995) and Clarida and Gali (1994) for structural VAR, identified with contemporaneous and long run restrictions respectively. The countries in our sample are: UK, France, Germany, Italy, Canada and Japan.

Eichenbaum and Evans (1995) consider three measures of monetary policy: the ratio of Non Borrowed Reserves to total reserves, the Federal Funds rate, and the Romer and Romer index of monetary policy. They look at the response of exchange rates to innovations in their measure of monetary policy. We replicated two of their specifications:

- $\{Y_{us}, CPI_{us}, NBRX_{us}, R^* - R_{us}, E\}$. In this specification, the innovations to the ratio of Non Borrowed Reserves represent monetary shocks. The Fed observes domestic industrial production and prices. All variables are in level form. Figure 1a-b show the impulse response of the interest rate differential (top row) and the nominal exchange rate (bottom row) for 6 G7 countries. In all cases, the path of interest differential is consistent with a Dornbusch-type experiment. The exchange rate response exhibits substantial inertia (especially in the cases of Japan, Canada and the UK). We report standard deviation bands around point estimates which were computed using a Monte Carlo method with 500 draws.

- $\{Y_{us}, CPI_{us}, Y^*, R^*, FF, NBRX_{us}, E\}$. In this specification, the Fed also observes foreign output and interest rates before setting the federal funds rate. We plot the response of Non Borrowed Reserves to total reserves (NBRX in the top row) and the nominal exchange rate (bottom row). See Figure 1c-d. Following an increase in the Federal Funds rate, the ratio of Non Borrowed Reserves dips down and then increases over time. The nominal exchange rate exhibits a pattern closely resembling the unconditional impulse response we obtained in section 3. On impact, the exchange rate appreciates, then it depreciates rapidly, sometimes falling below its original level. After 5 to 10 periods, the delayed overshooting pattern emerges.

B. Empirical Results on Interest Rate Differentials

We describe in this section the empirical procedure. The survey data are described in section 2 in the paper. We first run ARMA processes of various orders and select the ones that minimize Akaike's criterion. This procedure is unconstrained for both the order of the AR and MA component (the maximum order is 5 on each component). The

Table B.1: ARMA, Euro 3 months

	UK-US	France-US	Germany-US	Italy-US	Canada-US	Japan-US
λ_1	0.934	0.855	1.418	0.738	0.892	1.150
	(0.028)	0.035	(0.151)	(0.103)	(0.042)	(0.068)
λ_2			-0.427	0.358		-0.218
			(0.146)	(0.123)		(0.068)
λ_3				-0.155		
				(0.103)		
ϕ_1	-0.238		0.066			
	(0.075)		(0.132)			
ϕ_2	0.238		0.543			
	(0.075)		(0.073)			
AIC	591.32	849.58	536.20	232.75	186.33	575.27

Note: ARMA estimation by Maximum Likelihood. Sample Period: january 1974-september 1995, 261 monthly observations. Robust standard errors in parentheses.

results are reported in Table B.1, B.2 and B.3. Results using Euro-3 months interest rates report robust standard errors.

In all specifications but UK-US and Germany-US for the euro 3 months, the order of the moving average is inferior to the order of the autoregressive part and satisfies the overidentification restrictions. In a majority of cases, there is no moving average component indicating that there is no associated transitory component.

B.1. Kalman Filter Estimation

This subsection briefly derives the Kalman Filter equations. We postulate the following

Table B.2: ARMA, Money Market Rates

	UK-US	France-US	Germany-US	Italy-US	Canada-US	Japan-US
λ_1	0.742	0.901	0.946	0.9103	0.7249	1.391
	(0.061)	(0.028)	(0.019)	(0.027)	(0.043)	(0.056)
λ_2	0.187					-0.439
	(0.061)					(0.056)
ϕ_1		-0.283		-0.385		
		(0.063)		(0.062)		
AIC	1014.65	690.24	759.92	689.94	837.11	612.31

Note: ARMA estimation by Maximum Likelihood. Sample Period: january 1974-september 1995, 261 monthly observations. Standard errors in parentheses.

Table B.3: ARMA, PRIME RATE

	UK-US	Canada-US	Japan-US
λ_1	1.171	0.942	0.970
	(0.099)	(0.034)	(0.097)
λ_2	-0.195		0.308
	(0.100)		(0.137)
λ_3			-0.301
			(0.101)
AIC	138.70	134.98	105.24

Note: ARMA estimation by Maximum Likelihood. Sample Period: september 1987-october 1995, 97 monthly observations. Standard errors in parentheses.

process:

$$di_t = H'\xi_t + \nu_t \tag{B.1}$$
$$\xi_t = F\xi_{t-1} + \epsilon_t \tag{B.2}$$

where $\xi_t = \left(di_t^p, ..., di_{t-p+1}^p\right)'$, $H' = (1, 0, ..., 0)'$ is a px1 vector. ξ_t is the state vector for the process, (B.1) the measurement equation and (B.2) the space equation. Define the informations set $I_t = \{di_{t-i}, i \geq 0\}$, $\hat{\xi}_{t+1|t} = E\left[\xi_{t+1}|I_t\right]$, and $\hat{P}_{t+1|t} = E\left[\left(\xi_{t+1} - \hat{\xi}_{t+1|t}\right)\left(\xi_{t+1} - \hat{\xi}_{t+1|t}\right)'|I_t\right]$.

The filtering equations are:

$$\begin{aligned}\hat{\xi}_{t+1|t} &= F\hat{\xi}_{t-1|t} + F\hat{P}_{t|t-1}H\left(H'\hat{P}_{t|t-1}H + \sigma_\nu^2\right)^{-1}\left(di_t - H'\hat{\xi}_{t|t-1}\right) \\ \hat{P}_{t+1|t} &= F\left[\hat{P}_{t|t-1} - \hat{P}_{t|t-1}H\left(H'\hat{P}_{t|t-1}H + \sigma_\nu^2\right)^{-1}H'\hat{P}_{t|t-1}\right]F' + \sigma_\epsilon^2\end{aligned}$$

The smoother equations are:

$$\begin{aligned}\hat{\xi}_{t|T} &= \hat{\xi}_{t|t} + \hat{P}_{t|t}F'\hat{P}_{t+1|t}^{-1}\left(\hat{\xi}_{t+1|T} - \hat{\xi}_{t+1|t}\right) \\ \hat{P}_{t|T} &= \hat{P}_{t|t} + \left(\hat{P}_{t|t}F'\hat{P}_{t+1|t}^{-1}\right)\left(\hat{P}_{t+1|T} - \hat{P}_{t+1|t}\right)\left(\hat{P}_{t|t}F'\hat{P}_{t+1|t}^{-1}\right)'\end{aligned}$$

where

$$\hat{P}_{t|t} = \hat{P}_{t|t-1} - \hat{P}_{t|t-1}H\left(H'\hat{P}_{t|t-1}H + \sigma_\nu^2\right)^{-1}H'\hat{P}_{t|t-1}$$

C. Proofs.

Proof of Lemma 3.1:
In order to prove this Lemma we will use a standard result in Bayesian inference (see DeGroot (1970, section 9.5)). Suppose that y_t is an observation from a normal sampling distribution with unknown mean θ and known variance ρ^2. If the prior distribution of θ is normal with mean m_{t-1} and variance τ_{t-1}^2, then the posterior distribution of θ after observing y_t is normal with mean m_t and variance τ_t^2, where

$$m_t = (1-k_t)m_{t-1} + k_t y_t, \qquad \tau_t^2 = (1-k_t)\tau_{t-1}^2, \qquad k_t = \frac{\tau_{t-1}^2}{\tau_{t-1}^2 + \rho^2} \tag{C.1}$$

To apply this result to our model note that the observation y_t corresponds to $i_t - i^*$ and that θ corresponds to i_t^p, and recall the notation $E\left[i_{t-1}^p \,|I_{t-1}\right] = \tilde{\alpha}_{t-1}$ and $E\left[(\tilde{\alpha}_{t-1} - i_{t-1}^p)^2 \,|I_{t-1}\right] = \tilde{\sigma}_{t-1}^2$. First, since the young investors know that $i_{t-1}^p \,|I_{t-1} \sim N(\tilde{\alpha}_{t-1}, \tilde{\sigma}_{t-1})$, and since $i_t^p = \lambda i_{t-1}^p + \epsilon_t$, their prior about i_t^p (before observing i_t) is $i_t^p \,|I_{t-1} \sim N(\tilde{\alpha}_{t|t-1}, \tilde{\sigma}_{t|t-1}^2)$, where

$$\begin{aligned} \tilde{\alpha}_{t|t-1} &= E\left[i_t^p \,|I_{t-1}\right] = \lambda\tilde{\alpha}_{t-1} \\ \tilde{\sigma}_{t|t-1}^2 &= E\left[(\tilde{\alpha}_{t|t-1} - i_t^p)^2 \,|I_{t-1}\right] = E\left[(\lambda\tilde{\alpha}_{t-1} - \lambda i_{t-1}^p - \epsilon_t)^2 \,|I_{t-1}\right] = \lambda^2\tilde{\sigma}_{t-1}^2 + \tilde{\sigma}_\epsilon^2 \end{aligned} \tag{C.2}$$

Second, note that the observation $i_t - i^* = i_t^p + \nu_t$ is normal with unknown mean i_t^p and known variance $\tilde{\sigma}_\nu^2$. Lastly, using the above result it follows that after observing i_t the young investors' posterior is $i_t^p \,|I_t \sim N(\tilde{\alpha}_t, \tilde{\sigma}_t^2)$, where $\tilde{\alpha}_t$ and $\tilde{\sigma}_t^2$ are given by (3.7). The equations in (3.7) are obtained by substituting $(i_t - i^*,\ \lambda\tilde{\alpha}_{t-1},\ \tilde{\alpha}_t,\ \tilde{\sigma}_{t|t-1}^2,\ \tilde{\sigma}_t^2,\ \tilde{\sigma}_\nu^2)$ for $(y_t,\ \mu_{t-1},\ \mu_t,\ \tau_{t-1}^2,\ \tau_t^2,\ \rho^2)$ in (C.1).

Proof of Lemma 3.2
To prove this lemma, we simply solve for the constant gain and variance. Dividing both sides by $\tilde{\sigma}_\epsilon^2$ we find that the gain and variance only depend on the noise to signal ratio. The derivatives are straightforward.

Proof of Lemma 3.3
Using the conjectured price function to eliminate f_{t+1} from the objective function, we get:

$$\begin{aligned} U(c_t, c_{t+1}) = {} & -\exp(-\gamma c_t) - \\ & \delta \exp\left(-\gamma\left[W_t - f_t x_t - c_t\right]\left[1+i^*\right] - \gamma x_t\left[a + i_t\right]\right) E\left[\exp\left(-\gamma x_t B'\theta_{t+1}\right)\right] \end{aligned} \tag{C.3}$$

where $B' = (b, c)$ and $\theta_{t+1}' = (\tilde{\alpha}_{t+1}, i_{t+1})$. The vector θ_{t+1} is normally distributed with mean $\bar{\theta}_{t+1}$ and covariance matrix Σ

$$\bar{\theta}_{t+1} = \begin{bmatrix} \lambda\tilde{\alpha}_t \\ \lambda\tilde{\alpha}_t + i^* \end{bmatrix} \qquad \sum = \left(\lambda^2\tilde{\sigma}^2 + \tilde{\sigma}_\epsilon^2 + \tilde{\sigma}_\nu^2\right)\begin{bmatrix} \tilde{k}^2 & \tilde{k} \\ \tilde{k} & 1 \end{bmatrix} \tag{C.4}$$

Therefore, the expectation in (C.3) is given by

$$\begin{aligned} E\left[\exp(-\gamma x_t B'\theta_{t+1}) \,|I_t\right] &= \int_{-\infty}^{\infty} \frac{\exp\left(-\frac{1}{2}(\theta_{t+1} - \bar{\theta}_{t+1})'\Sigma^{-1}(\theta_{t+1} - \bar{\theta}_{t+1})\right)\exp(-\gamma x_t B'\theta_{t+1})}{|\Sigma|\sqrt{2\pi}} \\ &= \exp\left(-\gamma x_t B'\bar{\theta}_{t+1} + \frac{1}{2}\gamma^2 x_t^2 B'\Sigma B\right)\int_{-\infty}^{\infty} \frac{\exp\left(-\frac{1}{2}(\theta_{t+1} - m)'\Sigma^{-1}(\theta_{t+1} - m)\right)}{|\Sigma|\sqrt{2\pi}} \end{aligned} \tag{C.5}$$

where $m = \overline{\theta}_{t+1} - \Sigma\gamma x_t B$. The second equality follows from the fact that covariance matrix Σ is symmetric. Since the integrand in the second equation in (C.5) is the density function of a normal variable, it follows that the integral is equal to one. Thus, the expectation in (C.3) is equal to $\exp\left(-\gamma x_t B'\bar{\theta}_{t+1} + \frac{1}{2}\gamma^2 x_t^2 \pi^2\right)$ where $\pi^2 = B'\Sigma B$. Substituting this expression in (C.3) and taking the derivative with respect to x_t we obtain equation (3.12).

