The Rationality and Heterogeneity of Survey Forecasts of the Yen-Dollar Exchange Rate: A Reexamination

by Richard Cohen
College of Business and Public Policy
University of Alaska Anchorage,

Carl Bonham
Department of Economics
University of Hawaii at Manoa, and

Shigeyuki Abe
Center for Contemporary Asian Studies
Doshisha University

Working Paper No. 06-11
July 28, 2006

Abstract This paper examines the rationality and diversity of industry-level forecasts of the yen-dollar exchange rate collected by the Japan Center for International Finance. In several ways we update and extend the seminal work by Ito (1990). We compare three specifications for testing rationality: the "conventional" bivariate regression, the univariate regression of a forecast error on a constant and other information set variables, and an error correction model (ECM). We find that the bivariate specification, while producing consistent estimates, suffers from two defects: first, the conventional restrictions are sufficient but not necessary for unbiasedness; second, the test has low power. However, before we can apply the univariate specification, we must conduct pretests for the stationarity of the forecast error. We find a unit root in the six-month horizon forecast error for all groups, thereby rejecting unbiasedness and weak efficiency at the pretest stage. For the other two horizons, we find much evidence in favor of unbiasedness but not weak efficiency. Our ECM rejects unbiasedness for all forecasters at all horizons. We conjecture that these results, too, occur because the restrictions test sufficiency, not necessity. In our systems estimation and micro-homogeneity testing, we use an innovative GMM technique (Bonham and Cohen (2001)) that allows for forecaster cross-correlation due to the existence of common shocks and/or herd effects. Tests of micro-homogeneity uniformly reject the hypothesis that forecasters across the four industries exhibit similar rationality characteristics.

Keywords: Rational Expectations, Heterogeneity, Exchange Rate, Survey Forecast.

Adress: Richard Cohen: College of Business and Public Policy · University of Alaska Anchorage · 3211 Providence Drive · Anchorage, AK 99508. Phone: (907) 786-1949. Fax: (907) 786-4115. E-mail: afrc2@cbpp.uaa.alaska.edu
Carl Bonham E-mail: bonham@hawaii.edu
Shigeyuki Abe E-mail: sabe@mail.doshisha.ac.jp

1 Overview

This paper examines the rationality of industry-level survey forecasts of the yen-dollar exchange rate collected by the Japan Center for International Finance (JCIF). Tests of rationality take on additional significance when performed on asset market prices, since rational expectations is a necessary condition for market efficiency. In the foreign exchange market, tests of forward rate unbiasedness simultaneously test a zero risk premium in the exchange rate; hence this joint hypothesis is also called the risk-neutral efficient market hypothesis (RNEMH). The practical significance of such a hypothesis is that, if the forward rate is indeed an unbiased predictor of the future spot rate, then exchange risk can be costlessly hedged in the forward market. However, the RNEMH has been rejected nearly universally. Since the risk premium is unobservable, insight into the reason for the rejection of the RNEMH can be gained by separately testing for rationality using survey data on expectations. Because forecasters cannot be assumed to have identical information sets, we must use individual survey forecasts to avoid the aggregation bias inherent in the use of mean or median forecasts.

We use data from the same source as Ito (1990), the seminal study recognizing the importance of using individual data to test rationality hypotheses about the exchange rate. To achieve stationarity of the realizations and forecasts (which each have a unit root), Ito (1990) followed the conventional specification at the time of subtracting the current realization from each. These variables are then referred to as being in "return" form. To test unbiasedness he regressed the future rate of depreciation on the forecasted return and tested the joint restrictions that the intercept equalled zero and the slope coefficient equalled one. At the industry level he found approximately twice as many rejections (at the 1% level) at the longest horizon (six months) than at the two shorter horizons (one and three months).

We extend Ito's analysis in two principal respects: the specification of unbiasedness tests and inference in tests for micro-homogeneity of forecasters. One problem with the returns specification is that, since there is much more variation in the return on the realization than

in the forecast, there is a tendency to under-reject the part of the joint hypothesis that the coefficient on the forecast equals one. This is precisely what we would expect in tests of variables which are near random walks.

Second, and more fundamentally, Ito's (1990) bivariate (joint) regression test of unbiasedness is actually a test of sufficiency, not necessity as well as sufficiency. Following Holden and Peel (1990), the necessary and sufficient condition for unbiasedness is a mean zero forecast error. This is tested in a univariate regression by imposing a coefficient of unity on the forecast and testing the restriction that the intercept equals zero. This critique applies whether or not the forecast and realization are integrated in levels. However, when the realization and forecast are integrated in levels, we must conduct a pretest to determine whether the forecast error is stationary. If the forecast and realization are both integrated and cointegrated, then a necessary and sufficient condition for unbiasedness is that intercept and slope in the cointegrating regression (using levels of the realization and forecast) are zero and one, respectively. We test this hypothesis using Liu and Maddala's (1992) method of imposing the (0,1) vector, then testing the "restricted" cointegrating residual for stationarity.[1,2]

Third, we use the result from Engle and Granger (1987) that cointegrated variables have an error correction representation. First, we employ the specification and unbiasedness restrictions originally proposed by Hakkio and Rush (1989). However, the unbiasedness tests using the ECM specification produce more rejections over industry groups and horizons than the univariate or bivariate specifications. We conjecture that one possible explanation for this apparent anomaly is that, similar to the joint restrictions in the bivariate test, the ECM restrictions test sufficient conditions for unbiasedness, while the univariate restriction only tests a necessary and sufficient condition. Thus, the ECM has a tendency to over-reject. We then respecify the ECM, so that only the necessary and sufficient conditions are tested. We

[1] If in addition the residuals from the cointegrating regression are white noise, this supports a type of weak efficiency.

[2] Pretesting the forecast error for stationarity is a common practice in testing the RNMEH, but the only study we know of that applies this practice to survey forecasts of exchange rates is Osterberg (2000), and he does not test for a zero intercept in the cointegrating regression.

compare our results to those obtained using the sufficient conditions represented by the joint restrictions as well as the necessary and sufficient condition represented by the univariate restriction.

The second direction in which we extend Ito's (1990) analysis has to do with testing for differences among forecasters' ability to produce rational predictions.[3] We recognize, as does Ito, that differences among forecasters over time indicate that at least some individuals form biased forecasts. (The converse does not necessarily hold, since a failure to reject micro-homogeneity could conceivably be due to the same degree of irrationality of each individual in the panel.) Ito's heterogeneity test is a single equation test of deviations of individual forecasts from the mean forecast, where the latter may or may not be unbiased. In contrast, we test for differences in individual forecast performance using a micro-homogeneity test, i.e., imposing equal coefficients across the system of individual univariate rationality equations.

In our tests for micro-homogeneity, we expect cross-forecaster error correlation due to the possibility of common macro shocks and/or herd effects in expectations. To this end, we incorporate two innovations not previously used by investigators studying survey data on exchange rate expectations. First, in our micro-homogeneity tests we use a GMM system with a variance-covariance matrix that allows for cross-sectional as well as moving average and heteroscedastic errors. Here we follow the widely-used practice of modeling the individual regression residuals as an MA process of order h-1, where h is the number of periods in the forecast horizon. However, no other researchers have actually tested whether an MA process of this length is required to model the cross-sectional behavior of rational forecast errors. Thus, second, to investigate the nature of the actual MA processes, we use Pesaran's (2004) CD test to examine the statistical significance of the cross-sectional dependence of forecast errors, both contemporaneous and lagged.

[3]Market microstructure theories assume that there is a minimum amount of forecaster (as well as cross-sectional forecast) diversity. Also, theories of exchange rate determination that depend upon the interaction between chartists (or noise traders) and fundamentalists by definition require a certain structure of forecaster heterogeneity.

The organization of the rest of the paper is as follows. In section 2 we review some fundamental issues in testing rationality in the foreign exchange market. In sections 3 and 4 we conduct various rationality tests on the JCIF data. Section 5 contains our micro-homogeneity tests. Section 6 summarizes and discusses areas for future research.

2 Background: testing rationality in the foreign exchange market

The Rational Expectations Hypothesis (REH) assumes that economic agents know the true data generating process (DGP) for the forecast variable. This implies that the market's subjective probability distribution of the variable is identical to the objective probability distribution, conditional on a given information set, Φ_t. Equating first moments of the market, $E_m(s_{t+h}|\Phi_t)$, and objective, $E(s_{t+h}|\Phi_t)$, distributions,

$$E_m(s_{t+h}|\Phi_t) = E(s_{t+h}|\Phi_t), \tag{1}$$

where the right-hand-side can be shortened to $E_t(s_{t+h})$.

It follows that the REH implies that forecast errors have both unconditional and conditional means equal to zero. A forecast is unbiased if its forecast error has an unconditional mean of zero. A forecast is efficient if its error has a conditional mean of zero. The condition that forecast errors be serially uncorrelated is a subset of the efficiency condition where the conditioning information set consists of past values of the realization and current as well as past values of the forecast.[4]

In this paper we focus on testing whether forecasters can form rational expectations of future depreciation. If not, then at least part of the explanation for the failure of the RNEMH is due to the failure of the REH. There are two related interest parity conditions. Covered interest parity, an arbitrage condition, holds if $f_{t,h} - s_t = i_t - i_t^*$, i.e., the forward premium is equal to the interest differential between domestic and foreign risk free assets. Uncovered interest parity holds if $s_{t+h} - s_t^e = i_t - i_t^*$. Because uncovered interest parity assumes both

[4]It is important to note that the result from one type of rationality test does not have implications for the results from any other types of rationality tests. In this paper we test for unbiasedness and weak efficiency, leaving the more stringent tests of efficiency with respect to publicly available information for future analysis.

unbiased expectations and risk neutrality, some authors view it as equivalent to the RNEMH (see Phillips and Maynard (2001)).

The ability to decompose deviations from UIP into time-varying risk premium and systematic forecast error components also has implications for policymakers. Consider first the possibility of a violation of the risk neutrality hypothesis. According to the portfolio balance model, if a statistically significant time-varying risk premium component is found, this means that $i_t - i_t^*$ is time-varying, which in turn implies that foreign and domestic bonds are not perfect substitutes; changes in relative quantities (which are reflected in changes in current account balances) will affect the interest rate differential. In this way, sterilized official intervention can have significant effects on exchange rates. Second, consider the the possibility of a violation of the REH. If a statistically significant expectational error of the destabilizing (e.g., "bandwagon") type is found, and policymakers are more rational than speculators, a policy of "leaning against the wind" could have a stabilizing effect on exchange rate movements. (See Cavaglia et al. 1994.) More generally, monetary models of the exchange rate (in which the UIP condition is embedded), which assume model-consistent (i.e., rational) expectations with risk-neutrality, generally have not performed well empirically, especially in out-of-sample forecasting. (See, e.g., Bryant 1995.) One would like to be able to attribute the model failure to some combination of a failure of the structural assumptions (including risk neutrality) or a failure of the expectational assumption.

2.1 Why test rational expectations with *disaggregated* survey forecast data?

Beginning with Frankel and Froot (1987) and Froot and Frankel (1989), much of the literature examining exchange rate rationality in general, and the decomposition of deviations from the RNEMH in particular, has employed the representative agent assumption to justify using the mean or median survey forecast as a proxy for the market's expectation. In both studies, Frankel and Froot found significant evidence of irrationality. Subsequent research has found mixed results. Liu and Maddala (1992, p. 366) articulate the mainstream justification for using aggregated forecasts in tests of the REH. "Although ...data on indi-

viduals are important to throw light on how expectations are formed at the individual level, to analyze issues relating to market efficiency, one has to resort to aggregates." In fact, Muth's (1961)[p. 316] original definition of rational expectations seemed to allow for the possibility that rationality could be applied to an aggregate (e.g., mean or median) forecast. "...[E]xpectations of firms (or, more generally, the subjective probability distribution of outcomes) tend to be distributed, *for the same information set*, about the predictions of the theory (or the 'objective' probability distribution of outcomes.)" (Emphasis added.)

However, if individual forecasters have different information sets, Muth's definition does not apply. To take the simplest example, the (current) mean forecast is not in any forecaster's information set, since all individuals' forecasts must be made before a mean can be calculated. Thus, current mean forecasts contain private information (see MacDonald, 1992) and therefore cannot be tested for rationality.[5]

Using the mean forecast may also result in inconsistent parameter estimates. Figlewski and Wachtel (1983) were the first to show that, in the traditional bivariate unbiasedness equation, the presence of private information variables in the mean forecast error sets up a correlation with the mean forecast. This inconsistency occurs even if all individual forecasts are rational. In addition, Keane and Runkle (1990) pointed out that, when some forecasters are irrational, using the mean forecast may lead to false acceptance of the unbiasedness hypothesis, in the unlikely event that offsetting individual biases allow parameters to be consistently estimated. See also Bonham and Cohen (2001), who argue that, in the case of cointegrated targets and predictions, inconsistency of estimates in rationality tests using the mean forecast can be avoided if corresponding coefficients in the individual rationality tests

[5] A large theoretical literature relaxes Muth's assumption that all information relevant for forming a rational forecast is publicly available. Instead, this literature examines how heterogeneous individual expectations are mapped into an aggregate market expectation, and whether the latter leads to market efficiency. (See, e.g., Figlewski 1978, 1982, 1984; Kirman 1992; Haltiwanger and Waldman 1989.) Our paper focuses on individual rationality but allows for the possibility of synergism by incorporating not only heteroscedasticity and autocorrelation consistent standard errors in individual rationality tests but also cross-forecaster correlation in tests of micro-homogeneity. The extreme informational requirement of the REH lead Pesaran and Weale (2006 (forthcoming)) to propose a weaker form of the REH that is based on the (weighted) average expectation using only publicly available (i.e., common) information.

pass a test for micro-homogeneity.[6] Nevertheless, until the 1990s, few researchers tested for the rationality of individual forecasts, even when those data were available.

2.2 Rational reasons for the failure of the Rational Expectations Hypothesis using disaggregated data

Other than a failure to process available information efficiently, there are numerous explanations for a rejection of the REH. One set of reasons relates to measurement error in the individual forecast. Researchers have long recognized that forecasts of economic variables collected from public opinion surveys should be less informed than those sampled from industry participants. However, industry participants, while relatively knowledgeable, may not be properly motivated to devote the time and resources necessary to elicit their best responses. The opposite is also possible.[7] Having devoted substantial resources to produce a forecast of the price of a widely traded asset, such as foreign exchange, forecasters may be reluctant to reveal their true forecast before they have had a chance to trade for their own account.[8,9]

Second, some forecasters may not have the symmetric quadratic loss function embodied in typical measures of forecast accuracy, e.g., minimum mean squared error. (See Zellner 1986; Stockman 1987; Batchelor and Peel 1998.) In this case, the optimal forecast may

[6] The extent to which private information influences forecasts is more controversial in the foreign exchange market than in the equity or bond markets. While Chionis and MacDonald (1997) maintain that there is little or no private information in the foreign exchange market, Lyons (2002) argues that order flow explains much of the variation in prices. To the extent that one agrees with the market microstructure emphasis on the importance of the private information embodied in dealer order flow, the Figlewski-Wachtel critique remains valid in the returns regression.

[7] Elliott and Ito (1999) show that, although a random walk forecast frequently outperforms the JCIF survey forecasts using an MSE criterion, survey forecasts generally outperform the random walk, based on an excess profits criterion. This supports the contention that JCIF forecasters are properly motivated to produce their best forecasts.

[8] To mitigate the confidentiality problem in this case, the survey typically withholds individual forecasts until the realization is known, or (as with the JCIF) masks the individual forecast by only reporting some aggregate forecast (at the industry and total level) to the public.

[9] Furthermore, reported individual forecasts may not represent the mean of the forecaster's subjective probability distribution if that distribution is skewed and the forecaster reports another measure of central tendency, e.g., the median.

not be the MSE. In one scenario, related to the incentive aspect of the measurement error problem, forecasters may have strategic incentives involving product differentiation.[10]

In addition to strategic behavior, another scenario in which forecasters may deviate from the symmetric quadratic loss function is simply to maximize trading profits. This requires predicting the direction of change, regardless of MSE.[11]

Third, despite their best efforts, forecasters may find it difficult to distinguish between a temporary and permanent shift in the DGP. This difficulty underlies at least three theories of rational forecast errors: the peso problem, learning about past regime changes, and bubbles.