Derivation of Lemma 3.4

Using (3.7) to eliminate α_t from (3.14) and substituting $i^* + \lambda^t\epsilon_0$ for i_t, it follows that

$$\begin{aligned} f_t^p &= \bar{f} + \tfrac{\lambda^t\epsilon_0}{1+i^*} + \tfrac{\lambda^{t+1}}{[1+i^*][1+i^*-\lambda]}\left[\sum_{j=0}^{t} k(1-k)^j\epsilon_0 + \lambda(1-k)^{t+1}\alpha_{-1}\right] \\ &= \bar{f} + \tfrac{\lambda^t\epsilon_0}{[1+i^*][1+i^*-\lambda]}\left[1 + i^* - \lambda(1-k)^{t+1}\right] \end{aligned} \tag{C.6}$$

The second equality is obtained by carrying out the summation and setting $\alpha_{-1} = 0$. Equation (3.19) is obtained by using (C.6) to compute $f_{t+1}^p - f_t^p$ and by setting $\epsilon_0 = \kappa$.

Derivation of Proposition 3.2:

The first part of the proposition is trivial. To find the frontier of D_T, we set (3.20) equal to zero. This gives a second order equation in λ. We select the root smaller than 1, which is given by (3.21).

Derivation of Equations (3.23) and (3.24)

Denote by $f_t(\epsilon, \nu)$ the impulse response at time t to an initial persistent shock of size ϵ accompanied by a transitory shock of size ν. The unconditional path (3.23) is given by

$$\begin{aligned} f_t^u &= \int f_t(\epsilon_0, \kappa - \epsilon_0)\; d\varphi(\epsilon_0 | i_0 = \kappa) = \int (f_t(\epsilon_0, 0) + f_t(0, \kappa - \epsilon_0))\; d\varphi(\epsilon_0 | i_0 = \kappa) \\ &= f_t^p(\kappa)\; E[\epsilon_0 | i_0 = 1] + f_t^t(\kappa)\; E[\nu_0 | i_0 = 1] \end{aligned} \tag{C.7}$$

The second and third equalities follow from the linearity of the impulse response in the original shocks. To derive the expression for $E[\epsilon_0 | i_0 = i^* + \kappa]$ in (3.23) we use formula (C.1). Let us consider $i_0 - i^* = \lambda i_{-1}^p + \nu_0 + \epsilon_0$ as an observation from a distribution parameterized by ϵ_0. Note that since the system was in steady state (i.e., the expectation of i_{-1}^p was zero), it follows that $i_0 - i^* | \epsilon_0 \sim N(\epsilon_0,\, \lambda^2\sigma^2 + \sigma_\nu^2)$. Note also that the prior distribution of ϵ_0 is normal with mean zero and variance σ_ϵ^2. Thus, substituting ($i_0 - i^*$, 0, $E[\epsilon_0 | i_0 - i^*]$, σ_ϵ^2, $var[\epsilon_0 | i_0 - i^*]$, $\lambda^2\sigma^2 + \sigma_\nu^2$) for ($y_t$, μ_{t-1}, μ_t, τ_{t-1}^2, τ_t^2, ρ^2) in (C.1), it follows that the posterior distribution of ϵ_0 is $\epsilon_0 | i_0 - i^* \sim N\left(\frac{\sigma_\epsilon^2[i_0 - i^*]}{\lambda^2\sigma^2 + \sigma_\epsilon^2 + \sigma_\nu^2}, \frac{\sigma_\epsilon^2[\lambda^2\sigma^2 + \sigma_\nu^2]}{\lambda^2\sigma^2 + \sigma_\epsilon^2 + \sigma_\nu^2}\right)$. By setting $i_0 - i^* = \kappa$ we obtain (3.23). To obtain (3.24) note that the impulse response to a persistent shock is given by (C.6). Substituting $i_0 = i^* + \kappa$, $i_t = i^*$ for $t > 0$ and $\alpha_t = \lambda^t(1-k)^t k\kappa$ in (3.14), it follows that the impulse response to a transitory shock of size κ is:

$$f_t^t(\kappa) = \bar{f} + \frac{\lambda^{t+1}(1-k)^t k\kappa}{[1+i^*][1+i^*-\lambda]} \tag{C.8}$$

By substituting (C.6) and (C.8) in (C.7), and taking first differences we obtain (3.24).

Proof of Lemma 4.1:

Since $q = \frac{\sigma_\epsilon^2}{\lambda^2\sigma^2+\sigma_\epsilon^2+\sigma_\nu^2} < \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2+\sigma_\nu^2} = \frac{1}{1+\eta}$, it is sufficient to show that $\frac{1}{1+\eta} < k$. Using the definition of k, in (3.8), we get

$$\frac{1}{1+\eta} < k = \frac{1+\Delta-\eta\left(1-\lambda^2\right)}{1+\Delta+\eta\left(1+\lambda^2\right)}$$

By rearranging it follows that this inequality holds if and only if $1+\eta\left(1-\lambda^2\right) < \Delta$. Lastly, the definition of Δ in (3.8) implies that this inequality always holds. Thus, $q < k$.

Proof of Lemma 4.1:

The problem is the same as that of Lemma 3.1. The only difference is the distribution of $i_t - i^*$. We will show that $i_t - i^*$ is normally distributed with unknown mean i_t^p and known variance given by

$$E\left[\nu_{t-1}^2 \left| I_{t-1}\right.\right] = \xi_0 + \xi_1\left[\left(i_{t-1} - i^* - \alpha_{t-1}\right)^2 + \sigma_{t-1}^2\right] \tag{C.9}$$

Recall that $\alpha_{t-1} = \left[i_{t-1}^p \left| I_{t-1}\right.\right]$ and $\sigma_{t-1}^2 = E\left[\left(i_{t-1}^p - \alpha_{t-1}\right)^2 \left| I_{t-1}\right.\right]$. To derive (C.9) note first that from (4.1) and the law of iterated expectations it follows that

$$E\left[\nu_t^2 \left| I_{t-1}\right.\right] = E\left[E\left[\nu_t^2 \left|\nu_{t-1}, I_{t-1}\right.\right] \left| I_{t-1}\right.\right] = E\left[h_t^2 \left| I_{t-1}\right.\right] = \xi_0 + \xi_1 E\left[\nu_{t-1}^2 \left| I_{t-1}\right.\right] \tag{C.10}$$

Second, to derive $E\left[\nu_{t-1}^2 \left| I_{t-1}\right.\right]$ note that

$$\begin{aligned} E\left[\nu_{t-1}^2 \left| I_{t-1}\right.\right] &= E\left[\left(i_{t-1} - i^* - i_{t-1}^p\right)^2 \left| I_{t-1}\right.\right] & \text{(C.11)} \\ &= E\left[\left(i_{t-1} - i^*\right)^2 - 2\left(i_{t-1} - i^*\right) i_{t-1}^p + \left(i_{t-1}^p\right)^2 \left| I_{t-1}\right.\right] \\ \sigma_{t-1}^2 &= E\left[\left(i_{t-1}^p - \alpha_{t-1}\right)^2 \left| I_{t-1}\right.\right] = E\left[\left(i_{t-1}^p\right)^2 \left| I_{t-1}\right.\right] - \alpha_{t-1}^2 & \text{(C.12)} \end{aligned}$$

Using (C.12) to eliminate $E\left[\left(i_{t-1}^p\right)^2 \left| I_{t-1}\right.\right]$ from (C.11) and substituting the result in (C.10) we obtain (C.9). Next, following the same steps as in the proof of Lemma 1 and using the same notation as in (C.2), we have that the prior distribution of i_t^p is normal with mean $\alpha_{t|t-1} = \lambda\alpha_{t-1}$ and variance $\sigma_{t|t-1}^2 = \lambda^2\sigma_{t-1}^2 + \sigma_\epsilon^2$. Lastly, to obtain the mean and variance of the posterior distribution of i_t^p we substitute ($i_t - i^*$, $\alpha_{t|t-1}$, α_t, $\sigma_{t|t-1}^2$, σ_t^2, $E\left[\nu_t^2 \left| I_{t-1}\right.\right]$) for ($y_t$, μ_{t-1}, μ_t, τ_{t-1}^2, τ_t^2, ρ^2) in (C.1). The resulting expressions are given by (4.2).

FIGURE 1-a

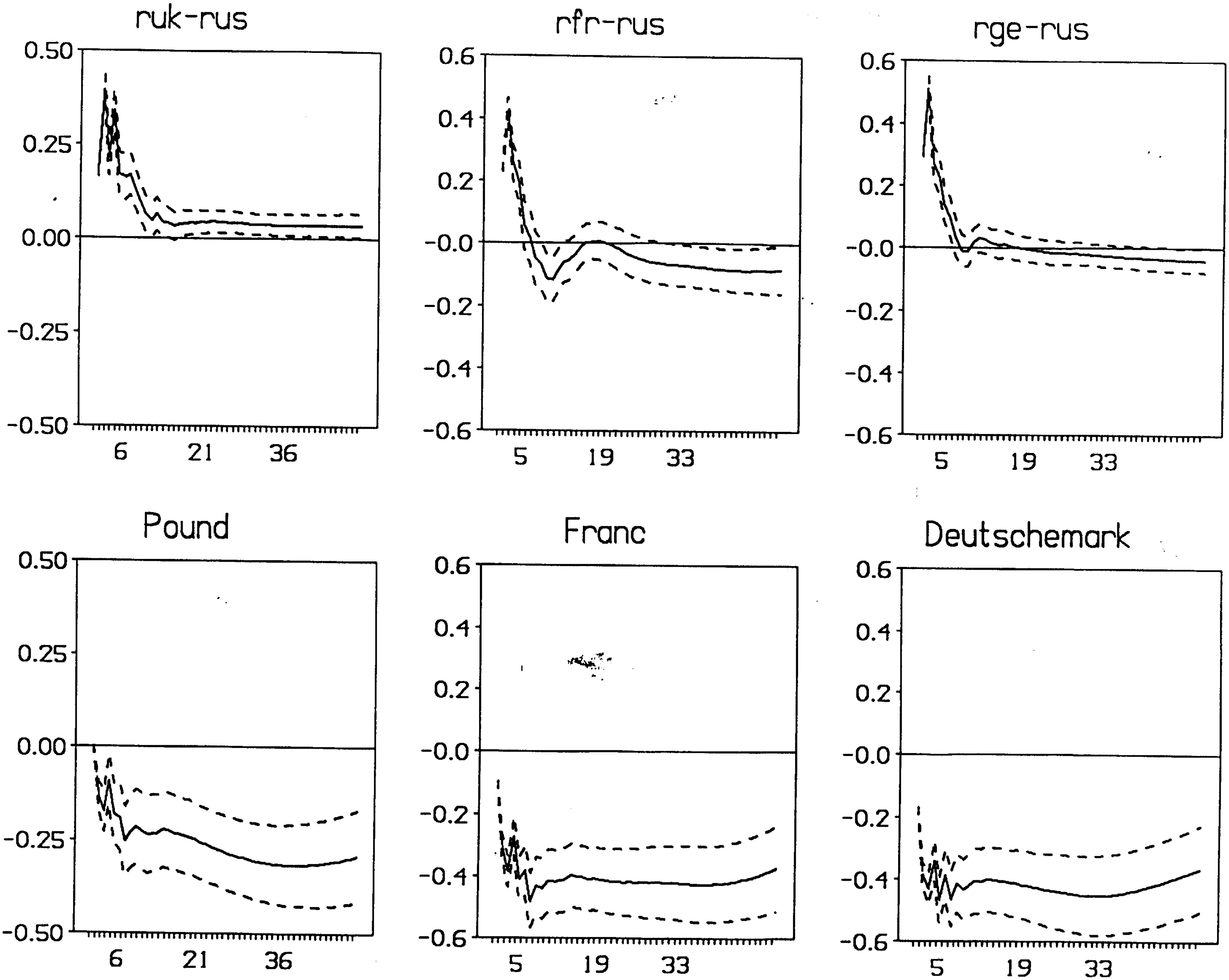
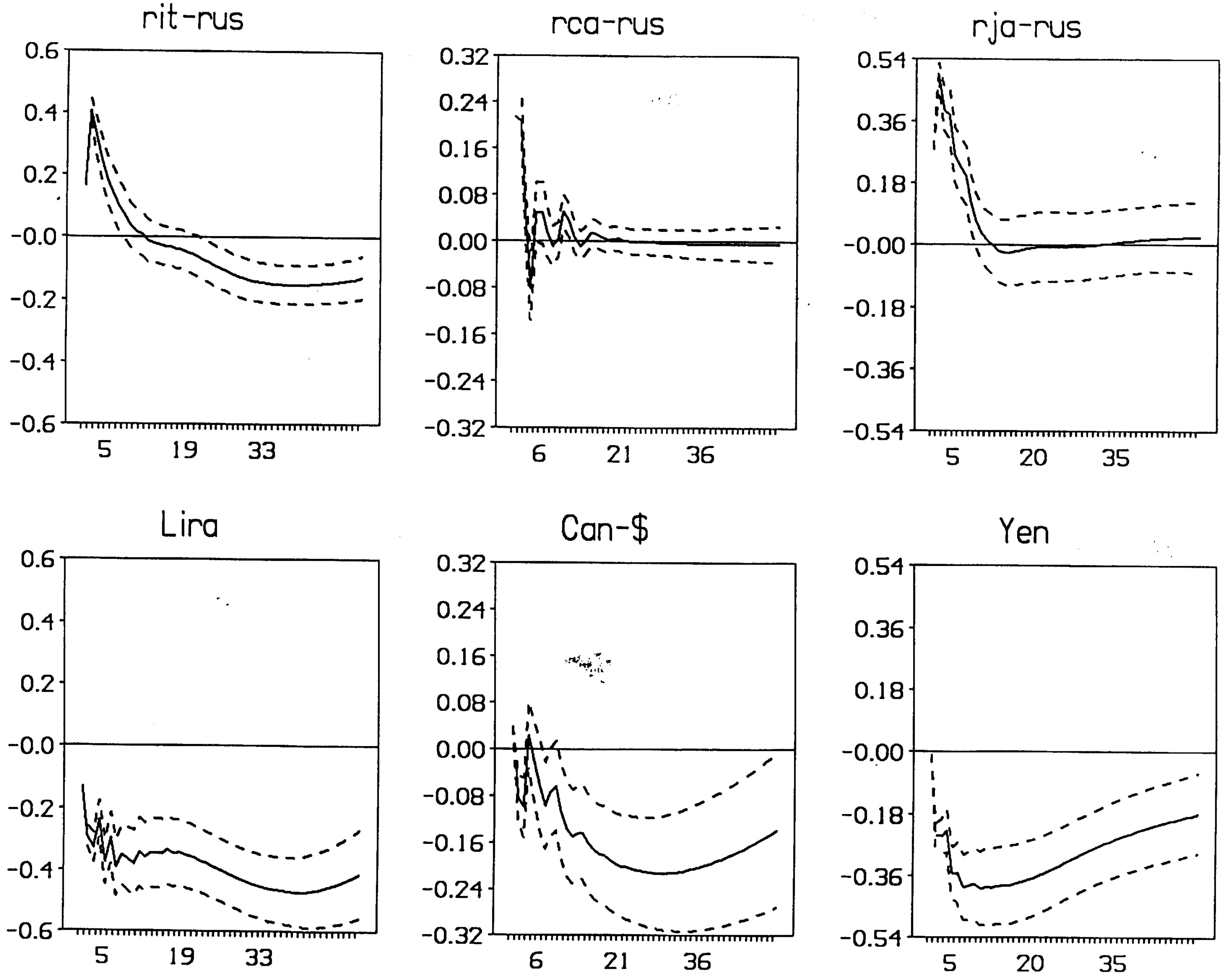
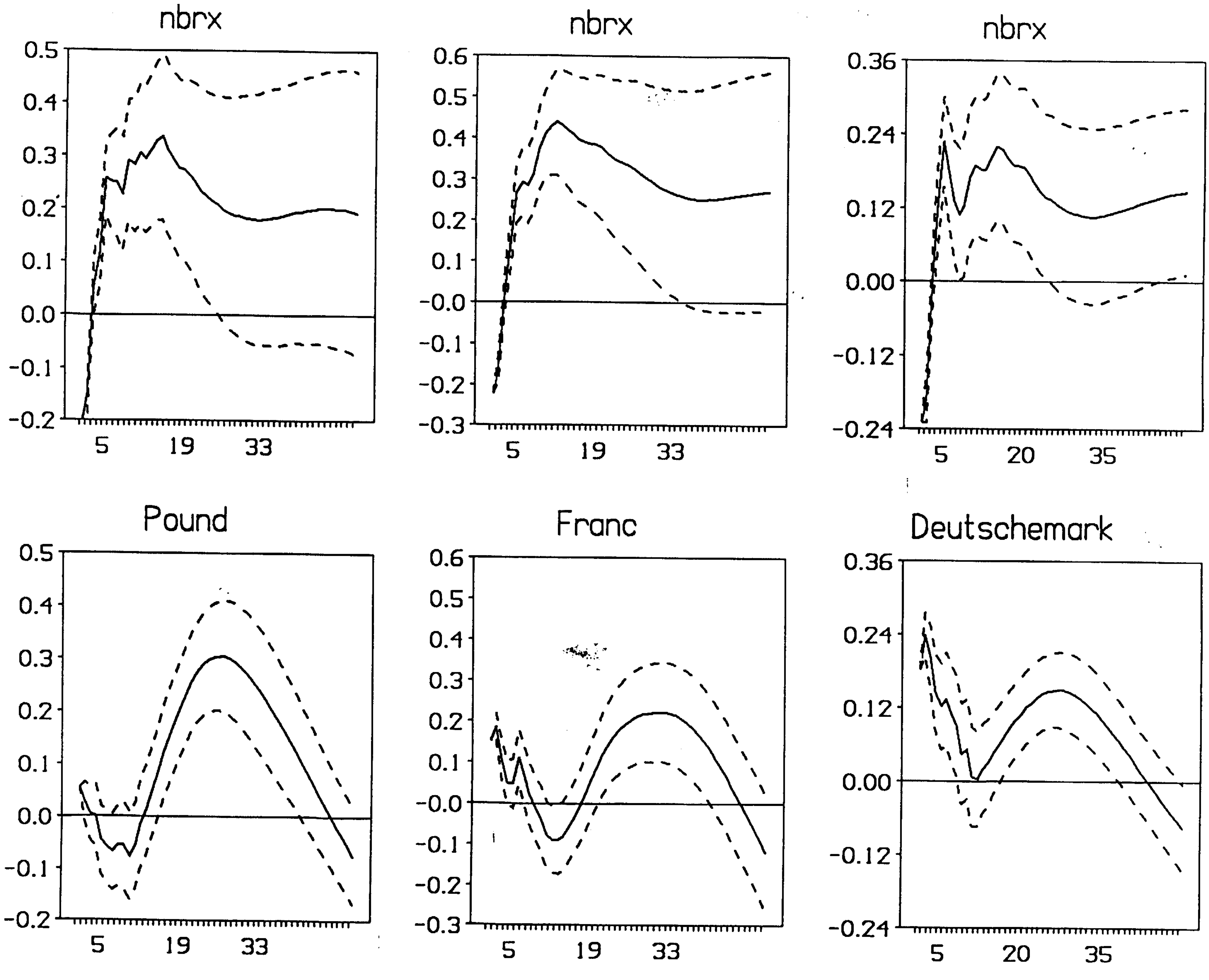

ruk-rus
0.50
0.25
0.00
-0.25
-0.50
6
21
36
rfr-rus
0.6
0.4
0.2
-0.0
-0.2
-0.4
-0.6
5
19
33
rge-rus
0.6
0.4
0.2
-0.0
-0.2
-0.4
-0.6
5
19
33
Pound
0.50
0.25
0.00
-0.25
-0.50
6
21
36
Franc
0.6
0.4
0.2
-0.0
-0.2
-0.4
-0.6
5
19
33
Deutschemark
0.6
0.4
0.2
-0.0
-0.2
-0.4
-0.6
5
19
33

FIGURE 1.b



rit-rus
0.6
0.4
0.2
-0.0
-0.2
-0.4
-0.6
5
19
33
rca-rus
0.32
0.24
0.16
0.08
0.00
-0.08
-0.16
-0.24
-0.32
6
21
36
rja-rus
0.54
0.36
0.18
-0.00
-0.18
-0.36
-0.54
5
20
35
Lira
0.6
0.4
0.2
-0.0
-0.2
-0.4
-0.6
5
19
33
Can-$
0.32
0.24
0.16
0.08
0.00
-0.08
-0.16
-0.24
-0.32
6
21
36
Yen
0.54
0.36
0.18
-0.00
-0.18
-0.36
-0.54
5
20
35

FIGURE 1.c



nbrx
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
5
19
33
nbrx
0.6
0.5
0.4
0.3
0.2
0.1
-0.0
-0.1
-0.2
-0.3
5
19
33
nbrx
0.36
0.24
0.12
0.00
-0.12
-0.24
5
20
35
Pound
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
5
19
33
Franc
0.6
0.5
0.4
0.3
0.2
0.1
-0.0
-0.1
-0.2
-0.3
5
19
33
Deutschemark
0.36
0.24
0.12
0.00
-0.12
-0.24
5
20
35