Below we conduct tests for structural change in estimated unbiasedness coefficients. When unbiasedness cannot be rejected, the structural change test may show certain subperiods in which unbiasedness did not hold. In the obverse case, when unbiasedness can be rejected, the structural change test may show certain subperiods in which unbiasedness cannot be rejected. Either situation would lend some support to the theories attributing bias to the difficulty of distinguishing temporary from permanent shifts.

3 Description of data

Every two weeks, the JCIF in Tokyo conducts telephone surveys of yen/dollar exchange rate expectations from 44 firms. The forecasts are for the future spot rate at horizons of one month, three months, and six months. Our data cover the period May 1985 to March 1996. This data set has very few missing observations, making it close to a true panel. For reporting purposes, the JCIF currently groups individual firms into four industry categories: 1) banks and brokers, 2) insurance and trading companies, 3) exporters, and 4) life insurance companies and importers. On the day after the survey, the JCIF announces

[10] Laster et al. (1999) called this practice "rational bias." Prominent references in this growing literature include Lamont (2002), Ehrbeck and Waldmann (1996), and Batchelor and Dua (1990a,b, 1992). Because we have access only to forecasts at the industry average level, we cannot test the strategic incentive hypotheses.

[11] See Elliott and Ito (1999), Boothe and Glassman (1987), LeBaron (2000), Leitch and Tanner (1991), Lai (1990), Goldberg and Frydman (1996), and Pilbeam (1995). This type of loss function may appear to be relevant only for relatively liquid assets such as foreign exchange, but not for macroeconomic flows. However, the directional goal is also used in models to predict business cycle turning points. Also, trends in financial engineering may lead to the creation of derivative contracts in macroeconomic variables, eg., CPI futures.

overall and industry average forecasts. (For further details concerning the JCIF database, see the descriptions in Ito (1990, 1994), Bryant (1995), and Elliott and Ito (1999).)

Figure 1 shows that, over the sample period (one of flexible exchange rates and no capital controls), the yen appreciated dramatically relative to the dollar, from a spot rate of approximately 270 yen/dollar in May 1985 to approximately 90 yen/dollar in March 1996. The path of appreciation was not steady, however. In the first two years of the survey alone, the yen appreciated to about 140 per dollar. The initial rapid appreciation of the yen is generally attributed to the Plaza meeting in September 1985, in which the Group of Five countries decided to let the dollar depreciate, relative to the other currencies. At the Louvre meeting in February 1987, the Group of Seven agreed to stabilize exchange rates by establishing soft target zones. These meetings may well be interpreted as unanticipated regime changes, since, as we will see below, forecasters generally underestimated the rapid appreciation following the Plaza meeting, then overestimated the value of the yen following the Louvre meeting. Thus, forecasts during these periods may have been subject to peso and learning problems. The period of stabilization lasted until about 1990, when yen appreciation resumed and continued through the end of the sample period.

4 Empirical tests of rationality

Early studies of the unbiasedness aspect of rationality regressed the level of the realization on the level of the forecast, testing the joint hypothesis that the intercept equalled zero and the slope equalled one.[12] However, since many macroeconomic variables have unit roots, realization and forecast typically share a common stochastic trend a rational forecast will be integrated and cointegrated with the target series. (See Granger, 1991, pp. 69-70.) . According to the modern theory of regressions with integrated processes (see, *inter alia* Banerjee et al., 1993), conventional OLS estimation and inference produce a slope coefficient

[12]The efficiency aspect of rationality is sometimes tested by including additional variables in the forecaster's information set, with corresponding hypotheses of zero coefficients on these variables. See, e.g., Keane and Runkle (1990) for a more recent study using the level specification and Bonham and Cohen (1995) for a critique of Keane and Runkle's integration accounting.

that is biased toward one and, therefore, a test statistic that is biased toward accepting the null of unbiasedness. The second generation studies of unbiasedness addressed this inference problem by subtracting the current realization from the forecast as well as the future realization, transforming the levels regression into a "returns" regression. In this specification of stationary variables, unbiasedness was still tested using the same (0,1) joint hypothesis as in the levels regression. However, an implication of Engle and Granger (1987) is that the levels regression is now interpreted as a cointegrating regression, with conventional t-statistics following nonstandard distributions which depend on nuisance parameters. After establishing that the realization and forecast are integrated and cointegrated, we perform two types of rationality tests. The first is a "restricted cointegration" test due to Liu and Maddala (1992). This is a cointegration test imposing the (0,1) restriction on the levels regression.

It is significant that, if realization and forecast are cointegrated, Liu and Maddala's (1992) technique is equivalent to regressing a stationary forecast error on a constant and then testing whether the coefficient equals zero (to test unbiasedness) and/or whether the residuals are white noise (to test a type of weak efficiency). Pretests for unit roots in the realization, forecast and forecast error are required for at least three reasons. First, univariate tests of unbiasedness are invalid if the forecast error is not stationary. Second, following Holden and Peel (1990), we show below (in section 4.1.1) that nonrejection of the joint test in the bivariate regression is sufficient but not necessary for unbiasedness, since the joint test is also an implicit test of weak efficiency with respect to the lagged forecast error. A zero intercept in the (correctly specified) univariate test is a necessary as well as sufficient condition for unbiasedness. Third, the Engel-Granger (1987) representation theorem proves that a cointegrating regression such as the levels joint regression ((2) below) has an error correction form that includes both differenced variables and an error correction term in levels. Under the joint null, the error correction term is the forecast error. While the returns form of the bivariate regression, is not, strictly speaking, misspecified (since the regressor subtracts s_t, not s^e_{t-1}, from s^e_t), the ECM specification may produce a better fit to the data

and, therefore, a more powerful test of the unbiasedness restrictions. We conduct such tests using a form of the ECM due to Hakkio and Rush (1989).

4.1 Joint tests of unbiasedness and weak efficiency

4.1.1 The lack of necessity critique

Many, perhaps most, empirical tests of the "unbiasedness" of survey forecasts are conducted using the bivariate regression equation

$$s_{t+h} - s_t = \alpha_{i,h} + \beta_{i,h}(s^e_{i,t,h} - s_t) + \varepsilon_{i,t,h}. \tag{2}$$

It is typical for researchers to interpret their non-rejection of the joint null $(\alpha_{i,h}, \beta_{i,h}) = (0, 1)$ as a necessary condition for unbiasedness. However, Holden and Peel (1990) show that this result is a sufficient, though not a necessary, condition for unbiasedness. The intuition for the lack of necessity comes from interpreting the right-hand-side of the bivariate unbiasedness regression as a linear combination of two potentially unbiased forecasts: a constant equal to the unconditional mean forecast plus a variable forecast, i.e., $s_{t+h} - s_t = (1 - \beta_{i,h}) \times E(s^e_{i,t,h} - s_t) + \beta_{i,h}(s^e_{i,t,h} - s_t) + \varepsilon_{i,t,h}$. Then the intercept is $\alpha_{i,h} = (1 - \beta_{i,h}) \times E(s^e_{i,t,h} - s_t)$. The necessary and sufficient condition for unbiasedness is that the unconditional mean of the subjective expectation $E[s^e_{i,t,h} - s_t]$ equal the unconditional mean for the objective expectation $E[s_{t+h} - s_t]$. However, this equality can be satisfied without $\alpha_{i,h}$ being equal to zero, i.e., $\beta_{i,h} = 1$.

Figure 2 shows that an infinite number of $\alpha_{i,h}, \beta_{i,h}$ estimates are consistent with unbiasedness. The only constraint is that the regression line intersect the 45 degree ray from the origin where the sample mean of the forecast and target are equal. Note that, in the case of differenced variables, this can occur at the origin, so that $\alpha_{i,h} = 0$, but $\beta_{i,h}$ is unrestricted (see Figure 3). It is easy to see why unbiasedness holds: in Figures 2 and 3 the sum of all horizontal deviations from the 45 degree line to the regression line, i.e., forecast errors, equal zero. However, when $\alpha_{i,h} \neq 0$, and $\alpha_{i,h} \neq (1 - \beta_{i,h}) \times E(s^e_{i,t,h} - s_t)$, there is bias regard-

less of the value of $\beta_{i,h}$. See Figure 4, where the bias, $E(s_{t+h} - s^e_{i,t,h})$, implies systematic underforecasts.

To investigate the rationality implications of different values for $\alpha_{i,h}$ and $\beta_{i,h}$, we follow Clements and Hendry (1998) and rewrite the forecast error in the bivariate regression framework of (2) as

$$\eta_{i,t,h} = s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + (\beta_{i,h} - 1)(s^e_{i,t,h} - s_t) + \varepsilon_{i,t,h} \tag{3}$$

A special case of weak efficiency occurs when the forecast and forecast error are uncorrelated, i.e.,

$$\begin{aligned} E[\eta_{i,t,h}(s^e_{i,t,h} - s_t)] &= 0 \qquad (4) \\ &= \alpha_{i,h}E(s^e_{i,t,h} - s_t) + (\beta_{i,h} - 1)E(s^e_{i,t,h} - s_t)^2 + E[\varepsilon_{i,t,h}(s^e_{i,t,h} - s_t)] \end{aligned}$$

Thus, satisfaction of the joint hypothesis ($\alpha_{i,h}$, $\beta_{i,h}$) = (0,1) is also sufficient for weak efficiency with respect to the current forecast. However, it should be noted that (4) may still hold even if the joint hypothesis is rejected. Thus, satisfaction of the joint hypothesis represents sufficient conditions for both unbiasedness and this type of weak efficiency, but necessary conditions for neither.

If $\beta_{i,h}$= 1, then, whether or not $\alpha_{i,h}$= 0, the variance of the forecast error equals the variance of the bivariate regression residual, since then $var(\eta_{i,t,h}) = (\beta_{i,h} - 1)^2 var(s^e_{i,t,h} - s_t) + var(\varepsilon_{i,t,h}) + 2(\beta_{i,h} - 1)cov[(s^e_{i,t,h} - s_t), \varepsilon_{i,t,h}] = var(\varepsilon_{i,t,h})$. Figure 4 illustrates this point. Mincer and Zarnowitz (1969) required only that $\beta_{i,h}$= 1 in their definition of forecast efficiency. If in addition to $\beta_{i,h}$= 1, $\alpha_{i,h}$= 0, then the mean square forecast error also equals the variance of the forecast. Mincer and Zarnowitz emphasized that, as long as the loss function is symmetric, as is the case with a minimum mean square error criterion, satisfaction of the joint hypothesis implies optimality of forecasts.

4.1.2 Empirical results of joint tests

Since Hansen and Hodrick (1980), researchers have recognized that, when data are sampled more frequently than the forecast horizon (h), forecast errors may follow an h-1 period moving average process. The typical procedure has been to use a variance-covariance matrix which allows for generalized serial correlation. Throughout this paper, we use the Newey and West (1987) procedure, with the number of lagged residuals set to h-1. To ensure a positive semi-definite VCV matrix, we use a Bartlett window (see Hamilton, 1994, pp. 281-84.)

In Tables 1.1 we report results for the joint unbiasedness tests. We reject the joint hypothesis $(\alpha_{i,h}, \beta_{i,h}) = (0, 1)$ at the 5% significance level for all groups at the one-month horizon (indicating the possible role of inefficiency with respect to the current forecast), but only for the exporters at the three- and six- month horizons.

Now consider the results of the separate tests of the joint hypothesis. The significance of the $\alpha_{i,h}$'s in the joint regressions (2) generally deteriorates with horizon. There is only one rejection at the 5% level for each of the two shorter horizons. However, the $\alpha_{i,h}$'s are all rejected at this significance level for the six-month horizon. The test results for the $\beta_{i,h}$'s follow the opposite pattern with respect to horizon. The null that $\beta_{i,h}= 1$ is rejected for all groups at the one-month horizon, but only for the exporters at the three- and six-month horizons. This implies that weak efficiency with respect to the current forecast fails at the one-month horizon, but not at the longer horizons, with only two exceptions. Thus, it appears that the pattern of rejection of the joint hypothesis is predominantly influenced by tests of whether the slope coefficient equals one. That is, tests of the joint hypothesis at the one-month horizon are rejected due to failure of this type of weak efficiency, not simple unbiasedness.

For this reason, Mincer and Zarnowitz (1969) and Holden and Peel (1990) suggest that, if one begins by testing the joint hypothesis, rejections in this first stage should be followed by tests of the simple unbiasedness hypothesis in a second stage. Only if unbiasedness is rejected

in this second stage should one conclude that forecasts are biased. For reasons described below (in section 4.2), our treatment eliminates the first stage, so that unbiasedness and weak efficiency are separately assessed using the forecast error as the dependent variable.

Finding greater efficiency at the longer-horizon is unusual, because forecasting difficulty is usually thought to increase with horizon. However, the longer-horizon result may not be as conclusive as the $\beta_{i,h}$ statistics suggest. For all tests at all horizons, the null hypothesis that $\beta_{i,h}$ equals *zero* also cannot be rejected. Thus, for the longer two horizons (with just the one exception for exporters at the three-month horizon), hypothesis testing cannot distinguish between the null hypotheses that $\beta_{i,h}$ equals one or zero. Therefore, we cannot conclude that weak efficiency with respect to the current forecast holds while unbiasedness does not. The failure to precisely estimate the slope coefficient also produces R^2s that are below 0.05 in all regressions.[13] The conclusion is that testing only the joint hypothesis has the potential to obscure the difference in performance between the unbiasedness and weak efficiency tests. This conclusion is reinforced by an examination of figures 5 - 7, the scatter plots and regression lines for the bivariate regressions.[14] All three scatter plots have a strong vertical orientation. With this type of data, it is easy to find the vertical midpoint and test whether it is different from zero. Thus, (one-parameter) tests of simple unbiasedness are feasible. However, it is difficult to fit a precisely estimated regression line to this scatter, because the small variation in the forecast variable inflates the standard error of the slope coefficient. This explains why the $\beta_{i,h}$'s are so imprecisely estimated that the null hypotheses that $\beta_{i,h} = 1$ and 0 are simultaneously not rejected. This also explains why the R^2s are so low. Thus, examination of the scatter plots also reveal why bivariate regressions are potentially misleading about weak efficiency as well as simple unbiasedness. Therefore, in

[13] As we report in section 5, this lack of power is at least consistent with the failure to reject microhomogeneity at all three horizons.

[14] Note that, for illustrative purposes only, we compute the expectational variable as the four-group average percentage change in the forecast. However, recall that, despite the failure to reject micro-homogeneity at any horizon, the Figlewski-Wachtel critique implies that these parameter estimates are inconsistent in the presence of private information. (See the last paragraph in this subsection.)

contrast to both Mincer and Zarnowitz (1969) and Holden and Peel (1990), we prefer to separate tests for unbiasedness from tests for (all types of) weak efficiency at the initial stage. This obviates the need for a joint test. In the next section, we conduct such tests, making use of cointegration between forecast and realization where it exists.[15]

More fundamentally, the relatively vertical scatter of the regression observations around the origin is consistent with an approximately unbiased forecast of a random walk in exchange rate levels.[16] In figures 11, 12, and 13, we observe a corresponding time series pattern of variation between the forecasts and realizations in return form. As Bryant lamented in reporting corresponding regressions using a shorter sample from the JCIF, "the regression...is...not one to send home proudly to grandmother" (Bryant, 1995, p. 51). He drew the conclusion that "analysts should have little confidence in a model specification [e.g., uncovered interest parity] setting [the average forecast] exactly equal to the next-period value of the model...[M]odel-consistent expectations...presume a type of forward-looking behavior [e.g., weak efficiency] that is not consistent with survey data on expectations" (Bryant, 1995, p. 40).

4.2 Pretests for rationality: the stationarity of the forecast error

To test the null hypothesis of a unit root, we estimate the augmented Dickey-Fuller(1979) (ADF) regression

$$\Delta y_{t+1} = \alpha + \beta y_t + \gamma t + \sum_{k=1}^{p} \theta_k \Delta y_{t+1-k} + \epsilon_{t+1} \tag{5}$$

where y is the level and first difference of the spot exchange rate, the level and first difference of each group forecast, the residual from the (unrestricted) cointegrating regression and the forecast error (i.e., the residual from the "restricted" cointegrating equation). The number

[15] However, in the general case of biased and/or inefficient forecasts, Mincer and Zarnowitz (1969, p. 11) also viewed the bivariate regression "as a method of correcting the forecasts ... to improve [their] accuracy ... Theil (1966, p.33) called it the 'optimal linear correction.'" That is, the correction would involve 1) subtracting $\alpha_{i,h}$, then 2) multiplying by $1/\beta_{i,h}$. Graphically, this is a translation of the regression line followed by a rotation, until the regression line coincides with the 45 degree line.