FIGURE 1-d

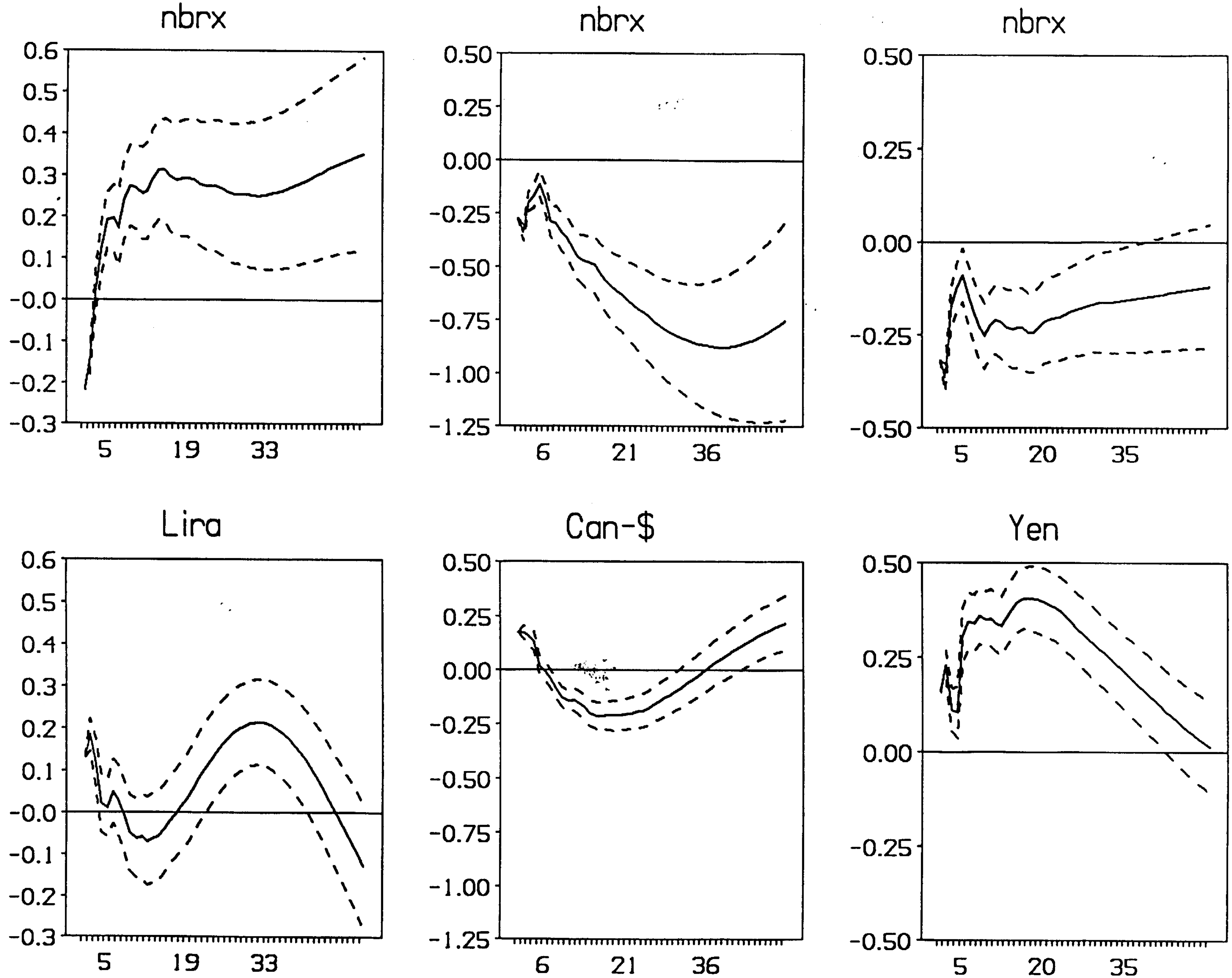


Figure 2.A: Money Market and Prime Rate
US

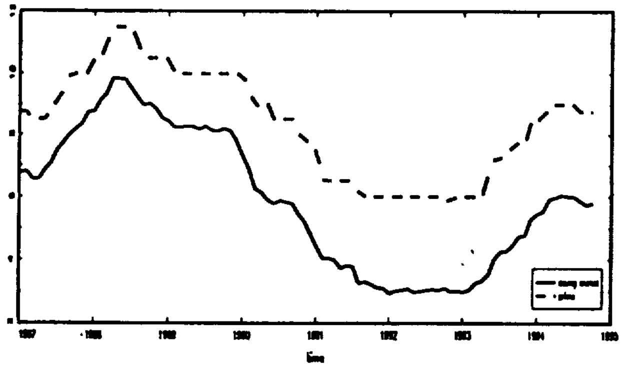


Figure 2.B: Money Market and Prime Rate
UK

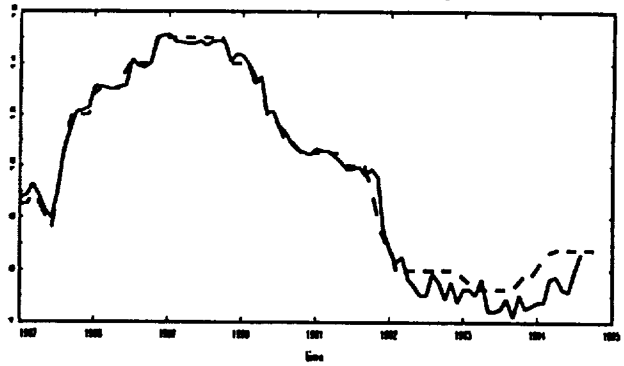


Figure 2.C: Money Market and Prime Rate
FR

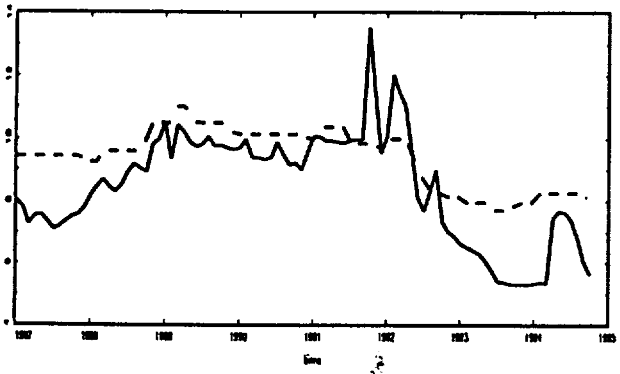


Figure 2.D: Money Market and Prime Rate
GE

Figure 2.E: Money Market and Prime Rate
IT

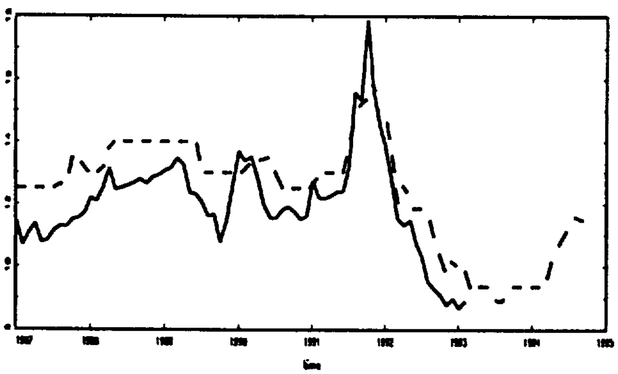


Figure 2.F: Money Market and Prime Rate
CA

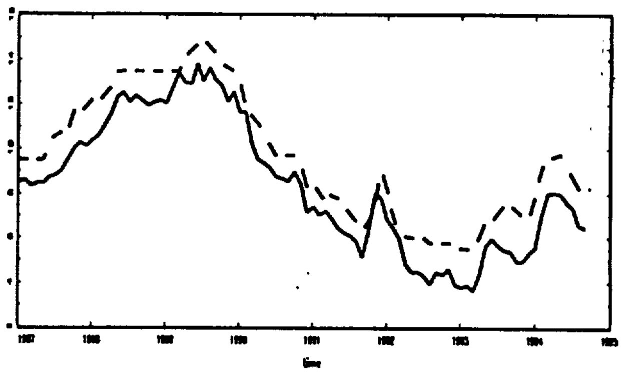


Figure 2.G: Money Market and Prime Rate
JA

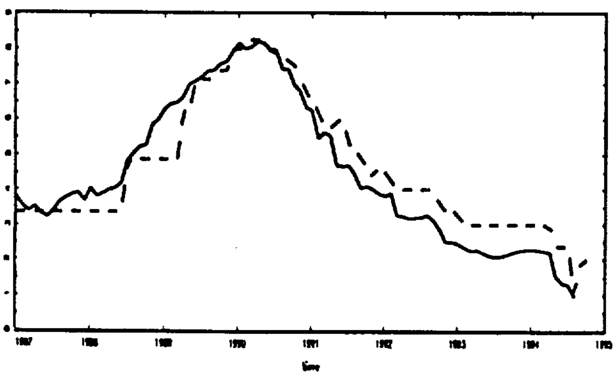


Figure 3.A: Money Market and Euro-3months
US

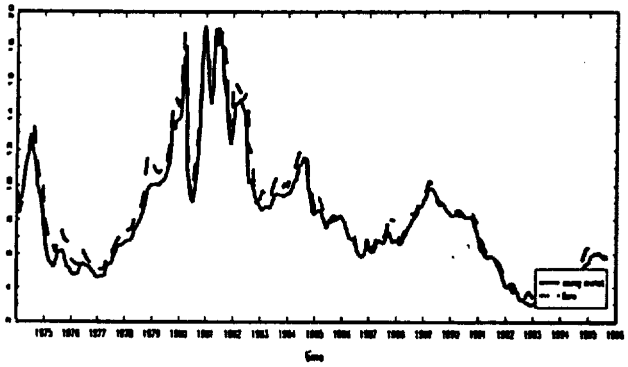


Figure 3.B: Money Market and Euro-3months
UK

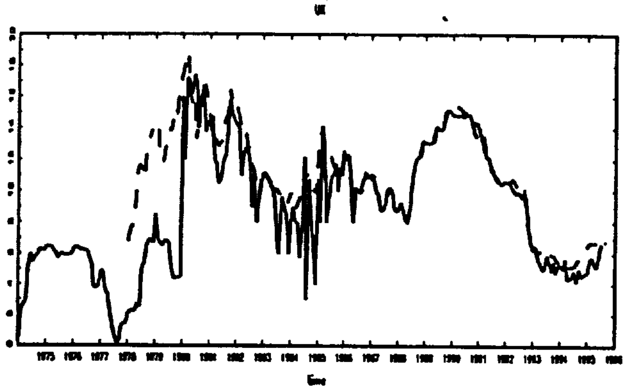


Figure 3.C: Money Market and Euro-3months
FR