[16] Other researchers (e.g., Bryant (1995)) have found similar vertical scatters for regressions where the independent variable, e.g., the forward premium/discount $f_{t,h} - s_t$, the "exchange risk premium" $f_{t,h} - s_{t+h}$, or the difference between domestic and foreign interest rates $(i - i^*)$, exhibits little variation.

of lagged differences to include in (5) is chosen by adding lags until a Lagrange Multiplier test fails to reject the null hypothesis of no serial correlation (up to lag 12). We test the null hypothesis of a unit root (i.e., $\beta = 0$) with the ADF t and z tests. We also test the joint null hypothesis of a unit root and no linear trend (i.e., $\beta = 0$ and $\gamma = 0$).

As can be seen in Tables 2.1 - 2.3, we fail to reject the null of a unit root in the log of the spot rate in two of the three unit root tests, but we reject the unit root in the hth difference for all three horizons. We conclude that the log of the spot rate is integrated of order one. Similarly, we conclude that the log of the forecast of each spot rate is integrated of order one. Thus, we can conduct cointegration tests on the spot rate and each corresponding forecast. The null of a unit root in the (unrestricted) residual in the "cointegrating regression" equation 6 is rejected at the 10% level or less for all groups and horizons except group three (exporters) at the six-month horizon. Thus, we can immediately reject unbiasedness for the latter group and horizon. Next, since a stationary forecast error is a necessary condition for unbiasedness, we test for unbiasedness (as well as) and weak efficiency in levels using Liu and Maddala's (1992) method of "restricted cointegration." This specification imposes the joint restriction $\alpha_{i,h} = 0$, $\beta_{i,h} = 1$ on the bivariate regression

$$s_{t+h} = \alpha_{i,h} + \beta_{i,h} s^e_{i,t,h} + \varepsilon_{i,t,h} \tag{6}$$

and tests whether the residual (the forecast error) is non-stationary. In a bivariate regression, any cointegrating vector is unique. Therefore, if we find that the forecast errors are stationary, then the joint restriction is not rejected, and (0,1) must be the unique cointegrating vector.[17] The advantage of the one-step restricted cointegration is that, if the joint hypothesis is true, then tests which impose this cointegrating vector have greater power than those which estimate a cointegrating vector. See, e.g., Maynard and Phillips (2001).

Note that the Holden and Peel (1990) critique does not apply in the I(1) case, because the intercept cannot be an unbiased forecast of a nonstationary variable. Thus, the cointegrating

[17] It is also possible to estimate the cointegrating parameters and jointly test whether they are zero and one. A variety of methods, such as those due to Saikkonen (1991) or Phillips and Hansen (1990), exist that allow for inference in cointegrated bivariate regressions.

regression line of the level realization on the level forecast must have both $\alpha = 0$ and $\beta = 1$ for unbiasedness to hold. This differs from Figure 3, the scatterplot in differences, where $\alpha_{i,h} = 0$ but $\beta_{i,h} \neq 1$. Intuitively, the reason for the difference in results is that the scatterplot in levels must lie in the first quadrant, i.e., no negative values of the forecast or realization.

At the one-month horizon, the null of a unit root in the residual of the restricted cointegrating regression (i.e., the forecast error) is rejected at the 1% level for all groups. We find nearly identical results at the three-month horizon; the null of a unit root in the forecast error is rejected at the 5% level for all groups. Thus, for these regressions we can conduct rationality tests by regressing the forecast error on a constant (hypothesized equal to zero for unbiasedness) and other information set variables (whose coefficients are hypothesized equal to zero for efficiency). (Recall just above that we failed to reject the null of a unit root in the unrestricted residual for the six-month forecasts of exporters.) Now, in the case of the restricted residual, the other three groups failed to reject a unit root at the 10% level in two out of three of the unit root tests.[18] (See figures 8, 9, and 10.) Thus, in contrast to the results for the two shorter horizons, at the six-month horizon, the evidence is clearly in favor of a unit root in the forecast error for all four groups. Therefore, we reject the null of simple unbiasedness because a forecast error with a unit root cannot be mean zero. In fact, given our finding of a unit root in the forecast errors, rationality tests regressing the forecast error on a constant and/or other information set variables would be invalid.

4.3 Univariate tests for unbiasedness

The unbiasedness equation is specified as

$$\eta_{i,t,h} = s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \epsilon_{i,t,h}, \tag{7}$$

where $\eta_{i,t,h}$ is the forecast error of individual i, for an h-period-ahead forecast made at time t. The results are reported in Tables 3.1 and 3.2. For the one-month horizon, unbiasedness cannot be rejected at conventional significance levels for any group. For the three-month

[18] As expected, exporters failed to reject at the 10% level in all three tests.

horizon, unbiasedness is rejected only for exporters (at a p-value of 0.03). As we saw in the previous subsection, rationality is rejected for all groups at the six-month horizon, due to nonstationary forecast errors.[19]

In these unbiasedness tests, as well as all others, it is possible that coefficient estimates for the entire sample are not stable over subsamples. The lower panels of 3.1 and 3.2 contain results of the test for equality of intercepts in four equal subperiods, each consisting of approximately 75 biweekly forecasts:

$$\eta_{i,t,h} = s_{t+h} - s^e_{i,t,h} = \alpha_{i,h,1} + \alpha_{i,h,2} + \alpha_{i,h,3} + \alpha_{i,h,4} + \epsilon_{i,t,h}. \tag{8}$$

For both one- and three-month horizons, all four forecaster groups undervalued the yen in the first and third subperiods. This is understandable, as both these subperiods were characterized by overall yen appreciation. (See Fig. 1.) Evidently, forecasters underestimated the degree of appreciation. Exporters were the only group to undervalue the yen in the last subperiod as well, although that was not one of overall yen appreciation. This is another perspective on the "wishful thinking" of exporters.[20]

The main difference between the two horizons is in the significance of the test for structural breaks. For the one-month horizon, the estimates of the individual break dummies generally do not reach statistical significance, and the test for their equality rejects only for the exporters. Thus, the exporters' bias was not constant throughout the sample. In contrast, for the three-month horizon, the test for no structural breaks is rejected at the 5% level for all groups, even though unbiasedness itself is rejected for the full sample only for exporters. Even setting aside the bias and variability of exporters' forecasts, our structural break tests allow us to conclude that there is considerably more variation around roughly zero

[19]The direction of the bias for exporters is negative; that is, they systematically underestimate the value of the yen, relative to the dollar. Ito (1990) found the same tendency using only the first two years of survey data (1985-1987). He characterized this depreciation bias as a type of "wishful thinking" on the part of exporters.

[20]Ito (1994) conducted a similar analysis for the aggregate of all forecasters, but without an explicit test for structural breaks.

mean forecast errors at the longer horizon. This probably reflects the additional uncertainty inherent in longer-term forecasts.[21]

4.4 Unbiasedness tests using Error Correction Models

As mentioned at the beginning of the previous subsection, the Error Correction Model provides an alternate specification for representing the relationship between cointegrated variables.

$$\begin{aligned} s_{t+h} - s_t &= \alpha_{i,h}(s_t - \gamma_{i,h} s^e_{i,t-h,h}) + \beta_{i,h}(s^e_{i,t,h} - s^e_{i,t-h,h}) \\ &\quad + \delta_i(\text{lags of } s_{t+h} - s_t) + \eta_i(\text{lags of } s^e_{i,t,h} - s^e_{i,t-h,h}) + \varepsilon_{i,t,h} \end{aligned} \tag{9}$$

The regressors include the smallest number of lagged dependent variables required such that we do not reject the hypothesis that the residuals are white noise. We impose $\gamma_{i,h} = 1$ when "restricted" cointegration of s_{t+h} and $s^e_{i,t,h}$ is not rejected. Recall that one- and three-month forecast errors were found to be stationary, so it was for these two horizons that estimation of the simple unbiasedness equation was possible. Although it would be valid to estimate the ECM at the six-month horizon using the (unrestricted) stationary cointegrating residual (i.e., for all groups but exporters), we elect not to, because the nonstationarity of the forecast error itself implies a failure of the unbiasedness restrictions.[22]

According to this specification of the ECM, the change in the spot rate is a function of the change in the forecast, interpreted as a short-run effect, and the current forecast error, interpreted as a long-run adjustment to past disequilibria. $\alpha_{i,h}$, the coefficient of the error correction term, represents the fraction of the forecast error observed at t-h that is corrected by time t. A negative coefficient indicates a stabilizing adjustment of expectations. This formulation of the ECM has the advantage that the misspecification (due to omitted variable

[21] This is consistent with the finding of nonstationary forecast errors for all groups at the six-month horizon.

[22] Our empirical specification of the ECM also includes an intercept. This will help us to determine whether there are structural breaks in the ECM.

bias) of the regression of the differenced future spot rate on the differenced current forecast can be gauged by the statistical significance of the error correction term.[23]

Then, as first asserted by Hakkio and Rush (1989), the unbiasedness restriction is represented by the joint hypothesis that $-\alpha_{i,h} = \beta_{i,h} = 1$ and all δ and η coefficients equal zero.[24] (The hypothesized coefficient on the error correction term of -1 reflects the unbiasedness requirement that the entire forecast error is corrected within the forecast horizon h.) We also test unbiasedness without including lagged dependent variables but incorporating robust standard errors which allow for generalized serial correlation and heteroscedasticity. This allows comparison with the univariate and bivariate unbiasedness equations.

First, we compare the ECM results to the joint unbiasedness restrictions in the returns regressions, using robust standard errors in both cases. Although the estimated coefficient of the error correction term is generally negative, indicating a stable error correction mechanism,[25] the coefficient does not reach a 5% significance level in any of the regressions. Thus, there is little evidence that the error correction term plays a significant role in the long-run dynamics of exchange rate changes. The R^2s in the ECM, while never more than 0.044, still are greater than in the joint unbiasedness specification, typically by factors of three to five.

Second, we compare the ECM results to the univariate simple unbiasedness regressions, again using robust standard errors in both cases. The ECM unbiasedness restrictions are rejected at a 5% level more often than in the simple unbiasedness tests. Whereas the only rejection of simple unbiasedness at the shorter two horizons is for exporters at the three-

[23]Zacharatos and Sutcliffe (2002) note that the inclusion of the contemporaneous spot forecast (in their paper, the forward rate) as a regressor assumes that the latter is weakly exogenous; that is, deviations from unbiasedness are corrected only by movements in the realized spot rate. These authors prefer a bivariate ECM specification, in which the change in the future spot rate and the change in the contemporaneous forecast are functions of an error correction term and lags of the dependent variables. However, Zivot (1998) points out that, if the spot rate and forecast are contemporaneously correlated, then our single equation specification does not make any assumptions about the weak exogeneity of the forecast.

[24]Since we include an intercept, we also test the restriction that the intercept equals zero–both individually and as part of the joint unbiasedness hypothesis.

[25]The only exception is for exporters at the one-month horizon.

month horizon, the ECM restrictions are rejected for three out of four groups at the one-month horizon as well as for exporters at the three-month horizon.

While it is uncontroversial that, for testing unbiasedness, the ECM is preferred to the conventional bivariate specification in returns, it is not at all clear that the ECM is preferred to the simple univariate test of unbiasedness. Can the more decisive rejections of unbiasedness using the ECM versus the simple univariate specification be reconciled?[26]

One way to proceed is to determine whether the unbiasedness restrictions imposed on the ECM are necessary as well as sufficient, as is the case for the simple unbiasedness test, or just sufficient, as is the case for the bivariate unbiasedness test. Thus, it is possible that the stronger rejections of unbiasedness in the ECM specification are due to the implicit test of weak efficiency with respect to the current forecast. That is, the Holden and Peel (1990) critique applies to the Hakkio and Rush (1989) test in (9), as well as the joint unbiasedness test in the returns regression. Setting $\beta_{i,h}$, the coefficient of the contemporaneous differenced forecast, equal to one produces an ECM in which the dependent variable is the forecast error:

$$s_t - s^e_{i,t,h} = (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h}) \tag{10}$$

Thus, in the ECM the necessary and sufficient condition for unbiasedness is that $\alpha_{i,h}$equals -1.[27] Table 4.1 contains tests of this conjecture. Here the joint hypothesis that the intercept equals zero and $\alpha_{i,h}$ equals minus one produces exactly the same results as in the simple unbiasedness tests.[28] It is interesting that, even when we can decouple the test for weak efficiency with respect to the current forecast from the unbiasedness test, the test of unbiasedness using this ECM specification still requires weak efficiency with respect to the current forecast error.[29]

[26]The standard errors in the univariate regression are about the same as those for the ECM. (By definition, of course, the R^2s for the univariate regression equal zero.)

[27]Since we estimate the restricted ECM with an intercept, unbiasedness also requires the intercept to be equal to zero.

[28]Since the intercept in equation 10 is not significant in any regression, the simple hypothesis that $\alpha_{i,h}$ equals one also fares the same as the simple unbiasedness tests.

[29]For purposes of comparison with both the bivariate joint and simple unbiasedness restrictions, we have used the ECM results using the robust standard errors. In all cases testing the ECM restrictions using

4.5 Explicit tests of weak efficiency

The literature on rational expectations exhibits even less consensus as to the definition of efficiency than it does for unbiasedness. In general, an efficient forecast incorporates all available information—private as well as public. It follows that there should be no relationship between forecast error and any information variables known to the forecaster at the time of the forecast. Weak efficiency commonly denotes the orthogonality of the forecast error with respect to functions of the target and prediction. For example, there is no contemporaneous relationship between forecast and forecast error which could be exploited to reduce the error. Strong efficiency denotes orthogonality with respect to the remaining variables in the information set. Below we perform two types of weak efficiency tests. In the first type, we regress each group's forecast error on three sets of weak efficiency variables.[30]

1. Single and cumulative lags of the mean forecast error (lagged one period):

$$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{k=h+1}^{h+7} \beta_{i,t+h-k}(s_{t+h-k} - s^e_{m,t+h-k,h}) + \epsilon_{i,t,h} \tag{11}$$

2. Single and cumulative lags of mean expected depreciation (lagged one period):

$$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{k=h+1}^{h+7} \beta_{i,t+h-k}(s^e_{m,t+h-k,h} - s_{t-k}) + \epsilon_{i,t,h} \tag{12}$$

3. Single and cumulative lags of actual depreciation:

$$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{k=h}^{h+6} \beta_{i,t+h-k}(s_{t+h-k} - s_{t-k}) + \epsilon_{i,t,h} \tag{13}$$

For each group and forecast horizon, we regress the forecast error on the most recent seven lags of the information set variable, both singly and cumulatively. We use a Wald test of

F-statistics based on whitened residuals produces rejections of all restrictions, simple and joint, except a zero intercept. Hakkio and Rush (1989) found similarly strong rejections of (9), where the forecast was the forward rate.

[30] Notice that the first two sets of weak efficiency variables include the *mean* forecast, rather than the individual group forecast. Our intention is to allow a given group to incorporate information from other groups' forecasts via the *prior* mean forecast. This requires an extra lag in the information set variables, relative to a contemporaneously available variable such as the realized exchange rate depreciation.

the null hypothesis $\alpha_{i,h} = \beta_{i,t+h-k} = 0$ and report chi-square test statistics, with degrees of freedom equal to the number of regressors excluding the intercept. If we were to perform only simple regressions (i.e., on each lag individually), estimates of coefficients and tests of significance could be biased toward rejection due to the omission of relevant variables. If we were to perform only multivariate regressions, tests for joint significance could be biased toward nonrejection due to the inclusion of irrelevant variables. It is also possible that joint tests are significant but individual tests are not. This will be the case when the linear combination of (relatively uncorrelated) regressors spans the space of the dependent variable, but individual regressors do not.

In the only reported efficiency tests on JCIF data, Ito (1990) separately regressed the forecast error (average, group, and individual firm) on a single lagged forecast error, lagged forward premium, and lagged actual change. He found that, for the 51 biweekly forecasts between May 1985 and June 1987, rejections increased from a relative few at the one- or three-month horizons to virtual unanimity at the six-month horizon. When he added a second lagged term for actual depreciation, rejections increased "dramatically" for all horizons.

The second type of weak efficiency tests uses the Breusch-Godfrey(1978) LM test for the null of no serial correlation of order k=h or greater, up to order k=h+6, in the residuals of the forecast error regression, equation (11).[31,32] Results for all efficiency tests for the one- and three-month horizons are presented in Tables 5.1 - 5.8. (Recall that the nonstationarity of the forecast errors at the six-month horizon is an implicit rejection of weak efficiency.) For each group, horizon, and variable, there are seven individual tests, i.e., on a single lag, and six joint tests, i.e., on multiple lags. These 13 tests are multiplied by four groups times two horizons times three weak efficiency variables for a total of 312 efficiency tests.

[31]This is a general test, not only because it allows for an alternative hypothesis of higher-order serial correlation of specified order, but also because it allows for serial correlation to be generated by AR, MA or ARMA processes.

[32]We use the F-statistic because the χ^2 test statistics tend to over-reject, while the F-tests have more appropriate significance levels (see Kiviet (1987)).

Using approximately nine more years of data than Ito (1990), we find many rejections. In some cases, nearly all single lag tests are rejected, yet few if any joint tests are rejected. (See, e.g., expected depreciation at the three-month horizon.) In other cases, nearly all joint tests are rejected, but few individual tests. (See, e..g, actual depreciation at the three-month horizon.) Remarkably, all but one LM test for serial correlation at a specified lag produces a rejection at less than a 10% level, with most at less than a 5% level. Thus, it appears that the generality of the alternative hypothesis in the LM test permits it to reject at a much greater rate than the conventional weak efficiency tests, in which the variance-covariance matrix incorporates the Newey-West-Bartlett correction for heteroscedasticity and serial correlation. Finally, unlike Ito (1990), we find no strong pattern between horizon length and number of rejections.

5 Microhomogeneity tests

In addition to testing the rationality hypotheses at the individual level, we are interested in the degree of heterogeneity of coefficients across forecasters. Demonstrating that individual forecasters differ systematically in their forecasts (and forecast generating processes) has implications for the market microstructure research program. As Frankel and Froot (1990, p. 182) noted, "the tremendous volume of foreign exchange trading is another piece of evidence that reinforces the idea of heterogeneous expectations, since it takes differences among market participants to explain why they trade."

Micro-homogeneity should have implications for rationality, as well. Intuitively, if all forecasters pass rationality tests, then their corresponding regression coefficients should be equal. However, the converse is not necessarily true: if all forecasters have equal regression coefficients, they will not satisfy rationality conditions if they are all biased or inefficient to the same degree with respect to the same variables. For the univariate unbiasedness regressions, the null of micro-homogeneity is given by $H_0 : \alpha_{ih} = \alpha_{jh}$, for all $i, j \neq i$. Before testing for homogeneous intercepts in equation (7) we must specify the form for our GMM system variance-covariance matrix. Keane and Runkle (1990) first accounted for

cross-sectional correlation (in price level forecasts) using a GMM estimator on pooled data. Bonham and Cohen (2001) tested the pooling specification by replacing Zellner's (1962) SUR variance-covariance matrix with a GMM counterpart that incorporates the Newey-West single equation corrections (used in our individual equation tests above) plus allowances for corresponding cross-covariances, both contemporaneous and lagged. Bonham and Cohen (2001) constructed a Wald statistic for testing the micro-homogeneity of individual forecaster regression coefficients in a system.[33]

Keane and Runkle (1990) provided some empirical support for their modeling of cross-sectional correlations, noting that the average covariance between a pair of forecasters is 58% of the average forecast variance. In contrast, we use Pesaran's (2004) CD (cross-sectional dependence) test to check for lagged as well as contemporaneous correlations of forecast errors among pairs of forecasters.

$$CD = \sqrt{\frac{2T}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij}, \tag{14}$$

where T is the number of time periods, N = 4 is the number of individual forecasters, and $\hat{\rho}_{ij}$ is the sample correlation coefficient between forecasters i and j, $i \neq j$. Under the null hypothesis of no cross-correlation, $CD \overset{\mathbf{a}}{\sim} N(0,1)$.[34] See Table 6.0 for CD test results. We tested for cross-correlation in forecast errors from lag zero up to lags four and eight, for the one-month and three month forecast horizons, respectively. (The nonstationarity of the six-month forecast error precludes using the CD test at that horizon.) At the one-month horizon, cross-correlations from lags zero through 4 are each significant at the 5% level. Since rational forecasts allow for (individual) serial correlation of forecast errors at lags of h-1 or less, and h= 2 for the one-month horizon, the cross-correlations at lags two through four indicate violations of weak efficiency. Similarly, at the three-month horizon, where h-1=5,

[33]Elliott and Ito (1999) used single equation estimation that incorporated a White correction for heteroscedasticity and a Newey-West correction for serial correlation. (See the discussion below of Ito's tests of forecaster heterogeneity.)

[34]Unlike Breusch and Pagan's (1980) LM test for cross-sectional dependence, Pesaran's (2004) CD test is robust to multiple breaks in slope coefficients and error variances, as long as the unconditional means of the variables are stationary and the residuals are symmetrically distributed.

there is significant cross-correlation at lag six.[35] However, it should be noted that, for many lags shorter than h, one cannot reject the null hypothesis that there are no cross-correlated forecast errors.

Nevertheless, in our micro-homogeneity tests, we follow Bonham and Cohen (2001), allowing for an MA(h-1) residual process, both individually and among pairs of forecast errors. (See the appendix for details.) By more accurately describing the panel's residual variance-covariance structure, we expect this systems approach to improve the consistency of our estimates. Consider first the four bivariate regressions. Recall that we rejected the joint hypothesis $(\alpha_{i,h}, \beta_{i,h}) = (0, 1)$ at the 5% significance level for all groups at the one-month horizon (indicating the possible role of inefficiency with respect to the current forecast), but only for the exporters at the three- and six-month horizons. However, there are no rejections of micro-homogeneity for any horizon.[36]

The micro-homogeneity test results are very different for the one- and three-month systems of univariate unbiasedness regressions. (Recall that unbiasedness was rejected for all groups at the six-month horizon due to the nonstationarity of the forecast error.) Despite having only one failure of unbiasedness at the 5% level for the two shorter horizons, micro-homogeneity is rejected at a level of virtually zero for both horizons. The rejection of micro-homogeneity at the one-month horizon occurs despite the failure to reject unbiasedness for any of the industry groups. We hypothesize that the consistent rejection of micro-homogeneity regardless of the results of individual unbiasedness tests is the result of sufficient variation in individual bias estimates as well as precision in these estimates. Ac-

[35] There are three instances of statistically significant negative test statistics for lags greater than h-1, none for lags less than or equal to h-1. Thus, some industries produce relatively high forecast errors several periods after others produce relative low forecast errors, and this information is not fully incorporated in some current forecasts.

[36] The nonrejection of micro-homogeneity in bivariate regressions does not, however, mean that one can avoid aggregation bias by using the mean forecast. Even if the bivariate regressions were correctly interpreted as joint tests of unbiasedness and weak efficiency with respect to the current forecast, and even if the regressions had sufficient power to reject a false null, the micro-homogeneity tests would be subject to additional econometric problems. According to the Figlewski-Wachtel (1983) critique, successfully passing a pre-test for micro-homogeneity does not ensure that estimated coefficients from such consensus regressions will be consistent. See section 2.1.

cording to these tests, aggregation of individual forecasts into a mean forecast is invalid at all horizons.

In addition to testing the weak efficiency hypothesis at the individual level, we are interested in the degree of heterogeneity of coefficients across forecasters. Here the null of micro-homogeneity is given by $H_0 : \phi_{il} = \phi_{jl}$, for $l = h, \ldots h+6$, for all $i, j \neq i$. As explained in the section on efficiency tests, there are 312 tests (not 468, due to a nonstationary forecast error for all four groups at the six-month horizon) / 4 groups = 83 micro-homogeneity tests. The null hypothesis of equal coefficients is $H_0 : \alpha_{i,h} = \alpha_{j,h}, \beta_{i,t+h-k} = \beta_{j,t+h-k}$ for all $i, j \neq i$. As with the micro-homogeneity tests for unbiasedness, our GMM variance-covariance matrix accounts for serial correlation of order h-1 or less, generalized heteroscedasticity, and cross-sectional correlation or order h-1 or less. We report $\chi^2(n)$ statistics, where n is the number of coefficient restrictions, with corresponding p-values. Rather than perform all 83 micro-homogeneity tests, we choose a sample consisting of the shortest and longest lag for which there are corresponding individual and joint tests (i.e., for the k=h+1st and k=h+6th lag). Thus, there are 4 tests (two individual and two corresponding joint tests) times two horizons times three variables for a total of 24 tests. Every one of the micro-homogeneity tests are rejected at the 0% level. As pointed out by Bryant (1995), a finding of micro-heterogeneity in unbiasedness and weak efficiency tests also casts doubt on the assumption of a rational representative agent commonly used in macroeconomic and asset pricing models.

5.1 Ito's heterogeneity tests

In Table 7.1, we replicate Ito's (1990) and Elliott and Ito's (1999) test for forecaster "heterogeneity." This specification regresses the deviation of the individual forecast from the cross-sectional average forecast on a constant. Algebraically, Ito's regression can be derived from the individual forecast error regression by subtracting the mean forecast error regression. Thus, because it simply replaces the forecast error with the individual deviation

from the mean forecast, it does not suffer from aggregation bias (c.f. Figlewski and Wachtel (1983)) or pooling bias (c.f. Zarnowitz (1985)). [37,38]

$$s^e_{i,t,h} - s^e_{m,t,h} = (\alpha_{i,h} - \alpha_m) + (\varepsilon_{i,t,h} - \varepsilon_{m,t}) \tag{15}$$

As above, we use the Newey-West-Bartlett variance-covariance matrix.

One may view Ito's "heterogeneity" tests as complementary to our micro-homogeneity tests. On the one hand, one is not certain whether a single (or pair of?) individual rejection(s) of, say, the null hypothesis of a zero mean deviation in Ito's test would result in a rejection of micro-homogeneity overall. On the other hand, a rejection of micro-homogeneity does not tell us which groups are the most significant violators of the null hypothesis. It turns out that Ito's mean deviation test produces rejections at a level of 6% or less for all groups at all horizons except for banks and brokers at the one-month horizon and life insurance and import companies at the six-month horizon.[39] Since Ito's regressions have a similar form (though not a similar economic interpretation) to the tests for univariate unbiasedness in Tables 3.1 and 3.2, it is not surprising that micro-homogeneity tests on the four-equation system of Ito equations produce rejections at a level of virtually zero for all three horizons.

6 Conclusions

In this paper, we undertake a reexamination of the rationality and diversity of JCIF forecasts of the yen-dollar exchange rate. In several ways we update and extend the seminal paper by Ito (1990). In particular, we have attempted to explore the nature of rationality tests on integrated variables. We show that tests based on the "conventional" bivariate regression in returns, while correctly specified in terms of integration accounting, have two

[37] Recall that our group results are not entirely comparable to Ito's (1990), since our dataset, unlike his, combines insurance companies and trading companies into one group, and life insurance companies and import-oriented companies into another group.

[38] Chionis and MacDonald (1997) performed an Ito-type test on individual expectations data from Consensus Forecasts of London.

[39] Elliott and Ito (1999), who have access to forecasts for the 42 individual firms in the survey, find that, for virtually the same sample period as ours, the null hypothesis of a zero deviation from the mean forecast is rejected at the 5% level by 17 firms at the one-month horizon, 13 firms for the three-month horizon, and 12 firms for the six-month horizon. These authors do not report results by industry group.

major shortcomings. First, following Holden and Peel (1990), they are misspecified as unbiasedness tests, because rejection of the (0,1) restriction on the slope and intercept is a sufficient, not a necessary, condition for unbiasedness. Only a zero restriction on the intercept in a regression of the forecast error on a constant is both necessary and sufficient for unbiasedness. Second, tests using the bivariate specification suffer from a lack of power. Yet, this is exactly what we would expect in an asset market whose price is a near random walk: the forecasted change is nearly unrelated to (and varies much less than) the actual change.

In contrast, we conduct pretests for rationality based on determining whether the realization and forecast are each integrated and cointegrated. In this case, following Liu and Maddala (1992), a "restricted" cointegration test, which imposes a (0,1) restriction on the cointegrating vector, is necessary for testing unbiasedness. (We show that the Holden and Peel (1990) critique does not apply if the regressor and regressand are cointegrated.) If a unit root in the restricted residual is rejected, then the univariate test which regresses the forecast error on a constant is equivalent to the restricted cointegration test. Testing this regression for white noise residuals is one type of weak efficiency test. Testing other stationary regressors in the information set for zero coefficients produces additional efficiency tests.

In the univariate specification, we find that, for each group, the ability to produce unbiased forecasts deteriorates with horizon length: no group rejects unbiasedness at the one-month horizon, but all groups reject at the six-month horizon, because the forecast errors are nonstationary. Exporters consistently perform worse than the other industry groups, with a tendency toward depreciation bias. Using only two years of data, Ito (1990) found the same result for exporters, which he described a a type of "wishful thinking."

The unbiasedness results are almost entirely reversed when we test the hypothesis using the conventional bivariate specification. That is, the joint hypothesis of zero intercept and unit slope is rejected for all groups at the one-month horizon, but only for exporters and the three- and six-month horizons. Thus, in stark contrast to the univariate unbiasedness

tests, as well as Ito's (1990) bivariate tests, forecast performance does not deteriorate with increases in the horizon.

Also, since Engle and Granger (1987) have showed that cointegrated variables have an error correction representation, we impose joint "unbiasedness" restrictions first used by Hakkio and Rush (1989) on the ECM. However, we show that these restrictions also represent sufficient, not necessary, conditions, so these tests could tend to over-reject. We then develop and test restrictions which are both necessary and sufficient conditions for unbiasedness. The test results confirm that the greater rate of rejections of the joint "unbiasedness" restrictions in the ECM is caused by the failure of the implicit restriction of weak efficiency with respect to the lagged forecast. When we impose the restriction that the coefficient of the forecast equals one, the ECM unbiasedness test results mimic those of the simple univariate unbiasedness tests. For this dataset, at least, it does not appear that an ECM provides any value added over the simple unbiasedness test. Furthermore, since the error correction term is not statistically significant in any regressions, it is unclear whether the ECM provides any additional insight into the long-run adjustment mechanism of exchange rate changes.

The failure of more general forms of weak efficiency is borne out by two types of explicit tests for weak efficiency. In the first type, we regress the forecast error on single and cumulative lags of mean forecast error, mean forecasted depreciation, and actual depreciation. We find many rejections of unbiasedness. In the second type, we use the Godfrey (1978) LM test for serial correlation of order h through h+6 in the residuals of the forecast error regression. Remarkably, all but one LM test at a specified lag length produces a rejection at less than a 10% level, with most at less than a 5% level. (As in the case of the univariate unbiasedness test, all weak efficiency tests at the six-month horizon fail due to the nonstationarity of the forecast error.)

Whereas Ito (1990) and Elliott and Ito (1999) measured diversity as a statistically significant deviation of an individual's forecast from the cross-sectional average forecast, we perform a separate test of micro-homogeneity for each type of rationality test–unbiasedness

as well as weak efficiency–that we first conducted at the industry level. In order to conduct the systems estimation and testing required for the micro-homogeneity test, our GMM estimation and inference makes use of an innovative variance-covariance matrix that extends the Keane and Runkle (1990) counterpart from a pooled to an SUR-type structure. Our variance-covariance matrix takes into account not only serial correlation and heteroscedasticity at the individual level (via a Newey-West-Bartlett correction), but also forecaster cross-correlation up to h-1 lags. We document the statistical significance of the cross-sectional correlation using Pesaran's (2004) CD test.

In the univariate unbiasedness tests, we find that, irrespective of the ability to produce unbiased forecasts at a given horizon, micro-homogeneity is rejected at virtually a 0% level for all horizons. We find this result to be somewhat counterintuitive, in light of our prior belief that micro-homogeneity would be more likely to obtain if there were no rejections of unbiasedness. Evidently, there is sufficient variation in the estimated bias coefficient across groups and/or high precision of these estimates to make the micro-homogeneity test quite sensitive. Micro-homogeneity is also strongly rejected in the weak efficiency tests.

In contrast to the results with the univariate unbiasedness specification, micro-homogeneity is not rejected at any horizon in the bivariate regressions. We conjecture that the imprecise estimation of the slope coefficient makes it difficult to reject joint hypotheses involving this coefficient.

In conclusion, we recommend that all rationality tests be undertaken using simple univariate specifications at the outset (rather than only if the joint bivariate test is rejected, as was suggested by Mincer and Zarnowitz (1969) and Holden and Peel (1990) and employed by Gavin (2003)). Before conducting such rationality tests, one should test the restricted cointegrated regression residuals, i.e., the forecast error, for stationarity. Clearly, integration accounting and regression specification matter for rationality testing.