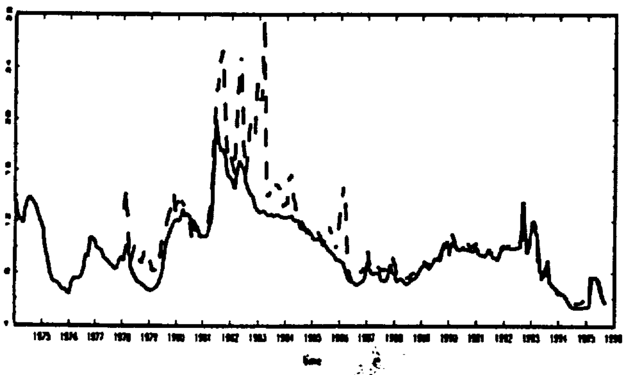


Figure 3.D: Money Market and Euro-3months
GE

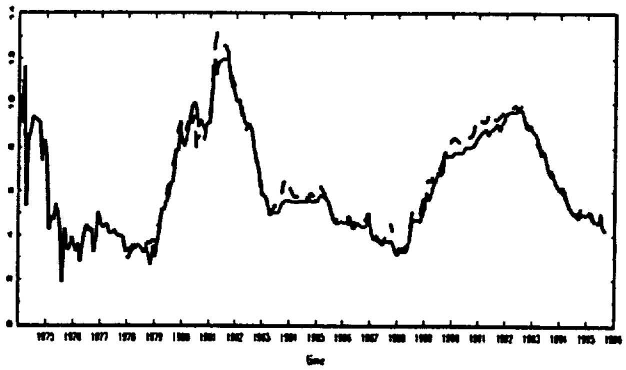


Figure 3.E: Money Market and Euro-3months
IT

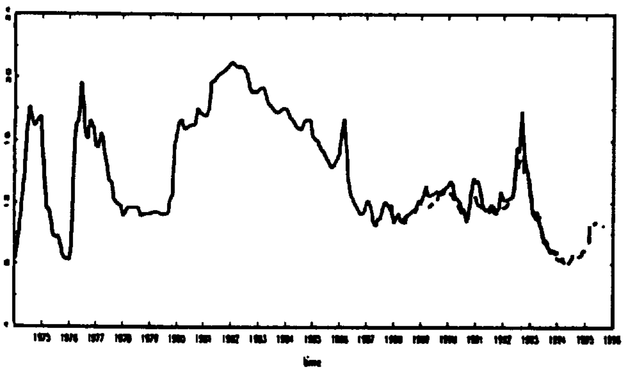


Figure 3.F: Money Market and Euro-3months
CA

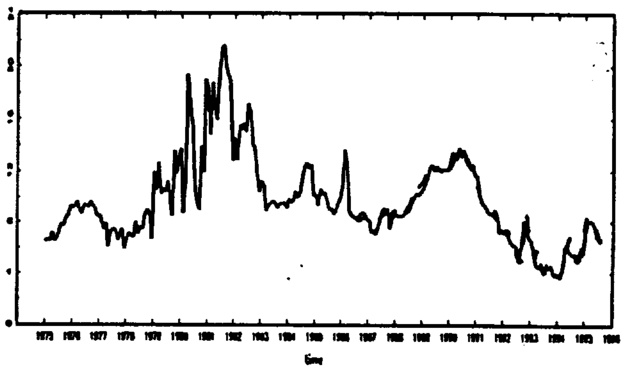


Figure 3.G: Money Market and Euro-3months
JA

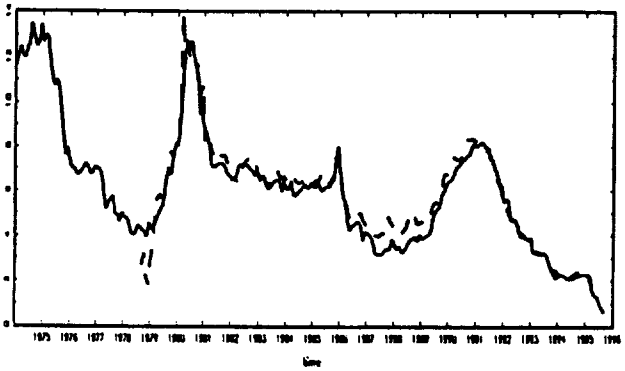


Figure 4.A: Euro 3 months and Forecasts
US-UK

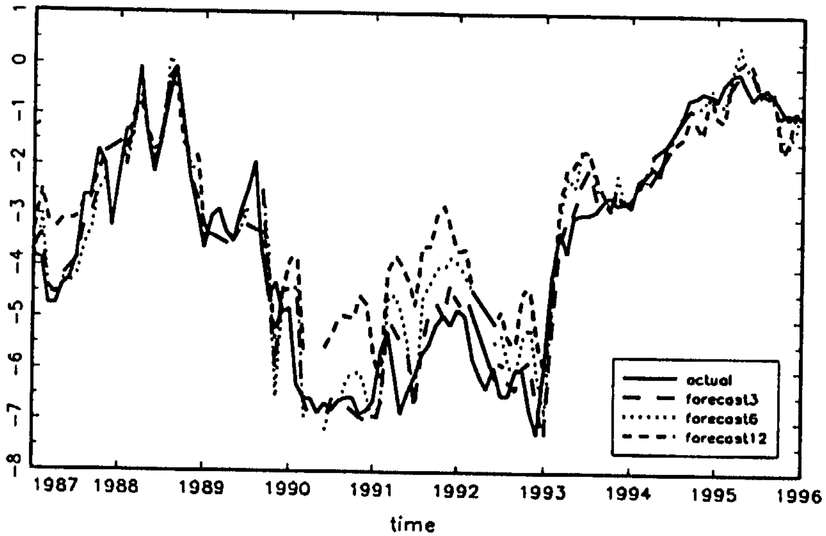


Figure 4.B: Euro 3 months, Forecast 3 Months
US-UK; AR=3

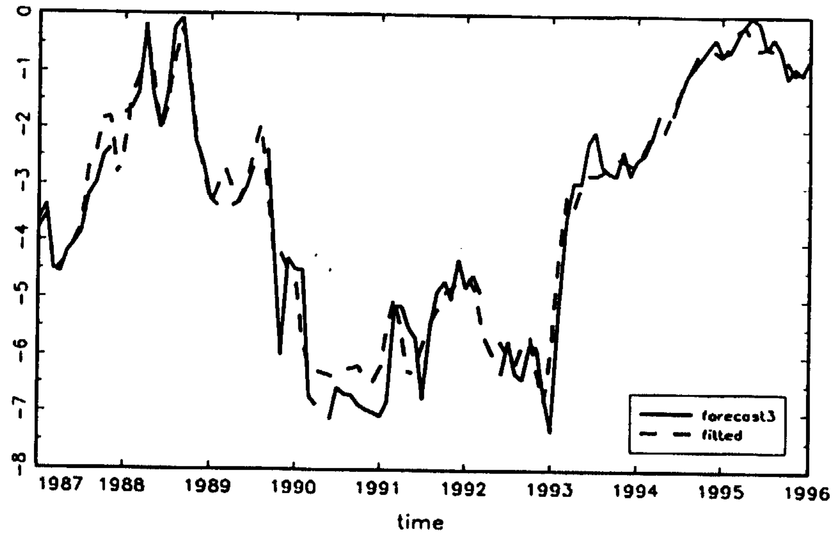


Figure 4.C: Euro 3 months, Forecast 6 Months
US-UK; AR=3

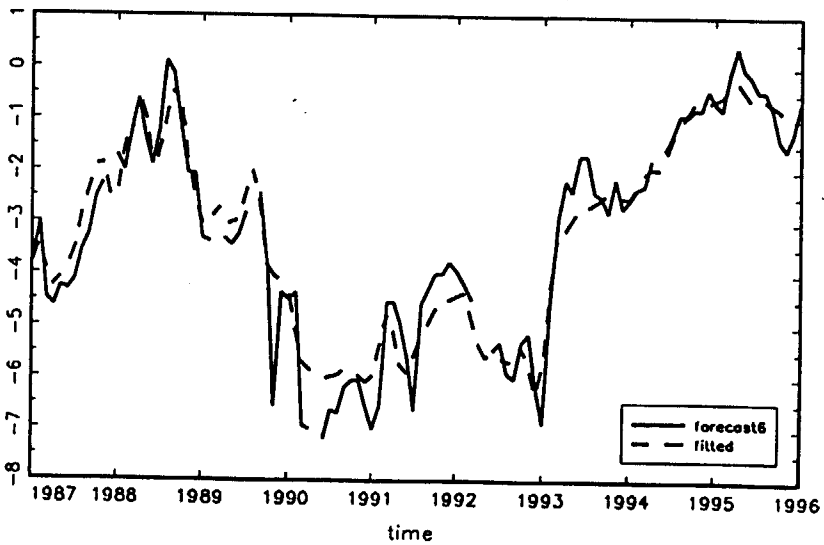


Figure 4.D: Euro 3 months, Forecast 12 Months
US-UK; AR=3

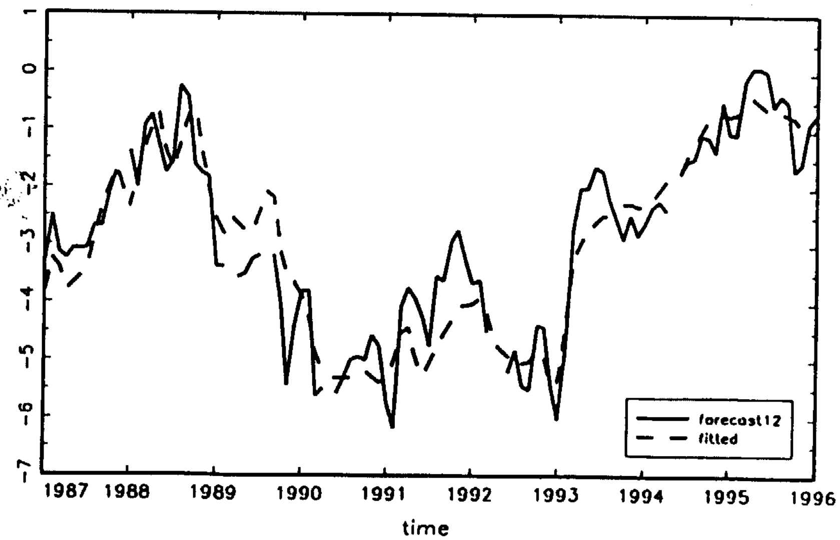


Figure 5: Delayed Overshooting

$\varepsilon_0=1;\ \lambda=0.98;\ k=0.2;\ i^*=0.05$