While our rationality tests do not attempt to explain cross-sectional dispersion, the widespread rejection of micro-homogeneity in different specifications of unbiasedness and

weak efficiency tests[40] provide more motivation for the classification of forecasters into types (e.g., fundamentalist and chartist/noise traders) than for simply assuming a representative agent (with rational expectations).

There are characteristics of forecasts other than rationality which are of intrinsic interest. Given our various rejections of rational expectations, it is natural to explore what expectational mechanism the forecasters use. Ito (1994) tested the mean JCIF forecasts for extrapolative and regressive expectations, as well as a mixture of the two.[41] Cohen and Bonham (2006) extend this analysis using individual forecast generating processes and additional specifications of learning models. In addition, much of the literature on survey forecasts has analyzed the accuracy of predictions, typically ranking forecasters by MSE (raw or relative). One relatively unexplored issue is the statistical significance of the ranking, regardless of loss function. However, other loss functions, especially nonsymmetric ones, are also reasonable. For example, Elliott and Ito (1999) have ranked individual JCIF forecasters using a profitability criterion. As mentioned in section 2.2, the loss function may incorporate strategic considerations that result in "rational bias." Such an exploration would require more disaggregated data than the JCIF industry forecasts to which we have access.

[40] We put less weight on the results of the weaker tests for micro-homogeneity in the bivariate regression framework.

[41] He also included regressors for adaptive expectations and the forward premium.

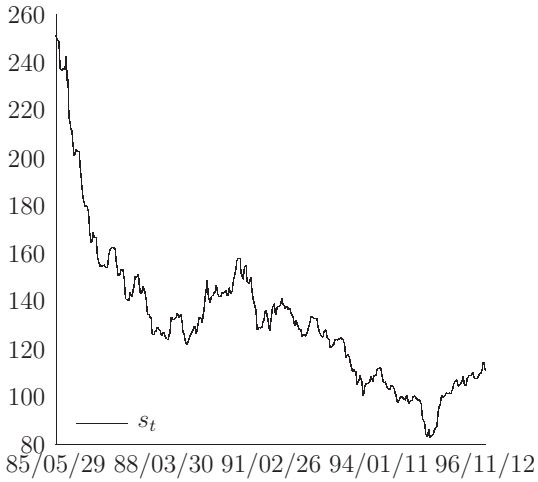

Figure 1 ¥/$ vs Time

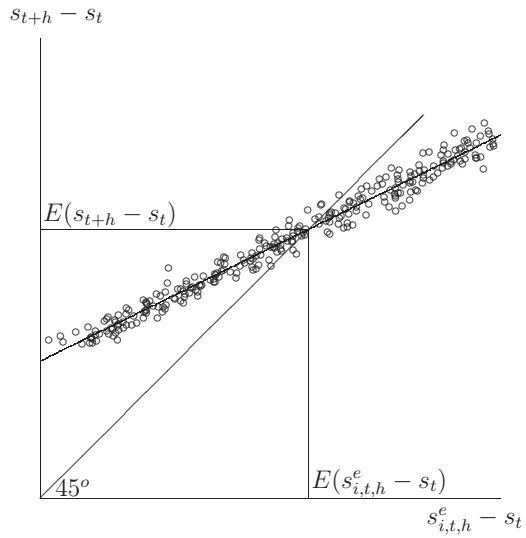

Figure 2 Unbiasedness with $\alpha \neq 0$, $\beta \neq 1$

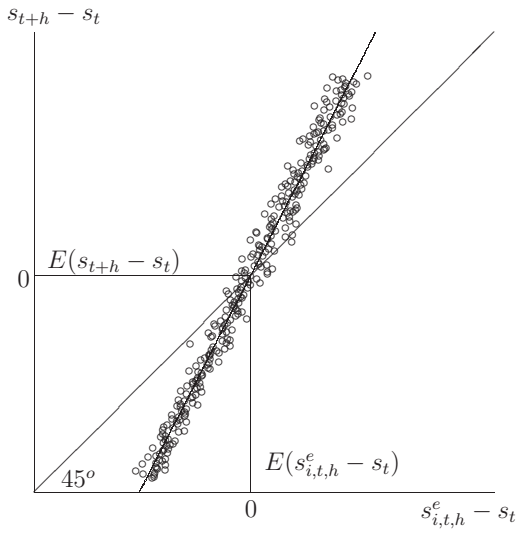

Figure 3 Unbiasedness with $\alpha = 0$, $\beta \neq 1$

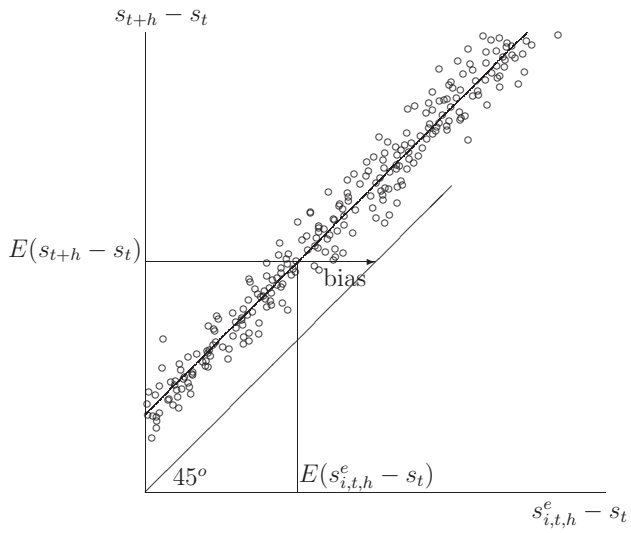

Figure 4 Bias with $\alpha \neq 0$, $\beta = 1$

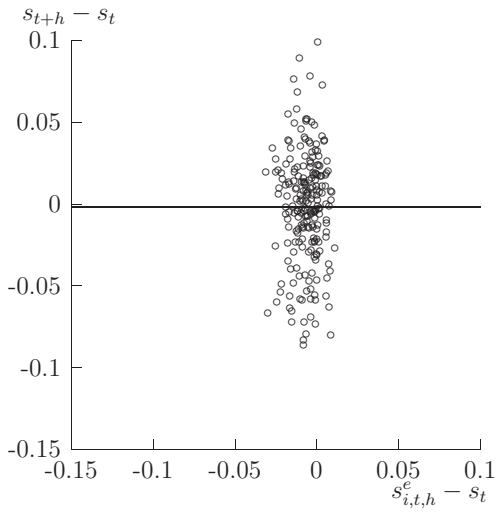

Figure 5 Actual vs Expected Depreciation, 1 Month-Ahead Forecast

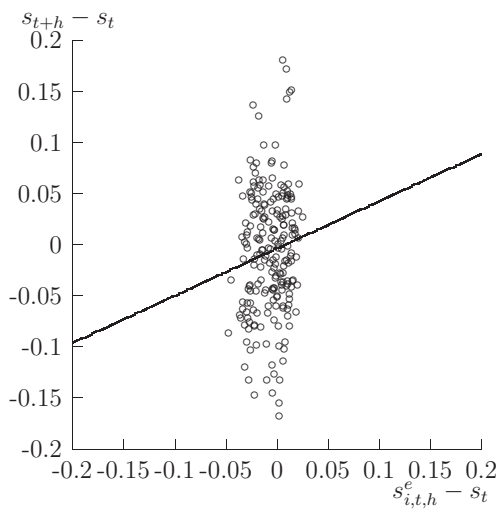

Figure 6 Actual vs Expected Depreciation, 3 Month-Ahead Forecast

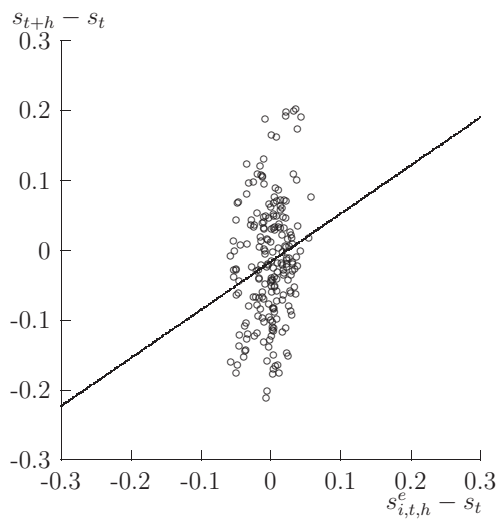

Figure 7 Actual vs Expected Depreciation: 6 Month-Ahead Forecast

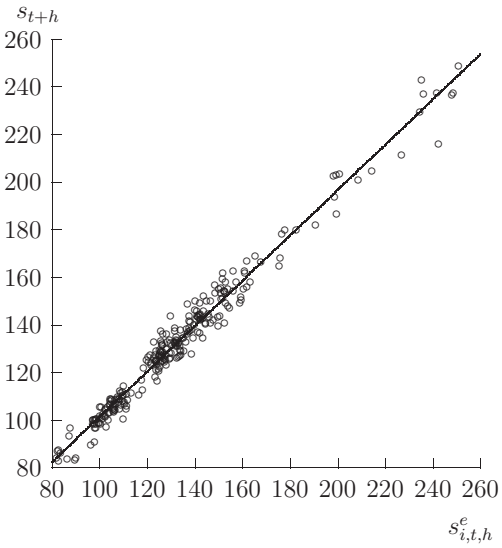

Figure 8 Actual vs Expected Exchange Rate: 1 Month-Ahead Forecast

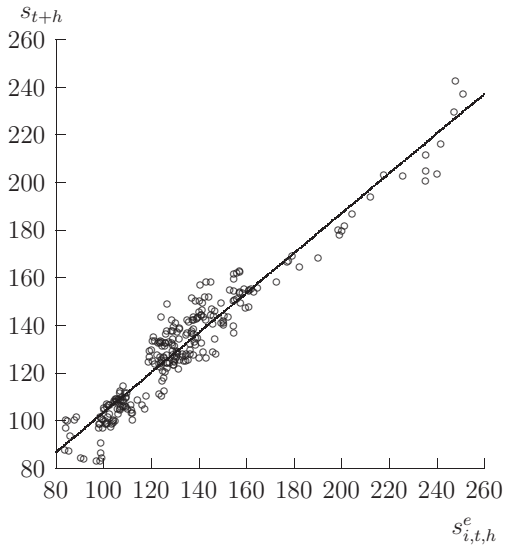

Figure 9 Actual vs Expected Exchange Rate: 3 Month-Ahead Forecast

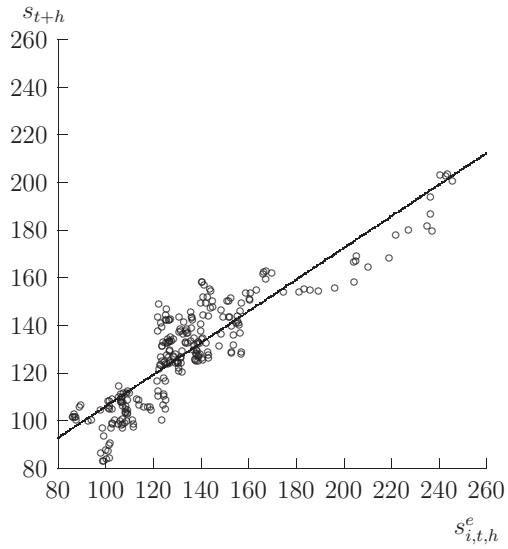

Figure 10 Actual vs Expected Exchange Rate: 6 Month-Ahead Forecast

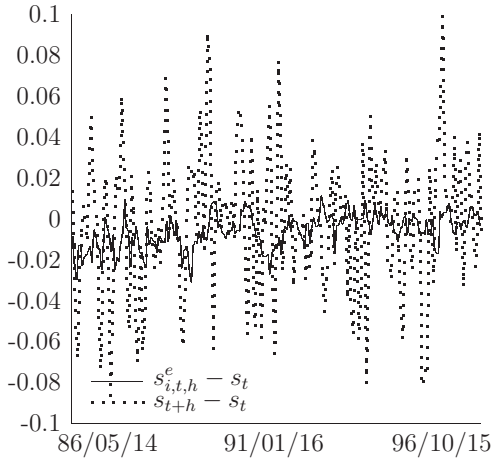

Figure 11 Actual & Expected Depreciation, 1 Month-Ahead Forecast

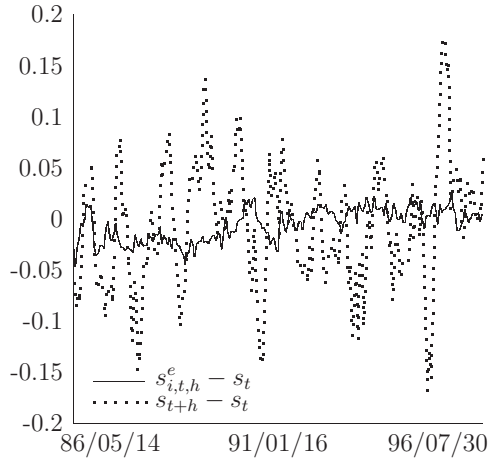

Figure 12 Actual & Expected Depreciation, 3 Month-Ahead Forecast

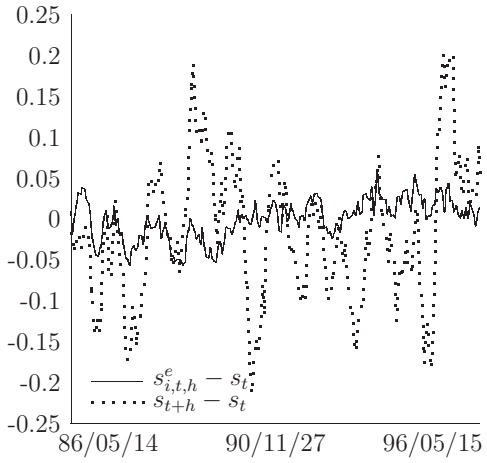

Figure 13 Actual & Expected Depreciation, 6 Month-Ahead Forecast

Table 1.1 Joint Unbiasedness Tests (1 month forecasts)

Individual regressions

$$s_{t+h} - s_t = \alpha_{i,h} + \beta_{i,h}(s^e_{i,t,h} - s_t) + \varepsilon_{i,t,h} \quad \text{for} \quad h = 2 \qquad (2)$$

degrees of freedom = 260

	i=1 Banks & Brokers	i=2 Insurance & Trading Cos.	i=3 Export Industries	i=4 Life Ins. & Import Cos.
$\alpha_{i,h}$	-0.003	-0.004	-0.007	-0.006
t (NW)	-1.123	-1.428	-2.732	-1.903
p-value	0.262	0.153	0.006	0.057
$\beta_{i,h}$	0.437	0.289	-0.318	0.008
t (NW)	1.674	1.382	-1.237	-0.038
R^2	0.014	0.008	0.007	0.000
$H_0 : \beta_{i,h} = 1, \text{for } i = 1, 2, 3, 4$				
χ^2	4.666	4.666	4.666	4.666
p-value	0.031	0.001	0.000	0.000
Unbiasedness Tests: $H_0 : \alpha_{i,h} = 0, \beta_{i,h} = 1$, for $i = 1, 2, 3, 4$				
χ^2(NW)	4.696	11.682	29.546	19.561
p-value	0.096	0.003	0.000	0.000
MH tests $H_0 : \alpha_{i,h} = \alpha_j, \beta_{i,h} = \beta_j$ for all $i, j \neq i$				
χ^2(GMM)	9.689			
p-value	0.138			

Table 1.2 Joint Unbiasedness Tests (3 month forecasts)

Individual regressions

$$s_{t+h} - s_t = \alpha_{i,h} + \beta_{i,h}(s^e_{i,t,h} - s_t) + \varepsilon_{i,t,h} \quad \text{for} \quad h = 6 \qquad (2)$$

degrees of freedom = 256

	i=1 Banks & Brokers	i=2 Insurance & Trading Cos.	i=3 Export Industries	i=4 Life Ins. & Import Cos.
$\alpha_{i,h}$	-0.013	-0.011	-0.017	-0.014
t (NW)	-1.537	-1.362	-2.060	-1.517
p-value	0.124	0.173	0.039	0.129
$\beta_{i,h}$	0.521	0.611	0.082	0.484
t (NW)	1.268	1.868	0.215	1.231
R^2	0.018	0.026	0.001	0.016
$H_0 : \beta_{i,h} = 1, \text{for } i = 1, 2, 3, 4$				
χ^2	1.362	1.415	5.822	1.728
p-value	0.243	0.234	0.016	0.189
Unbiasedness Tests: $H_0 : \alpha_{i,h} = 0, \beta_{i,h} = 1$, for $i = 1, 2, 3, 4$				
χ^2(NW)	2.946	2.691	11.156	3.023
p-value	0.229	0.260	0.004	0.221
MH tests $H_0 : \alpha_{i,h} = \alpha_j, \beta_{i,h} = \beta_j$ for all $i, j \neq i$				
χ^2(GMM)	5.783			
p-value	0.448			

Table 1.3 JOINT UNBIASEDNESS TESTS (6 MONTH FORECASTS)

Individual regressions

$$s_{t+h} - s_t = \alpha_{i,h} + \beta_{i,h}(s^e_{i,t,h} - s_t) + \varepsilon_{i,t,h} \quad \text{for} \quad h = 12 \qquad (2)$$

degrees of freedom = 256

	i=1 Banks & Brokers	i=2 Insurance & Trading Cos.	i=3 Export Industries	i=4 Life Ins. & Import Cos.
$\alpha_{i,h}$	-0.032	-0.032	-0.039	-0.034
t (NW)	-1.879	-1.831	-2.099	-1.957
p-value	0.060	0.067	0.036	0.050
$\beta_{i,h}$	0.413	0.822	0.460	0.399
t (NW)	0.761	1.529	0.911	-0.168
R^2	0.01	0.044	0.021	0.012
$H_0 : \beta_{i,h} = 1, \text{for } i = 1, 2, 3, 4$				
χ^2	1.166	0.110	1.147	1.564
p-value	0.280	0.740	0.284	0.211
Unbiasedness Tests: $H_0 : \alpha_{i,h} = 0, \beta_{i,h} = 1$, for $i = 1, 2, 3, 4$				
χ^2(NW)	4.332	3.5	7.899	5.006
p-value	0.115	0.174	0.019	0.082
MH tests $H_0 : \alpha_{i,h} = \alpha_j, \beta_{i,h} = \beta_j$ for all $i, j \neq i$				
χ^2(GMM)	7.071			
p-value	0.314			

Table 2.1 Unit Root Tests (1 Month Forecasts $h = 2$)

$$\Delta y_t = \alpha + \beta y_{t-1} + \gamma t + \sum_{k=1}^{p} \theta_t \Delta y_{t-k} + \epsilon_t \quad \text{for } h = 2 \tag{5}$$

	Lags	ADF t test	ADF z test	Joint test
Log of spot rate (n=276)	0	-2.828	-9.660	6.820**
H^{th} difference of log of spot rate	12	-4.306***	-167.473***	9.311***
	Group 1 Banks & Brokers			
Log of forecast	0	-2.274*	-5.072	6.327**
H^{th} difference of log of forecast	0	-7.752***	-96.639***	30.085***
Forecast error				
Restriced CI eq.	1	-11.325***	-249.389***	64.676***
Unrestriced CI eq.	1	65.581***		
	Group 2 Insurance & Trading Cos.			
Log of forecast	0	-2.735*	-5.149	6.705***
H^{th} difference of log of forecast	0	-7.895***	-94.986***	31.252***
Forecast error				
Restriced CI eq.	1	-11.624***	-270.302***	68.053***
Unrestriced CI eq.	1	-11.750***		
	Group 3 Export Industries			
Log of forecast	1	-2.372	-4.806	5.045**
H^{th} difference of log of forecast	0	-8.346***	-111.632***	34.889***
Forecast error				
Restriced CI eq.	1	-10.324***	-211.475***	53.757***
Unrestriced CI eq.	1	-10.392***		
	Group 4 Life Ins. & Import Cos.			
Log of forecast	0	-2.726*	-5.009	6.438**
H^{th} difference of log of forecast	1	-5.216***	-52.911***	3.630***
Forecast error				
Restriced CI eq.	1	-10.977***	-231.837***	60.820***
Unrestriced CI eq.	1	-10.979***		

* rejection at 10% level
** rejection at 5% level
*** rejection at 1% level

Table 2.2 Unit Root Tests (3 Month Forecasts $h = 6$)

$$\Delta y_t = \alpha + \beta y_{t-1} + \gamma t + \sum_{k=1}^{p} \theta_t \Delta y_{t-k} + \epsilon_t \quad \text{for } h = 6 \tag{5}$$

	Lags	ADF t test	ADF z test	Joint test
Log of spot rate (n=276)	0	-2.828	-9.660	6.820**
H^{th} difference of log of spot rate	2	-4.760***	-49.769***	11.351***
	Group 1 Banks & Brokers			
Log of forecast	0	-2.840*	-4.852	7.610***
H^{th} difference of log of forecast	0	-5.092***	-48.707***	12.990***
Forecast error				
Restriced CI eq.	6	-3.022**	-29.429***	4.673**
Unrestriced CI eq.	6	-3.343*		
	Group 2 Insurance & Trading Cos.			
Log of forecast	0	-2.778*	-4.533	8.858***
H^{th} difference of log of forecast	0	-6.514***	-71.931***	21.588***
Forecast error				
Restriced CI eq.	6	-3.068**	-31.038***	4.956**
Unrestriced CI eq.	2	-4.539***		
	Group 3 Export Industries			
Log of forecast	0	-3.105**	-4.549	9.090***
H^{th} difference of log of forecast	1	-4.677***	-41.524***	10.944***
Forecast error				
Restriced CI eq.	6	-3.317**	-31.207***	5.659**
Unrestriced CI eq.	5	-5.115***		
	Group 4 Life Ins. & Import Cos.			
Log of forecast	0	-2.863*	-4.400	8.161***
H^{th} difference of log of forecast	1	-4.324***	-39.870***	9.352***
Forecast error				
Restriced CI eq.	5	-4.825***	-118.586***	11.679***
Unrestriced CI eq.	4	-5.123***		

* rejection at 10% level
** rejection at 5% level
*** rejection at 1% level

Table 2.3 Unit Root Tests (6 Month Forecasts $h = 12$)

$$\Delta y_t = \alpha + \beta y_{t-1} + \gamma t + \sum_{k=1}^{p} \theta_t \Delta y_{t-k} + \epsilon_t \quad \text{for } h = 12 \tag{5}$$

	Lags	ADF t test	ADF z test	Joint test
Log of spot rate (n=276)	0	-2.828	-9.660	6.820**
H^{th} difference of log of spot rate	17	-3.189**	-26.210***	5.500**
	Group 1 Banks & Brokers			
Log of forecast	0	-2.947**	-4.254	9.131***
H^{th} difference of log of forecast	0	-4.772***	-44.018***	11.389***
Forecast error				
Restriced CI eq.	1	-2.373	-13.577*	2.947
Unrestriced CI eq.	7	-3.285**		
	Group 2 Insurance & Trading Cos.			
Log of forecast	0	-2.933**	-4.004	9.531***
H^{th} difference of log of forecast	0	-6.007***	-64.923***	18.044***
Forecast error				
Restriced CI eq.	1	-2.114	-11.464*	2.399
Unrestriced CI eq.	1	-2.684*		
	Group 3 Export Industries			
Log of forecast	0	-3.246**	-4.059	10.704***
H^{th} difference of log of forecast	12	-4.961***	-44.532***	12.331***
Forecast error				
Restriced CI eq.	12	-1.515	-5.601	1.466
Unrestriced CI eq.	0	-2.931		
	Group 4 Life Ins. & Import Cos.			
Log of forecast	0	-3.133**	-4.196	9.549***
H^{th} difference of log of forecast	0	-4.795***	-44.062***	11.537***
Forecast error				
Restriced CI eq.	2	-2.508	-14.535**	3.148
Unrestriced CI eq.	1	-2.851*		

* rejection at 10% level
** rejection at 5% level
*** rejection at 1% level

Table 3.1 Simple Unbiasedness Tests (1 month forecasts)

Individual regresions (h=2): $s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \varepsilon_{i,t,h}$ (7)
degrees of freedom = 261

	i=1 Banks & Brokers	i=2 Insurance & Trading Cos.	i=3 Export Industries	i=4 Life Ins. & Import Cos.
$\alpha_{i,h}$	0.000	0.001	-0.002	0.002
t (NW)	0.115	0.524	-0.809	0.720
p-value	0.909	0.600	0.418	0.472
MH tests $H_0 : \alpha_{i,h} = \alpha_j$, for all $i, j \neq i$				
χ^2(GMM)	41.643		p-value	0.000

Individual regresions (h=2): $s_{t+h} - s^e_{i,t,h} = \alpha_{i,h,1} + \alpha_{i,h,2} + \alpha_{i,h,3} + \alpha_{i,h,4} + \varepsilon_{i,t,h}$ (?)

1985:05:29 - 1988.03:16	i=1 Banks & Brokers	i=2 Insurance & Trading Cos.	i=3 Export Industries	i=4 Life Ins. & Import Cos.
$\alpha_{i,h,1}$	-0.007	-0.005	-0.013	-0.003
p-value	0.213	0.353	0.017	0.573
1988:03:30 - 1991:01:16				
$\alpha_{i,h,2}$	0.009	0.008	0.010	0.010
p-value	0.134	0.164	0.090	0.114
1991:01:29 - 1993:11:16				
$\alpha_{i,h,3}$	-0.002	-0.001	-0.004	-0.001
p-value	0.598	0.810	0.374	0.757
1993:11:30 - 1996:10:15				
$\alpha_{i,h,4}$	0.002	0.004	-0.001	0.003
p-value	0.675	0.433	0.814	0.519
Structrual Break tests $H_0 : \alpha_{i,h,1} = \alpha_{i,h,2} = \cdots = \alpha_{i,h,4}$				
χ^2	4.245	3.267	8.425	2.946
p-value	0.236	0.352	0.038	0.400

Table 3.2 Simple Unbiasedness Tests (3 month forecasts)

Individual regresions (h=6): $s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \varepsilon_{i,t,h}$ (7)				
degrees of freedom = 257				
	i=1 Banks & Brokers	i=2 Insurance & Trading Cos.	i=3 Export Industries	i=4 Life Ins. & Import Cos.
$\alpha_{i,h}$	-0.01	-0.008	-0.019	-0.01
t (NW)	-1.151	-0.929	-2.165	-1.121
p-value	0.25	0.353	0.03	0.262
MH tests $H_0 : \alpha_{i,h} = \alpha_j$, for all $i, j \neq i$				
χ^2(GMM)	40.16		p-value	0.000

Individual regresions (h=6): $s_{t+h} - s^e_{i,t,h} = \alpha_{i,h,1} + \alpha_{i,h,2} + \alpha_{i,h,3} + \alpha_{i,h,4} + \varepsilon_{i,t,h}$ (?)				
1985:05:29 - 1988.02:24	i=1 Banks & Brokers	i=2 Insurance & Trading Cos.	i=3 Export Industries	i=4 Life Ins. & Import Cos.
$\alpha_{i,h,1}$	-0.040	-0.039	-0.057	-0.039
p-value	0.005	0.007	0.000	0.008
1988:03:16 - 1990:12:11				
$\alpha_{i,h,2}$	0.023	0.025	0.022	0.021
p-value	0.169	0.111	0.187	0.229
1990:12:25 - 1993:09:28				
$\alpha_{i,h,3}$	-0.020	-0.018	-0.029	-0.020
p-value	0.064	0.125	0.023	0.094
1993:10:12 - 1996:0730				
$\alpha_{i,h,4}$	0.000	0.001	-0.013	0.001
p-value	0.994	0.941	0.466	0.970
Structrual Break tests $H_0 : \alpha_{i,h,1} = \alpha_{i,h,2} = \cdots = \alpha_{i,h,4}$				
χ^2	9.319	9.925	13.291	7.987
p-value	0.025	0.019	0.004	0.046

Table 4.1 Error Correction Models (1 Month Forecasts)

Group 1 Banks & Brokers				
$s_{t+h} - s_t = c_i + \alpha_i(s_t - \gamma_i s^e_{i,t-h,h}) + \beta_i(s^e_{i,t} - s^e_{i,t-h,h}) + \epsilon_{i,h}$				(9)
χ^2 tests of ECM with robust standard errors ($R^2 = 0.0195$)				
	Coeff	$\chi^2(n)$	n	p-value
constant	-0.002	0.377	1	0.539
$\alpha_{i,h} = 0$	-0.465	2.884	1	0.089
$\alpha_{i,h} = -1$	-0.465	3.813	1	0.051
$\beta_{i,h} = 1$	0.491	3.847	1	0.050
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$		3.910	3	0.271
F tests of ECM with whitened residuals ($R^2 = 0.605$)				
constant	-0.001	0.238	1	0.627
$\alpha_{i,h} = -1$	-0.025	14.696	1	0.000
$\beta_{i,h} = 1$	0.453	4.895	1	0.028
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$		6.790	3	0.000
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$ & coeffs on all lags of realizations & forecasts $= 0$		51.115	12	0.000
$s_{t+h} - s^e_{i,t,h} = d_i + (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h})$				(10)
χ^2 tests with robust standard errors and $\beta_{i,h} = 1$ imposed				
constant	0.002	0.293	1	0.582
$\alpha_{i,h} = -1$	-0.991	0.015	1	0.903
constant $= 0$ & $\alpha_{i,h} = -1$		0.294	2	0.863

Group 2 Insurance & Trading Cos.				
$s_{t+h} - s_t = c_i + \alpha_i(s_t - \gamma_i s^e_{i,t-h,h}) + \beta_i(s^e_{i,t} - s^e_{i,t-h,h}) + \epsilon_{i,h}$				(9)
χ^2 tests of ECM with robust standard errors ($R^2 = 0.009$)				
	Coeff	$\chi^2(n)$	n	p-value
constant	-0.003	1.168	1	0.280
$\alpha_{i,h} = 0$	-0.262	1.111	1	0.292
$\alpha_{i,h} = -1$	-0.262	8.807	1	0.003
$\beta_{i,h} = 1$	0.278	10.703	1	0.001
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$		10.965	3	0.012
F tests of ECM with whitened residuals ($R^2 = 0.596$)				
constant	-0.002	0.563	1	0.454
$\alpha_{i,h} = -1$	-0.036	13.614	1	0.000
$\beta_{i,h} = 1$	0.207	12.639	1	0.001
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$		6.792	3	0.000
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$ & coeffs on all lags of realizations & forecasts $= 0$		54.557	11	0.000
$s_{t+h} - s^e_{i,t,h} = d_i + (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h})$				(10)
χ^2 tests with robust standard errors and $\beta_{i,h} = 1$ imposed				
constant	0.003	0.787	1	0.3751
$\alpha_{i,h} = -1$	-1.026	0.113	1	0.7368
constant $= 0$ & $\alpha_{i,h} = -1$		1.052	2	0.591

Table 4.2 Error Correction Models (1 Month Forecasts)

Group 3 Export Industries				
$s_{t+h} - s_t = c_i + \alpha_i(s_t - \gamma_i s^e_{i,t-h,h}) + \beta_i(s^e_{i,t} - s^e_{i,t-h,h}) + \epsilon_{i,h}$				(9)
χ^2 tests of ECM with robust standard errors ($R^2 = 0.009$)				
	Coeff	$\chi^2(n)$	n	p-value
constant	-0.006	4.202	1	0.040
$\alpha_{i,h} = 0$	0.305	1.335	1	0.248
$\alpha_{i,h} = -1$	0.305	24.402	1	0.000
$\beta_{i,h} = 1$	-0.256	20.516	1	0.000
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$		27.207	3	0.000
F tests of ECM with whitened residuals ($R^2 = 0.602$)				
constant	-0.002	0.632	1	0.428
$\alpha_{i,h} = -1$	0.107	18.043	1	0.000
$\beta_{i,h} = 1$	-0.055	17.987	1	0.000
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$		8.455	3	0.000
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$ & coeffs on all lags of realizations & forecasts $= 0$		55.054	10	0.000
$s_{t+h} - s^e_{i,t,h} = d_i + (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h})$				(10)
χ^2 tests with robust standard errors and $\beta_{i,h} = 1$ imposed				
constant	-0.001	0.082	1	0.7749
$\alpha_{i,h} = -1$	-0.887	2.321	1	0.1277
constant $= 0$ & $\alpha_{i,h} = -1$		2.578	2	0.276