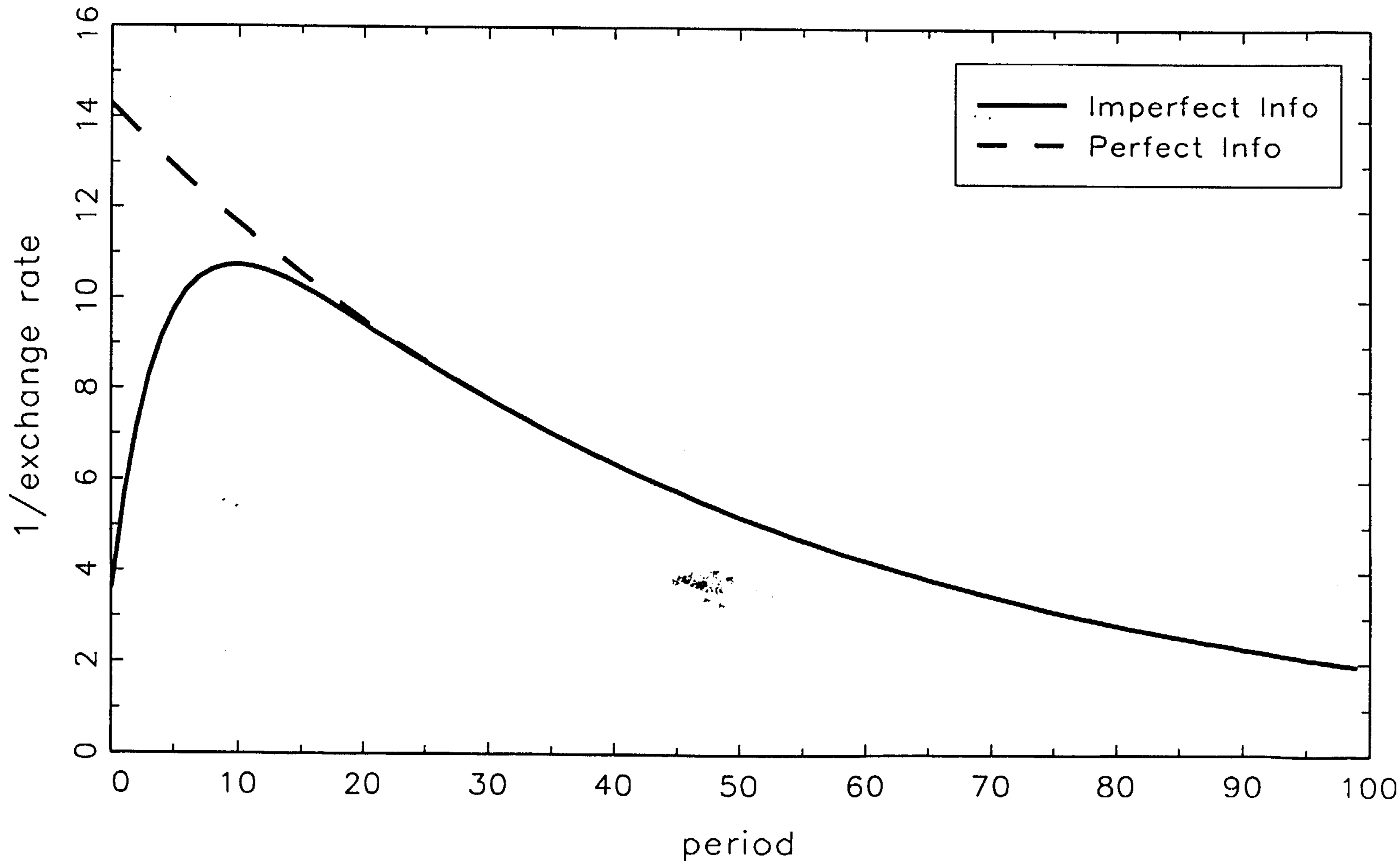


Figure 6: Market Belief

$\varepsilon_0=1$; $\lambda=0.98$; $k=0.2$; $i^*=0.05$

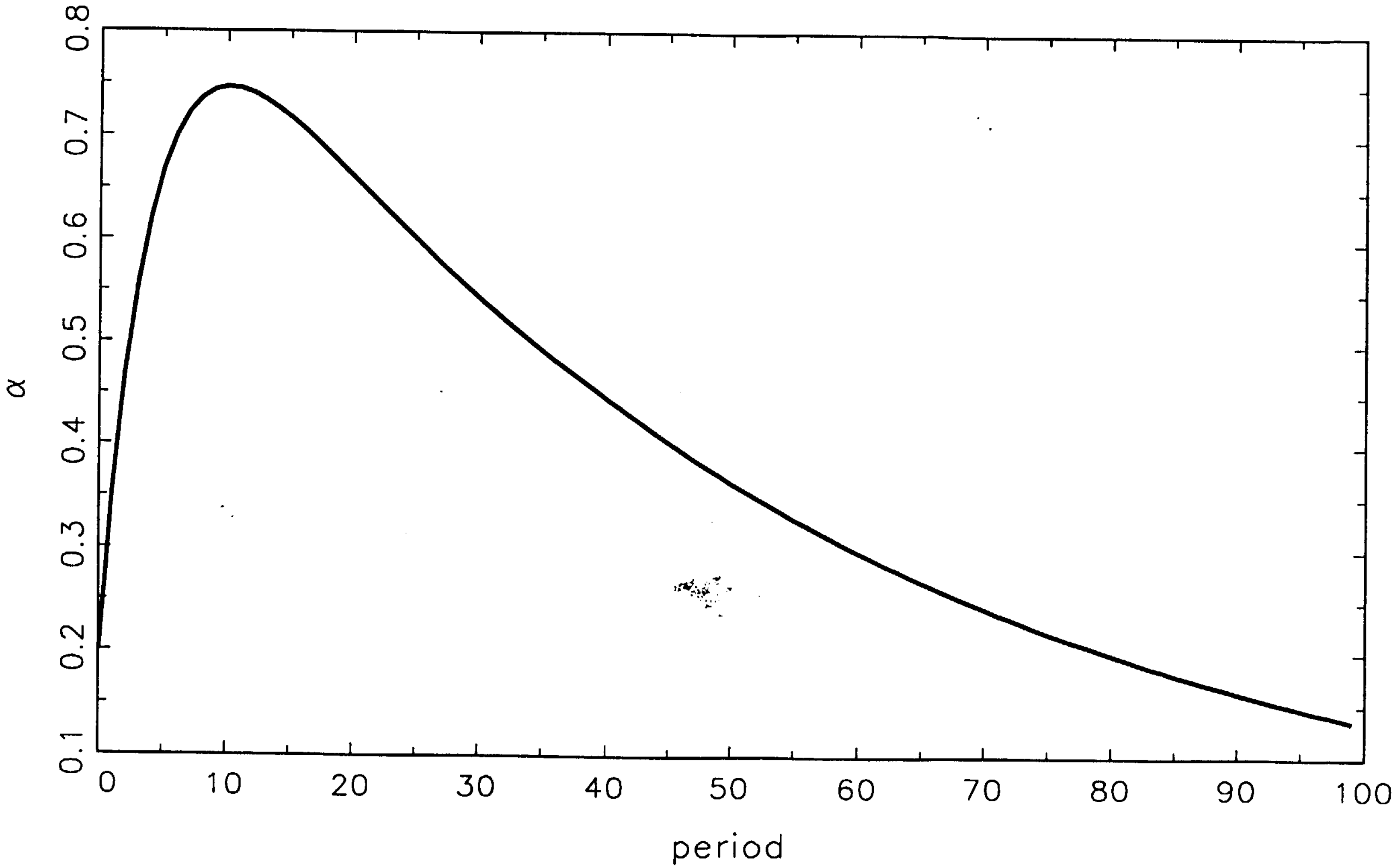


Figure 7.a: Delayed Overshooting Region

$\tau=1;\ \varepsilon_0=1;\ i^*=0.05$

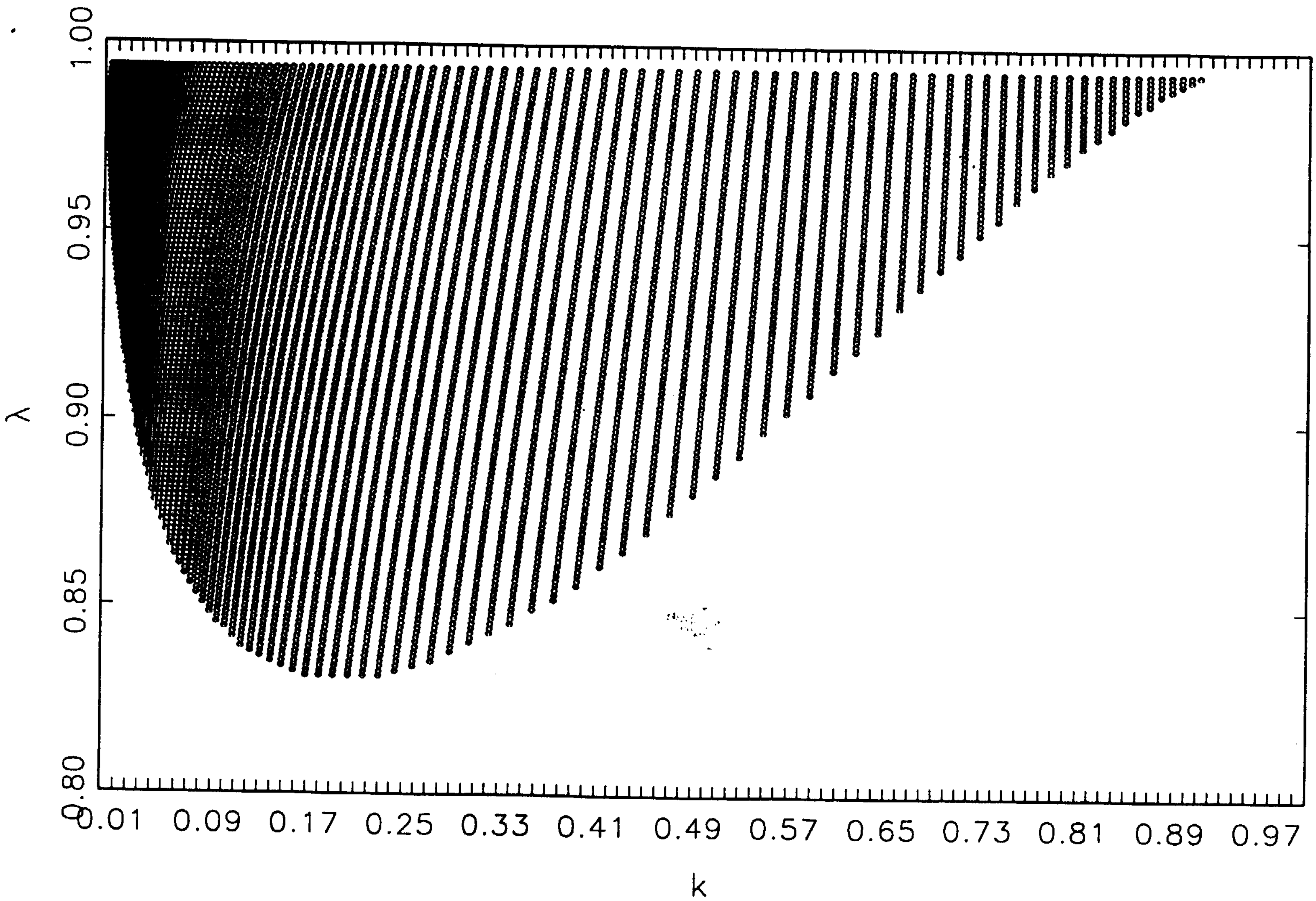
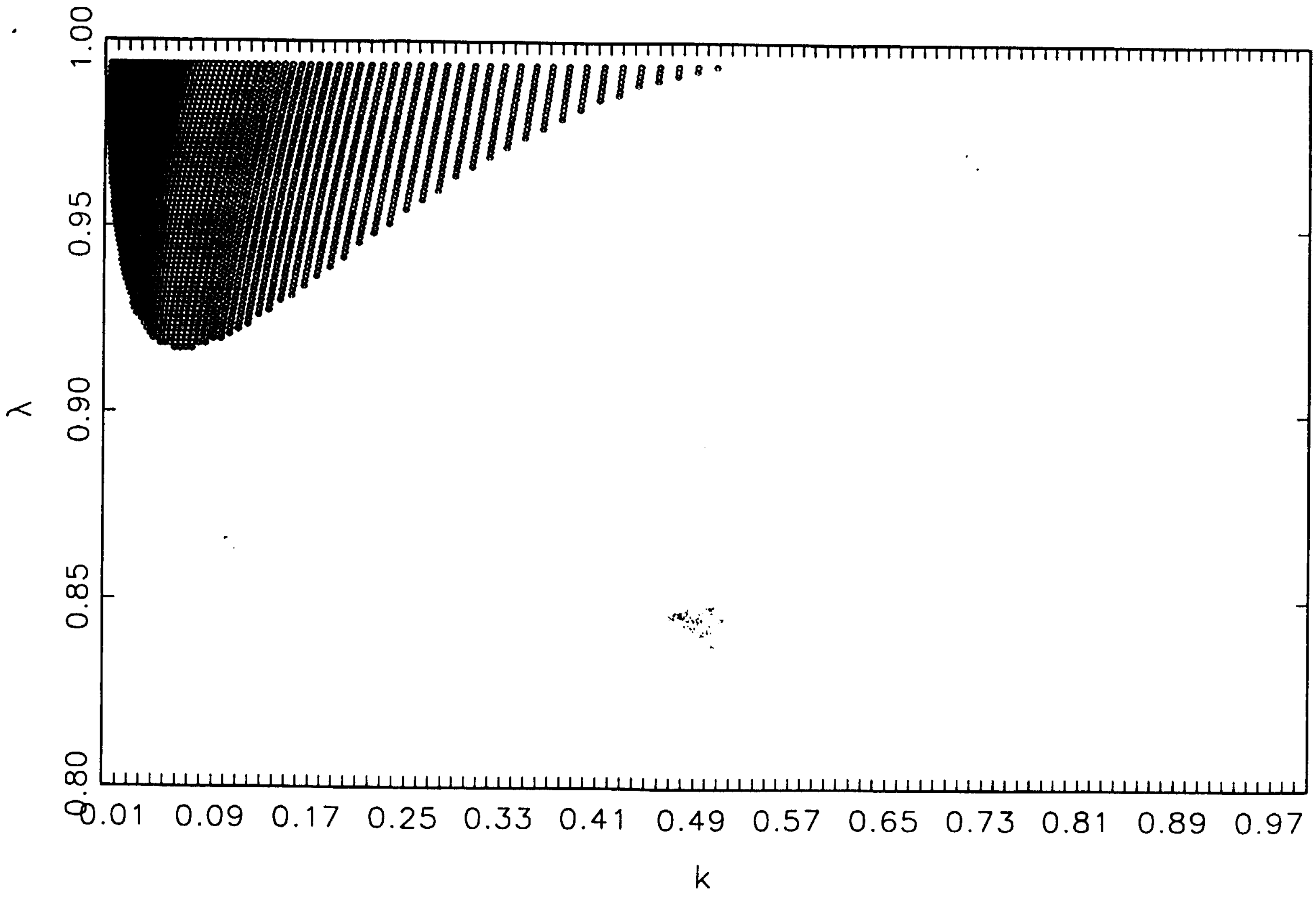


Figure 7.b: Delayed Overshooting Region

$\tau=5$; $\varepsilon_0=1$; $i^*=0.05$



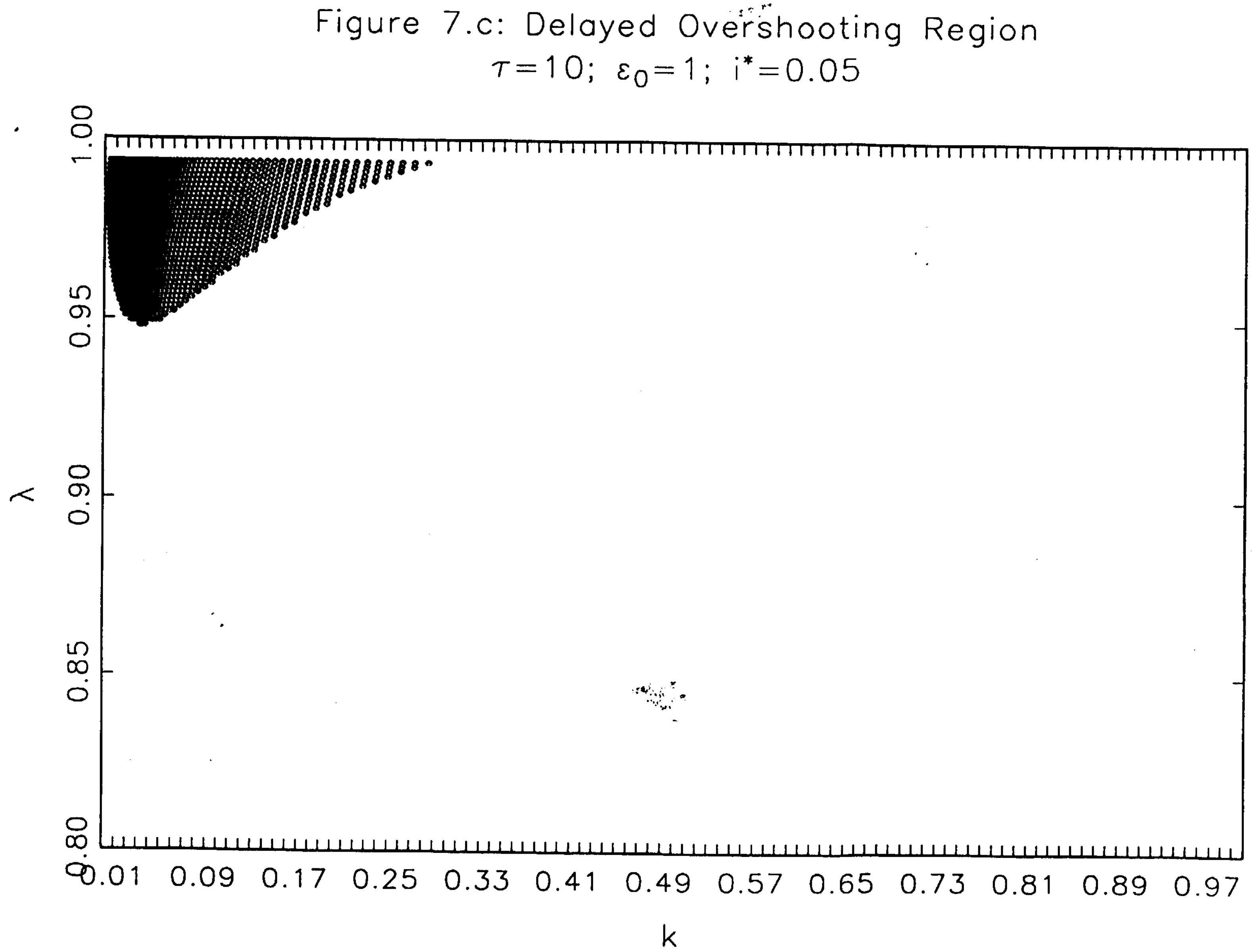


Figure 7.c: Delayed Overshooting Region
τ=10; ε0=1; i*=0.05
1.00
0.95
0.90
0.85
0.80
λ
0.01
0.09
0.17
0.25
0.33
0.41
0.49
0.57
0.65
0.73
0.81
0.89
0.97
k

Figure 8.a: Unconditional Delayed Overshooting Region

$\tau=1$; $\varepsilon_0=1$; $i^*=0.05$; $\eta=0.5$

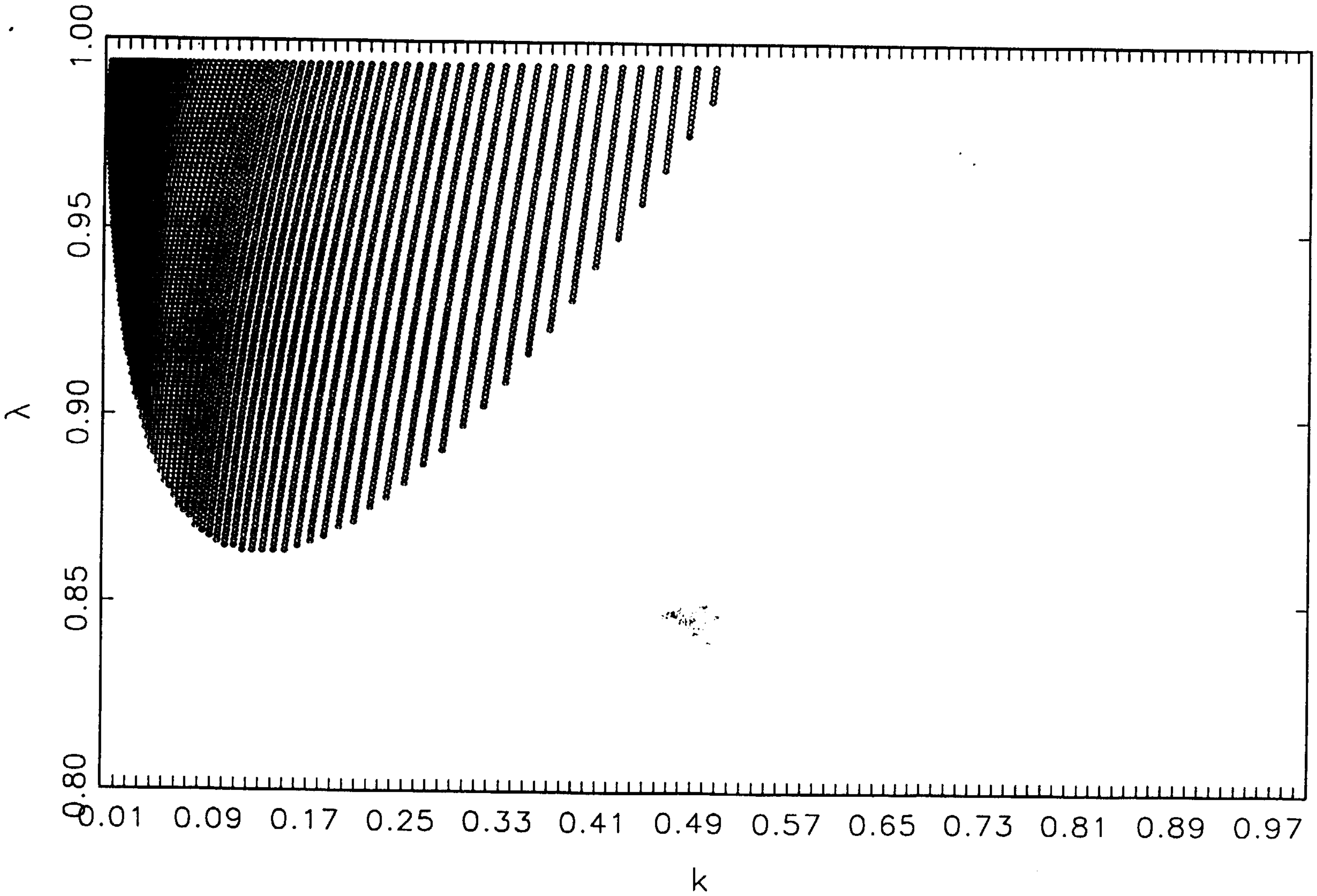


Figure 8.b: Unconditional Delayed Overshooting Region
$\tau=5$; $\varepsilon_0=1$; $i^*=0.05$; $\eta=0.5$

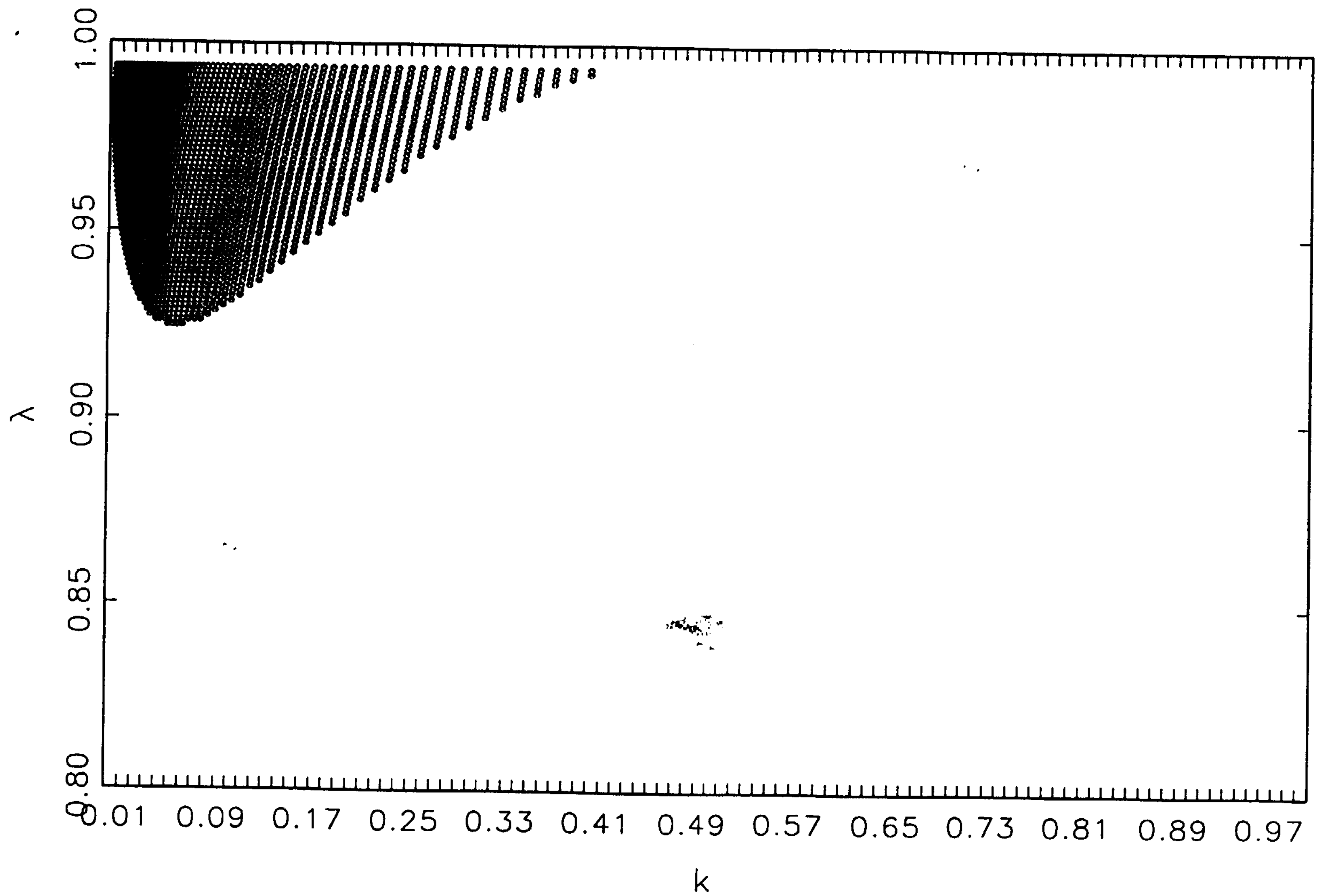


Figure 8.c: Unconditional Delayed Overshooting Region

$\tau=10$; $\varepsilon_0=1$; $i^*=0.05$; $\eta=0.5$

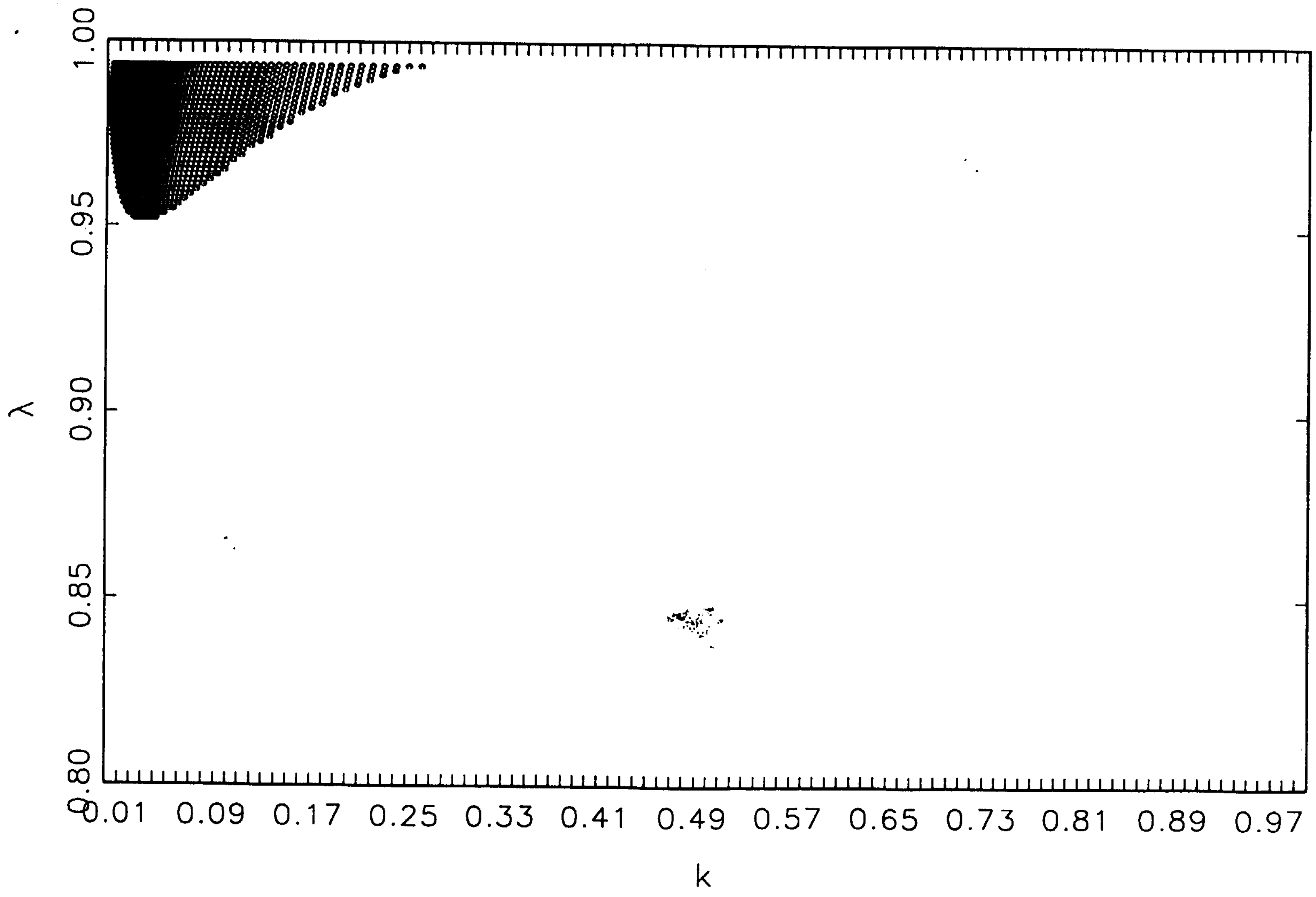


Figure 9: Monte Carlo Distribution of b_1
10,000 simulations

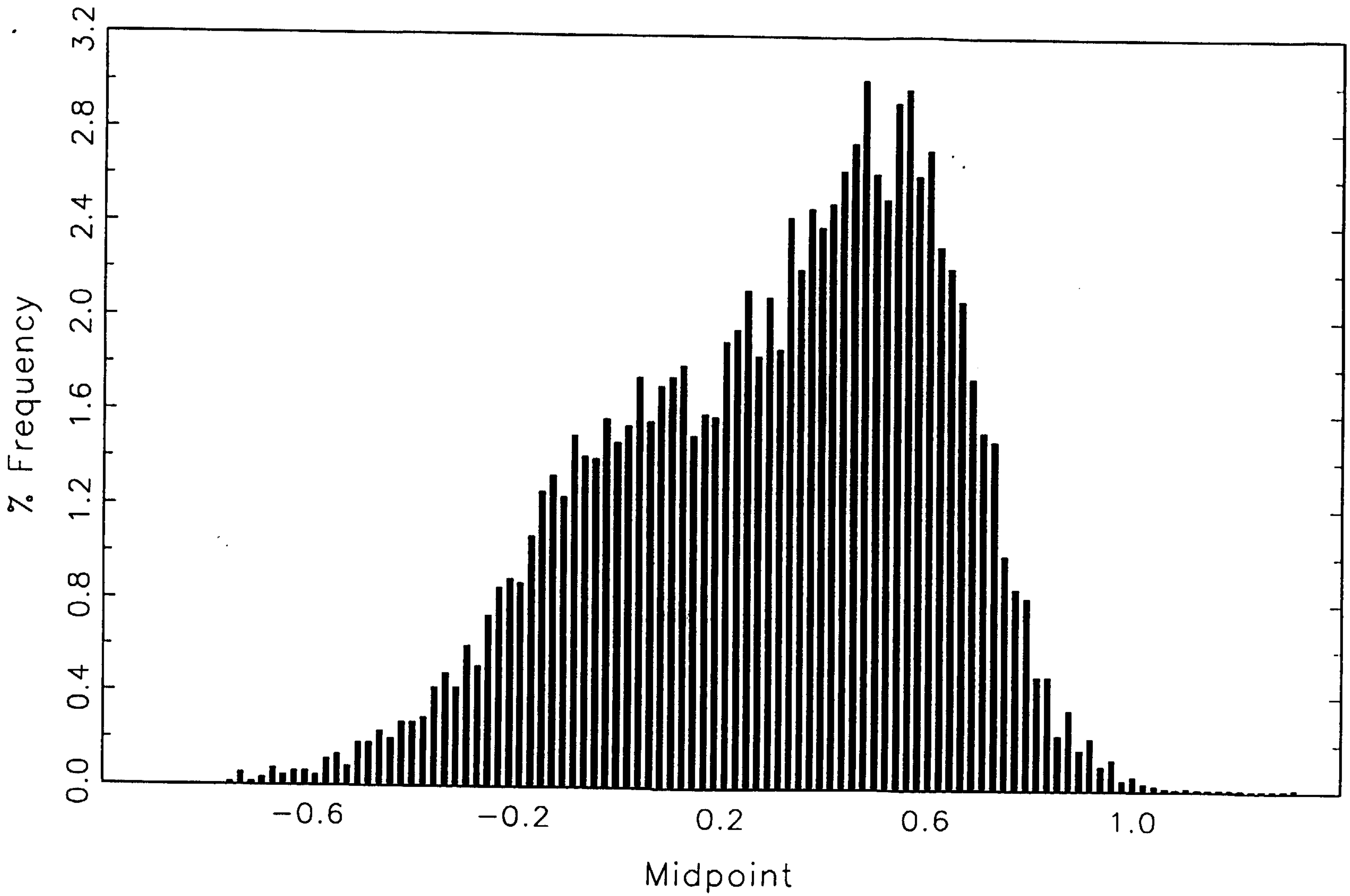


Figure 10: Monte Carlo Simulation
b_1 vs $se(b_1)$

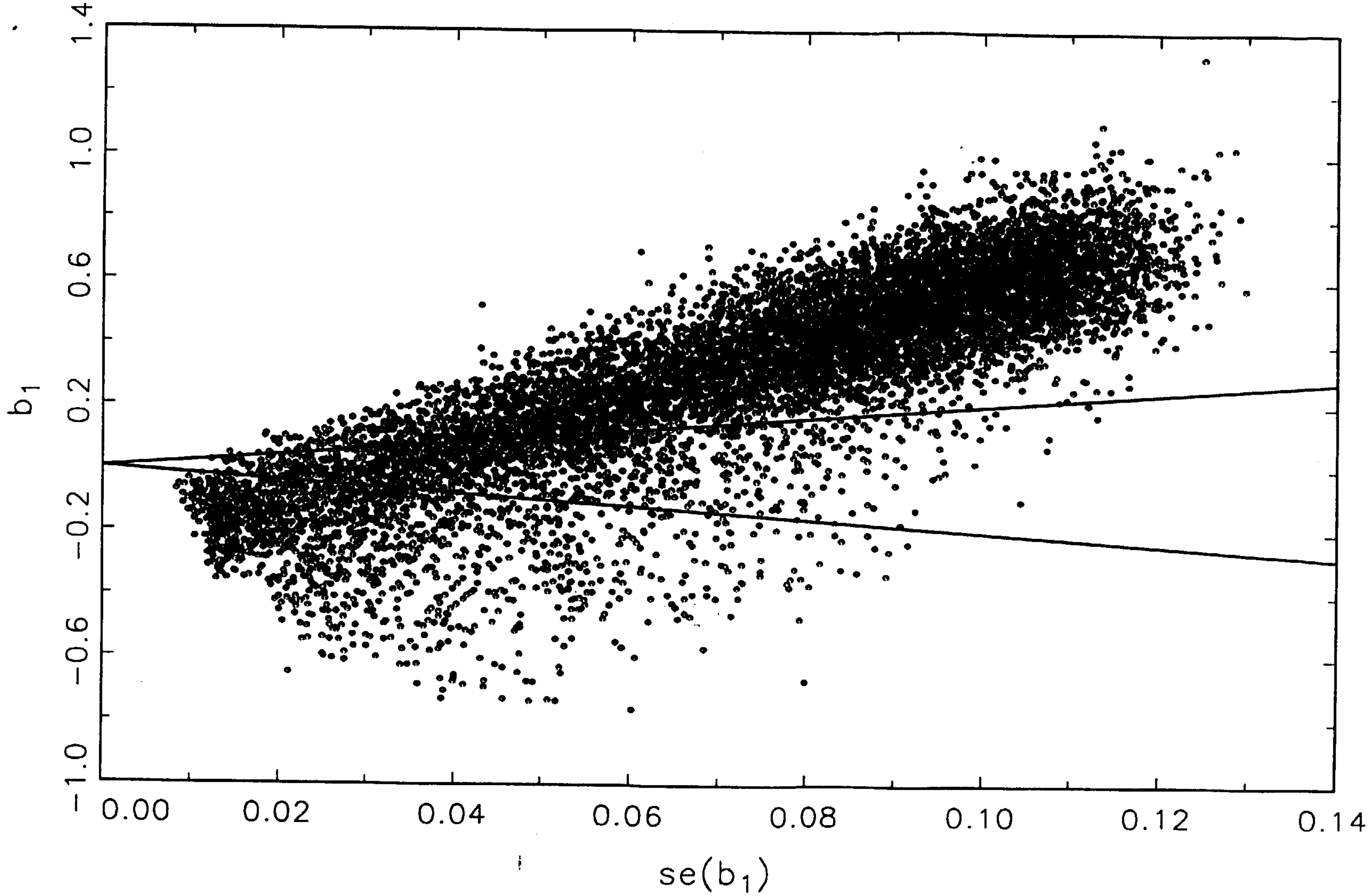


Figure 11.A: Interest Rate Differentials
UK-US

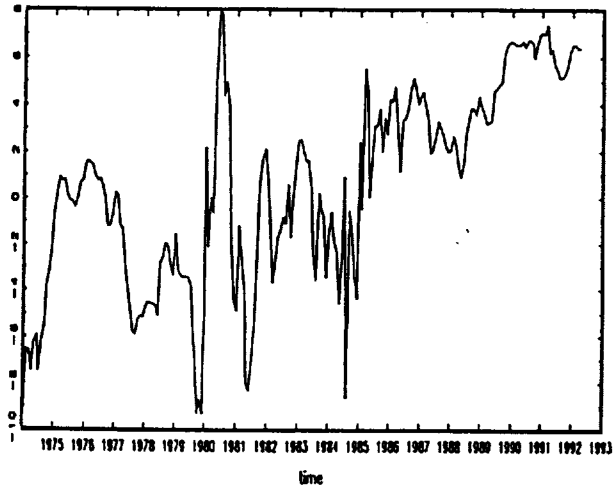


Figure 11.B: Interest Rate Differentials
FR-US

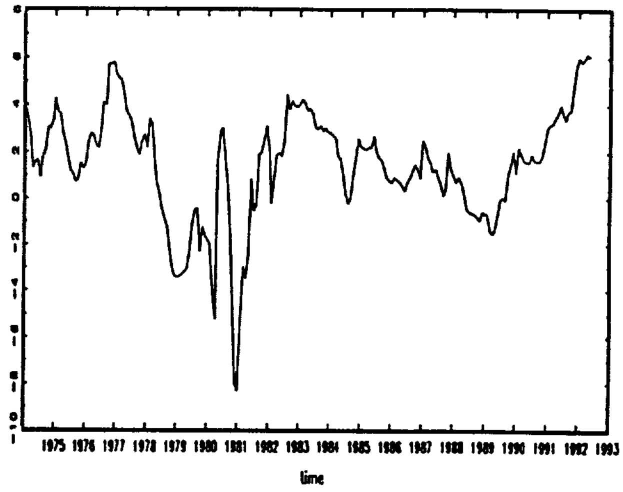


Figure 11.C: Interest Rate Differentials
GE-US

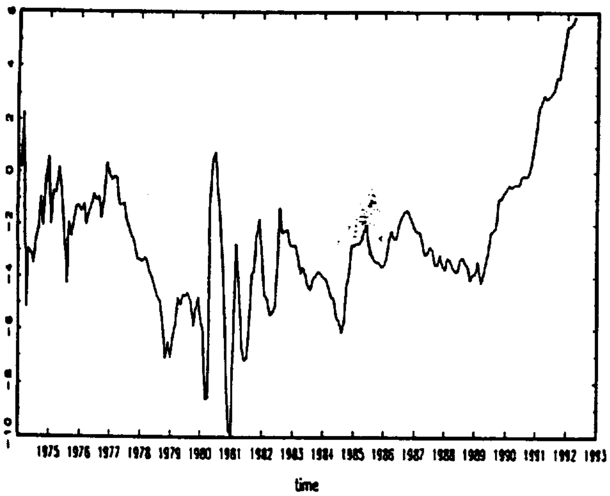


Figure 11.D: Interest Rate Differentials
IT-US