Group 4 Life Insurance & Import Cos.				
$s_{t+h} - s_t = c_i + \alpha_i(s_t - \gamma_i s^e_{i,t-h,h}) + \beta_i(s^e_{i,t} - s^e_{i,t-h,h}) + \epsilon_{i,h}$				(9)
χ^2 tests of ECM with robust standard errors ($R^2 = 0.003$)				
	Coeff	$\chi^2(n)$	n	p-value
constant	-0.004	1.734	1	0.188
$\alpha_{i,h} = 0$	-0.066	0.083	1	0.773
$\alpha_{i,h} = -1$	-0.066	16.501	1	0.000
$\beta_{i,h} = 1$	0.112	16.086	1	0.000
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$		17.071	3	0.001
F tests of ECM with whitened residuals ($R^2 = 0.607$)				
constant	-0.002	0.392	1	0.532
$\alpha_{i,h} = -1$	-0.026	20.268	1	0.000
$\beta_{i,h} = 1$	0.226	12.020	1	0.001
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$		10.794	3	0.000
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$ & coeffs on all lags of realizations & forecasts $= 0$		57.702	11	0.000
$s_{t+h} - s^e_{i,t,h} = d_i + (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h})$				(10)
χ^2 tests with robust standard errors and $\beta_{i,h} = 1$ imposed				
constant	0.003	1.254	1	0.2629
$\alpha_{i,h} = -1$	-0.949	0.481	1	0.4879
constant $= 0$ & $\alpha_{i,h} = -1$		1.628	2	0.443

Table 4.3 Error Correction Models (3 Month Forecasts)

Group 1 Banks & Brokers				
$s_{t+h} - s_t = c_i + \alpha_i(s_t - \gamma_i s^e_{i,t-h,h}) + \beta_i(s^e_{i,t} - s^e_{i,t-h,h}) + \epsilon_{i,h}$				(9)
χ^2 tests of ECM with robust standard errors ($R^2 = 0.036$)				
	Coeff	$\chi^2(n)$	n	p-value
constant	-0.010	1.306	1	0.253
$\alpha_{i,h} = 0$	-0.377	0.590	1	0.443
$\alpha_{i,h} = -1$	-0.377	1.604	1	0.205
$\beta_{i,h} = 1$	0.501	1.268	1	0.260
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$		2.348	3	0.503
F tests of ECM with whitened residuals ($R^2 = 0.863$)				
constant	-0.008	4.173	1	0.044
$\alpha_{i,h} = -1$	0.233	56.755	1	0.000
$\beta_{i,h} = 1$	-0.178	51.113	1	0.000
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$		28.974	3	0.000
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$ & coeffs on all lags of realizations & forecasts $= 0$		121.851	14	0.000
$s_{t+h} - s^e_{i,t,h} = d_i + (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h})$				(10)
χ^2 tests with robust standard errors and $\beta_{i,h} = 1$ imposed				
constant	-0.006	0.556	1	0.456
$\alpha_{i,h} = -1$	-0.889	0.896	1	0.344
constant $= 0$ & $\alpha_{i,h} = -1$		1.330	2	0.514

Group 2 Insurance & Trading Cos.				
$s_{t+h} - s_t = c_i + \alpha_i(s_t - \gamma_i s^e_{i,t-h,h}) + \beta_i(s^e_{i,t} - s^e_{i,t-h,h}) + \epsilon_{i,h}$				(9)
χ^2 tests of ECM with robust standard errors ($R^2 = 0.044$)				
	Coeff	$\chi^2(n)$	n	p-value
constant	-0.008	0.874	1	0.350
$\alpha_{i,h} = 0$	-0.556	2.061	1	0.151
$\alpha_{i,h} = -1$	-0.556	1.310	1	0.252
$\beta_{i,h} = 1$	0.663	0.965	1	0.326
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$		1.833	3	0.608
F tests of ECM with whitened residuals ($R^2 = 0.844$)				
constant	-0.005	1.400	1	0.239
$\alpha_{i,h} = -1$	0.080	31.425	1	0.000
$\beta_{i,h} = 1$	-0.167	40.346	1	0.000
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$		23.551	3	0.000
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$ & coeffs on all lags of realizations & forecasts $= 0$		148.338	10	0.000
$s_{t+h} - s^e_{i,t,h} = d_i + (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h})$				(10)
χ^2 tests with robust standard errors and $\beta_{i,h} = 1$ imposed				
constant	-0.005	0.291	1	0.589
$\alpha_{i,h} = -1$	-0.897	0.773	1	0.379
constant $= 0$ & $\alpha_{i,h} = -1$		0.945	2	0.623

Table 4.4 Error Correction Models (3 Month Forecasts)

Group 3 Export Industries				
$s_{t+h} - s_{i,t,h} = c_i + \alpha_i(s_t - \gamma_i s^e_{i,t-h,h}) + \beta_i(s^e_{i,t} - s^e_{i,t-h,h}) + \epsilon_{i,h}$				
χ^2 tests of ECM with robust standard errors ($R^2 = 0.026$)				
	Coeff	$\chi^2(n)$	n	p-value
constant	-0.013	2.303	0	0.129
$\alpha_{i,h} = 0$	-0.253	0.393	1	0.531
$\alpha_{i,h} = -1$	-0.253	3.422	1	0.064
$\beta_{i,h} = 1$	0.411	2.102	1	0.147
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$		7.663	3	0.054
F tests of ECM with whitened residuals ($R^2 = 0.856$)				
constant	-0.003	0.840	1	0.361
$\alpha_{i,h} = -1$	-0.006	29.512	1	0.000
$\beta_{i,h} = 1$	-0.205	40.582	1	0.000
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$		16.290	3	0.000
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$ & coeffs on all lags of realizations & forecasts $= 0$		182.912	10	0.000
$s_{t+h} - s^e_{i,t,h} = d_i + (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h})$				(10)
χ^2 tests with robust standard errors and $\beta_{i,h} = 1$ imposed				
constant	-0.012	1.971	1	0.160
$\alpha_{i,h} = -1$	-0.775	3.791	1	0.052
constant $= 0$ & $\alpha_{i,h} = -1$		6.337	2	0.042

Group 4 Life Insurance & Import Cos.				
$s_{t+h} - s_{i,t,h} = c_i + \alpha_i(s_t - \gamma_i s^e_{i,t-h,h}) + \beta_i(s^e_{i,t} - s^e_{i,t-h,h}) + \epsilon_{i,h}$				
χ^2 tests of ECM with robust standard errors ($R^2 = 0.038$)				
	Coeff	$\chi^2(n)$	n	p-value
constant	-0.009	0.993	1	0.319
$\alpha_{i,h} = 0$	-0.478	1.250	1	0.264
$\alpha_{i,h} = -1$	-0.478	1.488	1	0.223
$\beta_{i,h} = 1$	0.604	0.919	1	0.338
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$		2.451	3	0.484
F tests of ECM with whitened residuals ($R^2 = 0.845$)				
constant	-0.003	0.510	1	0.477
$\alpha_{i,h} = -1$	0.050	32.000	1	0.000
$\beta_{i,h} = 1$	0.062	32.469	1	0.000
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$		21.673	3	0.000
constant $= 0$ & $\alpha_{i,h} = -1$ & $\beta_{i,h} = 1$ & coeffs on all lags of realizations & forecasts $= 0$		169.286	9	0.000
$s_{t+h} - s^e_{i,t,h} = d_i + (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h})$				(10)
χ^2 tests with robust standard errors and $\beta_{i,h} = 1$ imposed				
constant	-0.006	0.455	1	0.500
$\alpha_{i,h} = -1$	-0.865	1.405	1	0.236
constant $= 0$ & $\alpha_{i,h} = -1$		1.726	2	0.422

Table 5.1 Weak Efficiency Tests (1 Month Forecasts)

$$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{p=h}^{h+6} \beta_{i,t+h-p}(s_{t+h-p} - s^e_{m,t+h-p,h}) + \epsilon_{i,t,h} \quad \text{for} \quad h = 2 \qquad (11)$$

	i = 1 Banks & Brokers		i = 2 Insurance & Trading Cos.		i = 3 Export Industry		i = 4 Life Ins. & Import Cos.	
Lags	χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value
Single								
2	0.132	0.895	0.139	0.709	1.871	0.171	0.334	0.563
3	-0.914	0.361	1.971	0.160	0.027	0.869	0.186	0.667
4	0.160	0.689	0.006	0.938	1.634	0.201	0.714	0.398
5	0.450	0.502	0.050	0.823	1.749	0.186	1.180	0.277
6	0.046	0.831	0.104	0.747	0.686	0.408	0.188	0.665
7	0.002	0.967	0.282	0.595	0.069	0.793	0.001	0.970
8	0.091	0.763	0.436	0.509	0.022	0.883	0.300	0.584
Cum.								
3	0.765	0.682	1.778	0.411	1.746	0.418	0.585	0.746
4	4.626	0.201	3.463	0.326	8.763	0.033	5.349	0.148
5	4.747	0.314	4.382	0.357	7.680	0.104	5.081	0.279
6	5.501	0.358	5.592	0.348	7.652	0.176	5.768	0.329
7	6.252	0.396	6.065	0.416	8.879	0.180	6.677	0.352
8	5.927	0.548	5.357	0.617	8.390	0.299	6.087	0.530

Selected micro-homogeneity tests
$H_0 : \alpha_{i,h} = \alpha_{j,h}, \beta_{i,t+h-p} = \beta_{j,t+h-p}$ for all $i, j \neq i$

	$\chi^2(n)$	p-value	n
Single			
2	122.522	0.000	6
8	43.338	0.000	6
Cum.			
3	136.830	0.000	9
8	201.935	0.000	24

VCV matrix incorporates Newey-West correction for serial correlation
χ^2 statistics for mean forecast error regressions (p-value underneath)
Degrees of freedom (n) represent number of regressors, excluding intercept
(n=1 for single lag, n=max. lag -2 for cumulative lags)

Table 5.2 Weak Efficiency Tests (1 Month Forecasts)

$$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{p=h}^{h+6} \beta_{i,t+h-p}(s^e_{m,t+h-p,h} - s_{t-p}) + \epsilon_{i,t,h} \quad \text{for} \quad h = 2 \qquad (12)$$

	i = 1 Banks & Brokers		i = 2 Insurance & Trading Cos.		i = 3 Export Industry		i = 4 Life Ins. & Import Cos.	
Lags	χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value
Single								
2	-2.325	0.020	5.641	0.018	3.658	0.056	7.011	0.008
3	4.482	0.106	3.519	0.061	3.379	0.066	5.877	0.015
4	3.162	0.075	2.580	0.108	2.805	0.094	4.911	0.027
5	3.956	0.047	2.993	0.084	3.102	0.078	7.467	0.006
6	6.368	0.012	4.830	0.028	5.952	0.015	9.766	0.002
7	8.769	0.003	6.786	0.009	7.755	0.005	12.502	0.000
8	5.451	0.020	4.114	0.043	4.417	0.036	7.564	0.006
Cum.								
3	5.592	0.061	6.138	0.046	4.116	0.128	7.508	0.023
4	5.638	0.131	5.896	0.117	4.283	0.232	7.888	0.048
5	5.189	0.268	4.964	0.291	3.784	0.436	8.009	0.091
6	6.025	0.304	5.068	0.408	4.847	0.435	8.401	0.136
7	7.044	0.317	5.746	0.452	5.940	0.430	9.434	0.151
8	10.093	0.183	8.494	0.291	7.919	0.340	12.530	0.084

Selected micro-homogeneity tests
$H_0 : \alpha_{i,h} = \alpha_{j,h}, \beta_{i,t+h-p} = \beta_{j,t+h-p}$ for all $i, j \neq i$

	$\chi^2(n)$	p-value	n
Single			
2	40.462	0.000	6
8	30.739	0.000	6
Cum.			
3	42.047	0.000	6
8	46.124	0.004	24

VCV matrix incorporates Newey-West correction for serial correlation
χ^2 statistics for mean forecast error regressions (p-value underneath)
Degrees of freedom (n) represent number of regressors, excluding intercept
(n=1 for single lag, n=max. lag -2 for cumulative lags)

Table 5.3 Weak Efficiency Tests (1 Month Forecasts)

$$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{p=h}^{h+6} \beta_{i,t+h-p}(s_{t+h-p} - s_{t-p}) + \epsilon_{i,t,h} \quad \text{for} \quad h = 2 \tag{13}$$

	i = 1 Banks & Brokers		i = 2 Insurance & Trading Cos.		i = 3 Export Industry		i = 4 Life Ins. & Import Cos.	
Lags	χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value
Single								
2	-0.328	0.743	0.639	0.424	1.249	0.264	0.023	0.879
3	1.621	0.203	3.060	0.080	0.000	0.993	0.550	0.458
4	0.002	0.964	0.335	0.562	0.819	0.366	0.146	0.702
5	0.086	0.770	0.042	0.837	1.001	0.317	0.344	0.557
6	0.165	0.685	0.916	0.339	0.029	0.864	0.095	0.758
7	0.850	0.357	1.861	0.172	0.329	0.566	1.152	0.283
8	0.597	0.440	1.088	0.297	0.317	0.574	1.280	0.258
Cum.								
3	1.978	0.372	3.169	0.205	1.940	0.379	1.132	0.568
4	3.304	0.347	3.501	0.321	5.567	0.135	3.318	0.345
5	3.781	0.436	4.248	0.373	5.806	0.214	3.598	0.463
6	3.651	0.601	4.646	0.461	5.756	0.331	3.819	0.576
7	4.493	0.610	5.609	0.468	6.608	0.359	5.040	0.539
8	5.619	0.585	6.907	0.439	7.907	0.341	6.521	0.480

Selected micro-homogeneity tests
$H_0 : \alpha_{i,h} = \alpha_{j,h}, \beta_{i,t+h-p} = \beta_{j,t+h-p}$ for all $i, j \neq i$

	$\chi^2(n)$	p-value	n
Single			
2	150.698	0.000	6
8	45.652	0.000	6
Cum.			
3	161.950	0.000	9
8	214.970	0.000	24

VCV matrix incorporates Newey-West correction for serial correlation
χ^2 statistics for mean forecast error regressions (p-value underneath)
Degrees of freedom (n) represent number of regressors, excluding intercept
(n=1 for single lag, n=max. lag -2 for cumulative lags)

Table 5.4 LM Test for Serial Correlation (1 Month Forecasts)

$H_0 : \beta_{i,t-h} = ... = \beta_{i,t-h-6} = 0,$ for $h = 2$
in $\hat{\epsilon}_{i,t,h} = \alpha_{i,h} + \sum_{k=1}^{h-1} \beta_{i,k}(s_{t+h-k} - s^e_{i,t-k,h}) + \sum_{l=h}^{h+6} \phi_{i,l}\hat{\epsilon}_{i,t-l,h} + \eta_{i,t,h},$
where ϵ is generated from
$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{k=1}^{h-1} \beta_{i,k}(s_{t+h-k} - s^e_{i,t-k,h}) + \epsilon_{i,t,h}$

Cum. lags (k)	2	3	4	5	6	7	8
$n-k$	219	205	192	179	166	153	144
			i = 1				
			Banks & Brokers				
$F(k, n-k)$	29.415	18.339	14.264	11.180	9.699	7.922	6.640
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
			i = 2				
			Insurance & Trading Cos.				
$F(k, n-k)$	30.952	19.506	15.372	11.661	9.695	8.120	7.050
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
			i = 3				
			Export Industries				
$F(k, n-k)$	32.387	20.691	16.053	12.951	10.628	9.418	7.520
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
			i = 4				
			Life Ins. & Import Cos.				
$F(k, n-k)$	29.694	18.606	14.596	11.093	9.586	9.154	7.937
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 5.5 Weak Efficiency Tests (3 Month Forecasts)

$$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{p=h}^{h+6} \beta_{i,t+h-p}(s_{t+h-p} - s^e_{m,t+h-p,h}) + \epsilon_{i,t,h} \quad \text{for} \quad h = 6 \qquad (11)$$

	i = 1 Banks & Brokers		i = 2 Insurance & Trading Cos.		i = 3 Export Industry		i = 4 Life Ins. & Import Cos.	
Lags	χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value
Single								
6	0.667	0.414	0.954	0.329	4.493	0.034	1.719	0.190
7	0.052	0.820	0.071	0.789	1.434	0.231	0.268	0.605
8	0.006	0.940	0.010	0.921	0.382	0.537	0.001	0.976
9	0.055	0.814	0.043	0.836	0.140	0.708	0.060	0.806
10	0.264	0.607	0.278	0.598	0.001	0.980	0.432	0.511
11	0.299	0.585	0.381	0.537	0.020	0.888	0.598	0.439
12	0.172	0.678	0.336	0.562	0.011	0.918	0.633	0.426
Cum.								
7	8.966	0.011	11.915	0.003	19.663	0.000	12.350	0.002
8	12.288	0.006	16.263	0.001	23.290	0.000	15.146	0.002
9	11.496	0.022	15.528	0.004	22.417	0.000	14.778	0.005
10	8.382	0.136	12.136	0.033	16.839	0.005	12.014	0.035
11	11.596	0.072	18.128	0.006	23.782	0.001	15.330	0.032
12	11.527	0.117	15.983	0.025	21.626	0.003	13.038	0.071