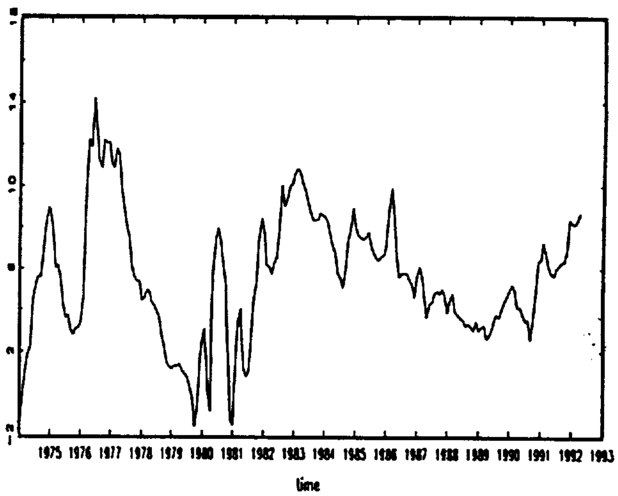


Figure 11.E: Interest Rate Differentials
CA-US

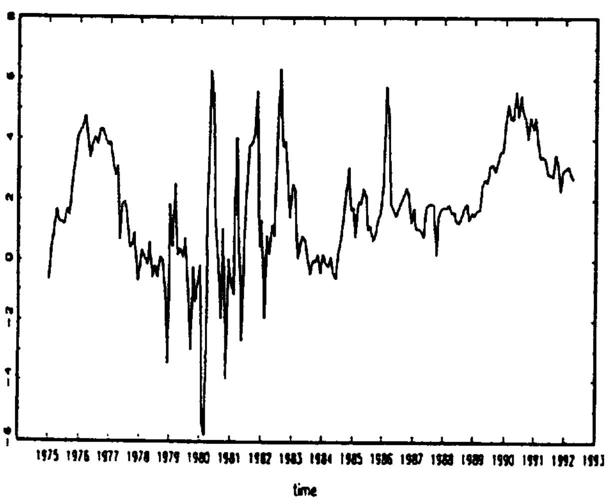


Figure 11.F: Interest Rate Differentials
JA-US

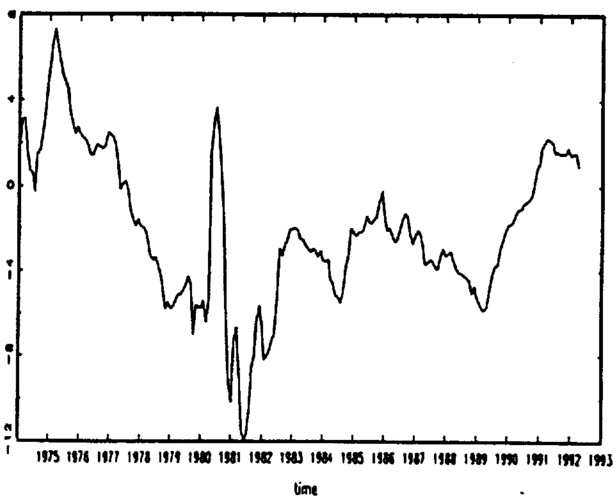


Figure 12.A: Interest Rate Differentials
UK-US

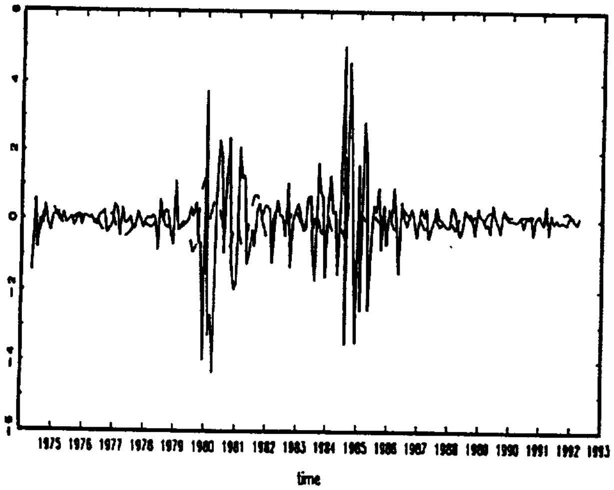


Figure 12.B: Interest Rate Differentials
FR-US

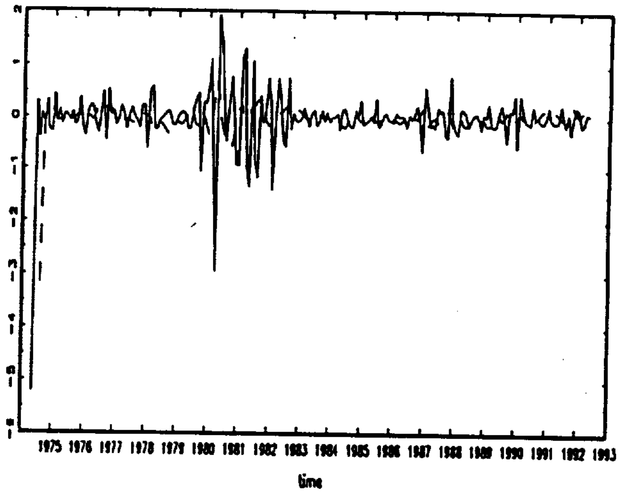


Figure 12.C: Interest Rate Differentials
GE-US

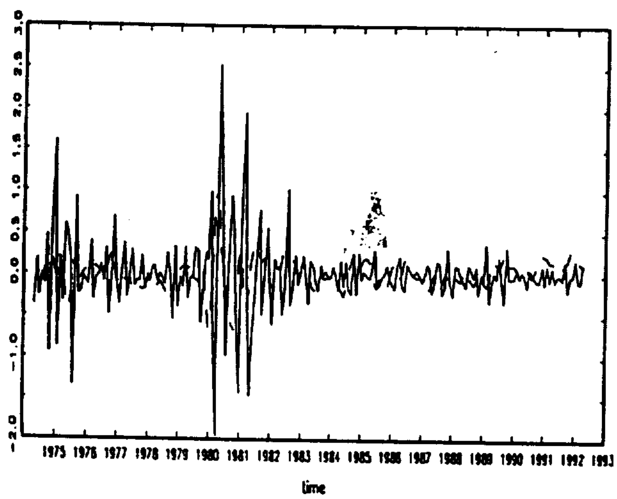


Figure 12.D: Interest Rate Differentials
IT-US

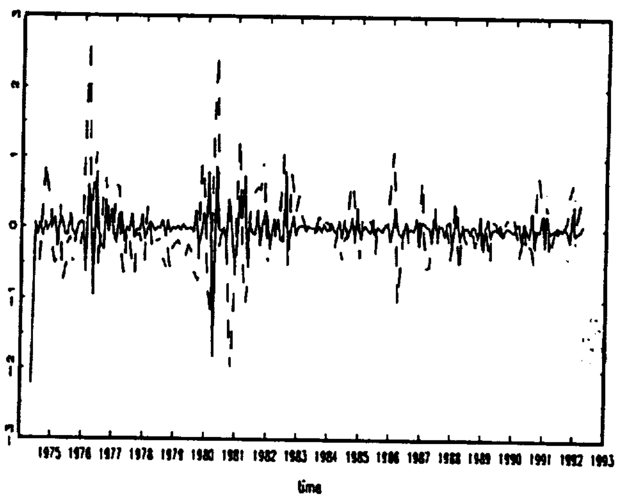


Figure 12.E: Interest Rate Differentials
CA-US

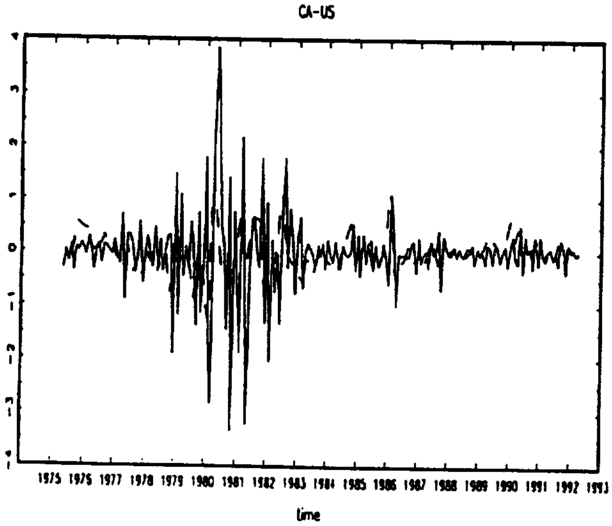


Figure 12.F: Interest Rate Differentials
JA-US