Selected micro-homogeneity tests
$H_0 : \alpha_{i,h} = \alpha_{j,h}, \beta_{i,t+h-p} = \beta_{j,t+h-p}$ for all $i, j \neq i$

	$\chi^2(n)$	p-value	n
Single			
6	188.738	0.000	6
12	63.364	0.000	6
Cum.			
7	217.574	0.000	9
12	229.567	0.000	24

VCV matrix incorporates Newey-West correction for serial correlation
χ^2 statistics for mean forecast error regressions (p-value underneath)
Degrees of freedom (n) represent number of regressors, excluding intercept
(n=1 for single lag, n=max. lag -2 for cumulative lags)

Table 5.6 Weak Efficiency Tests (3 Month Forecasts)

$$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{p=h}^{h+6} \beta_{i,t+h-p}(s^e_{m,t+h-p,h} - s_{t-p}) + \epsilon_{i,t,h} \quad \text{for} \quad h = 6 \qquad (12)$$

	i = 1 Banks & Brokers		i = 2 Insurance & Trading Cos.		i = 3 Export Industry		i = 4 Life Ins. & Import Cos.	
Lags	χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value
Single								
6	3.457	0.063	2.947	0.086	3.470	0.062	3.681	0.055
7	4.241	0.039	3.834	0.050	4.390	0.036	4.370	0.037
8	5.748	0.017	5.177	0.023	5.410	0.020	6.053	0.014
9	6.073	0.014	5.843	0.016	5.968	0.015	6.474	0.011
10	8.128	0.004	7.868	0.005	7.845	0.005	8.521	0.004
11	8.511	0.004	8.004	0.005	8.308	0.004	8.429	0.004
12	6.275	0.012	6.691	0.010	6.635	0.010	6.079	0.014
Cum.								
7	4.717	0.095	4.985	0.083	4.954	0.084	4.928	0.085
8	5.733	0.125	5.209	0.157	5.045	0.168	6.736	0.081
9	5.195	0.268	5.411	0.248	5.112	0.276	6.053	0.195
10	7.333	0.197	9.245	0.100	9.456	0.092	7.872	0.163
11	8.539	0.201	6.658	0.354	7.488	0.278	7.955	0.241
12	8.758	0.271	6.747	0.456	7.796	0.351	8.698	0.275

Selected micro-homogeneity tests
$H_0 : \alpha_{i,h} = \alpha_{j,h}, \beta_{i,t+h-p} = \beta_{j,t+h-p}$ for all $i, j \neq i$

	$\chi^2(n)$	p-value	n
Single			
6	57.130	0.000	6
12	58.230	0.000	6
Cum.			
7	63.917	0.000	9
12	126.560	0.000	24

VCV matrix incorporates Newey-West correction for serial correlation
χ^2 statistics for mean forecast error regressions (p-value underneath)
Degrees of freedom (n) represent number of regressors, excluding intercept
(n=1 for single lag, n=max. lag -2 for cumulative lags)

Table 5.7 Weak Efficiency Tests (3 Month Forecasts)

$$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{p=h}^{h+6} \beta_{i,t+h-p}(s_{t+h-p} - s_{t-p}) + \epsilon_{i,t,h} \quad \text{for} \quad h = 6 \tag{13}$$

	i = 1 Banks & Brokers		i = 2 Insurance & Trading Cos.		i = 3 Export Industry		i = 4 Life Ins. & Import Cos.	
Lags	χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value
Single								
6	0.268	0.604	0.450	0.502	3.657	0.056	1.065	0.302
7	0.055	0.814	0.037	0.848	0.599	0.439	0.003	0.957
8	0.331	0.565	0.305	0.581	0.029	0.864	0.230	0.632
9	0.513	0.474	0.482	0.488	0.022	0.883	0.577	0.448
10	1.038	0.308	1.077	0.299	0.318	0.573	1.344	0.246
11	1.335	0.248	1.532	0.216	0.563	0.453	1.872	0.171
12	1.184	0.276	1.620	0.203	0.616	0.433	1.979	0.159
Cum.								
7	6.766	0.034	8.767	0.012	15.683	0.000	10.052	0.007
8	8.752	0.033	11.784	0.008	18.330	0.000	11.162	0.011
9	8.654	0.070	11.588	0.021	18.929	0.001	11.309	0.023
10	9.421	0.093	12.890	0.024	19.146	0.002	12.275	0.031
11	9.972	0.126	13.137	0.041	19.597	0.003	13.003	0.043
12	8.581	0.284	11.823	0.107	17.670	0.014	11.431	0.121

Selected micro-homogeneity tests
$H_0 : \alpha_{i,h} = \alpha_{j,h}, \beta_{i,t+h-p} = \beta_{j,t+h-p}$ for all $i, j \neq i$

	$\chi^2(n)$	p-value	n
Single			
6	151.889	0.000	6
12	66.313	0.000	6
Cum.			
7	164.216	0.000	9
12	193.021	0.000	24

VCV matrix incorporates Newey-West correction for serial correlation
χ^2 statistics for mean forecast error regressions (p-value underneath)
Degrees of freedom (n) represent number of regressors, excluding intercept
(n=1 for single lag, n=max. lag -2 for cumulative lags)

Table 5.8 LM Test for Serial Correlation (3 Month Forecasts)

$H_0 : \beta_{i,t-h} = ... = \beta_{i,t-h-6} = 0,$ for $h = 6$
in $\hat{\epsilon}_{i,t,h} = \alpha_{i,h} + \sum_{k=1}^{h-1} \beta_{i,k}(s_{t+h-k} - s^e_{i,t-k,h}) + \sum_{l=h}^{h+6} \phi_{i,l}\hat{\epsilon}_{i,t-l,h} + \eta_{i,t,h},$
where ϵ is generated from
$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{k=1}^{h-1} \beta_{i,k}(s_{t+h-k} - s^e_{i,t-k,h}) + \epsilon_{i,t,h}$

Cum. lags (k)	6	7	8	9	10	11	12
$n-k$	126	117	108	99	94	89	84
			i = 1				
			Banks & Brokers				
$F(k, n-k)$	3.452	2.856	3.023	2.951	2.599	2.652	2.921
p-value	0.003	0.009	0.004	0.004	0.008	0.006	0.002
			i = 2				
			Insurance & Trading Cos.				
$F(k, n-k)$	3.499	2.850	3.408	2.907	2.492	2.584	2.341
p-value	0.003	0.009	0.002	0.004	0.011	0.007	0.012
			i = 3				
			Export Industries				
$F(k, n-k)$	4.687	3.956	4.409	3.572	2.928	2.819	2.605
p-value	0.000	0.001	0.000	0.001	0.003	0.003	0.005
			i = 4				
			Life Ins. & Import Cos.				
$F(k, n-k)$	2.352	2.482	2.501	2.168	1.866	1.794	1.811
p-value	0.035	0.021	0.016	0.031	0.060	0.067	0.059

Table 6.0 CD Tests for Contemporaneous Cross-Sectional Dependendence of Forecast Errors

$s_{t+h} - s^e_{i,t,h} = \alpha_{i.h} + \epsilon_{i,t,h}$ (7)

Lag length	CD	p-value
1 month horizon		
0	37.945	0.000
$h-1=1$	3.633	0.000
2	-4.261	0.000
3	2.063	0.039
4	-2.233	0.026
3 month horizon		
0	31.272	0.000
1	2.461	0.014
2	0.387	0.699
3	2.322	0.020
4	1.594	0.111
$h-1=5$	1.461	0.144
6	-5.887	0.000
7	0.340	0.734
8	1.456	0.145

$CD = \sqrt{\frac{2T}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij} \overset{a}{\sim} N(0,1)$
$N = 24$, $T = 276$, $\hat{\rho}_{ij}$ is the sample correlation coefficient between forecasters i and j, $i \neq j$.

Table 7.1 Ito tests (1 month forecasts)

Individual regresions
$s^e_{i,t,h} - s^e_{m,t,h} = (\alpha_{i,h} - \alpha_m) + (\varepsilon_{i,t,h} - \varepsilon_{m,t})$ for $h = 2$
degrees of freedom = 263

	i=1 Banks & Brokers	i=2 Insurance & Trading Cos.	i=3 Export Industries	i=4 Life Ins. & Import Cos.
$\alpha_{i,h}$	0	-0.001	0.003	-0.002
t (NW)	0.173	-2.316	5.471	-3.965
p-value	0.863	0.021	0	0

MH tests $H_0 : \alpha_{i,h} = \alpha_j, \text{for all} i, j \neq I$

χ^2(GMM)	40.946
MSL(GMM)	0

Table 7.2 Ito tests (3 month forecasts)

Individual regresions
$s^e_{i,t,h} - s^e_{m,t,h} = (\alpha_{i,h} - \alpha_m) + (\varepsilon_{i,t,h} - \varepsilon_{m,t})$ for $h = 6$
degrees of freedom = 263

	i=1 Banks & Brokers	i=2 Insurance & Trading Cos.	i=3 Export Industries	i=4 Life Ins. & Import Cos.
$\alpha_{i,h}$	-0.002	-0.003	0.008	-0.002
t (NW)	-2.307	-3.986	5.903	-1.883
p-value	0.021	0	0	0.06

MH tests $H_0 : \alpha_{i,h} = \alpha_j, \text{for all} i, j \neq I$

χ^2(GMM)	37.704
MSL(GMM)	0

Table 7.3 Ito tests (6 month forecasts)

Individual regresions
$s^e_{i,t,h} - s^e_{m,t,h} = (\alpha_{i,h} - \alpha_m) + (\varepsilon_{i,t,h} - \varepsilon_{m,t})$ for $h = 12$
degrees of freedom = 263

	i=1 Banks & Brokers	i=2 Insurance & Trading Cos.	i=3 Export Industries	i=4 Life Ins. & Import Cos.
$\alpha_{i,h}$	-0.004	-0.003	0.01	0
t (NW)	-3.52	-2.34	4.549	-0.392
p-value	0	0.019	0	0.695

MH tests $H_0 : \alpha_{i,h} = \alpha_j, \text{for all} i, j \neq I$

χ^2(GMM)	23.402
p-value	0.001

References

A. Banerjee, J. J. Dolado, J. W. Galbraith, and D. F. Hendry. *Co-integration, Error Correction and the Econometric Analysis of Non-Stationary Data*. Oxford: Oxford University Press, 1993.

Roy Batchelor and Pami Dua. Forecaster ideology, forecasting technique, and the accuracy of economic forecasts. *International Journal of Forecasting*, 6:3–10, 1990a.

Roy Batchelor and Pami Dua. Product differentiation in the economic forecasting industry. *International Journal of Forecasting*, 6:311–316, 1990b.

Roy Batchelor and Pami Dua. Conservatism and consensus-seeking among economic forecasters. *Journal of Forecasting*, 11:169–181, 1992.

Roy Batchelor and David A. Peel. Rationality testing under asymmetric loss. *Economics Letters*, 61:49–54, 1998.

Carl S. Bonham and Richard H. Cohen. To aggregate, pool, or neither: Testing the rational-expectations hypothesis using survey data. *Journal of Business & Economic Statstics*, 19 (3):278–291, July 2001.

Carl S. Bonham and Richard H. Cohen. Testing the rationality of price forecasts: Comment. *American Economic Review*, 85(1):284–289, March 1995.

Paul Boothe and Debra Glassman. Comparing exchange rate forecasting models. *International Journal of Forecasting*, 3:65–79, 1987.

T. S. Breusch. Testing for auitocorrelation in dynamic linear models. *Australian Economic Papers*, 17:334–355, 1978.

T. S. Breusch and A. R. Pagan. The lagrange multiplier test and its application to model specifications in econometrics. *Review of Economic Studies*, 47:239–253, 1980.

Ralph C. Bryant. The "exchange risk premium," uncovered interest parity, and the treatment of exchange rates in multicountry macroeconomic models. Technical report, The Brookings Institution, January 1995.

Stefano M. F. G. Cavaglia, Willem F. C. Verschoor, and Christian C. P. Wolff. On the biasedness of forward foreign exchange rates: Irrationality or risk premia? *Journal of Business*, 67(3):321–343, July 1994.

Dionysios Chionis and Ronald MacDonald. Some tests of market microstructure hypotheses in the foreign exchange market. *Journal of Multinational Financial Management*, 7:203–229, 1997.

Richard H. Cohen and Carl S. Bonham. Specifying the forecast generating process for exchange rate survey forecasts. Working paper, Department of Economics, University of Hawaii Manoa, June 2006.

A. Davies and K. Lahiri. A new framework for testing rationality and measuring aggregate shocks using panel data. *Journal of Econometrics*, 68:205–227, 1995.

D. A. Dickey and W. A. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74:427–31, 1979.

Tilman Ehrbeck and Robert Waldmann. Why are professional forecasters biased? agency versus behavioral explanations. *The Quarterly Journal of Economics*, 111(1):21–40, February 1996.

Graham Elliott and Takatoshi Ito. Heterogeneous expectations and tests of efficiency in the yen/dollar forward exchange rate market. *Journal of Monetary Economics*, 43:435–456, 1999.

Robert Engle and Clive W. S. Granger. Cointegration and error correction representation, estimation and testing. *Econometrica*, 55:251–276, March 1987.

Stephen Figlewski. Market "efficiency" in a market with heterogeneous information. *Journal of Political Economy*, 86(4):581–597, August 1978.

Stephen Figlewski. Information diversity and market behavior. *Journal of Finance*, 37(1): 87–102, March 1982.

Stephen Figlewski. Information diversity and market behavior: A reply. *Journal of Finance*, 39(1):299–302, March 1984.

Stephen Figlewski and Paul Wachtel. Rational expectations, informational efficiency, and tests using survey data: A reply. *Review of Economics and Statistics*, 65:529–531, 1983.

Jeffey A. Frankel and Kenneth A. Froot. Chartists, fundamentalists, and trading in the foreign exchange market. *American Economic Review*, 80(2):181–185, May 1990.

Jeffrey A. Frankel and Kenneth A. Froot. Using survey data to test standard propositions regarding exchange rate expectations. *American Economic Review*, 77(1):133–153, March 1987.

Kenneth A. Froot and Jeffrey A. Frankel. Forward discount bias: Is it an exchange risk premium? *Quarterly Journal of Economics*, 104(1):139–161, February 1989.

William T. Gavin. FOMC forecasts: Is all the information in the central tendency? *Economic Review, Federal Reserve Bank of St. Louis*, pages 27–46, May/June 2003.

L. G. Godfrey. Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables. *Econometrica*, 46:1303–1310, 1978.

Michael D. Goldberg and Roman Frydman. Imperfect knowledge and behaviour in the foreign exchange market. *Economic Journal*, 106(437):869–893, July 1996.

Clyde W. J. Granger. Developments in the study of cointegrated economic variables. In Robert F. Engle and C. W. J. Granger, editors, *Long Run Economic Relationships*, pages 65–80. Oxford University Press, 1991.

Craig S. Hakkio and Mark Rush. Market efficiency and cointegration: an application to the sterling deutschemark exchange markets. *Journal of International Money and Finance*, 8: 75–88, 1989.

John C. Haltiwanger and Michael Waldman. Rational expectations in the aggregate. *Economic Inquiry*, XXVII:619–636, 1989.

J. Hamilton. *Time Series Analysis*. Princeton: Princeton University Press, 1994.

Lars Peter Hansen and Robert J. Hodrick. Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. *Journal of Political Economy*, 88(5):829–853, 1980.

K. Holden and D. A. Peel. On testing for unbiasedness and efficiency of forecasts. *Manchester School*, LVIII(2):120–127, June 1990.

Takatoshi Ito. Foreign exchange rate expectations: Micro survey data. *American Economic Review*, 80(3):434–449, June 1990.

Takatoshi Ito. Short-run and long-run expectations of the yen/dollar exchange rate. *Journal of the Japanese and International Economies*, 8:119–143, 1994.

Michael P. Keane and David E. Runkle. Testing the rationality of price forecasts: New evidence from panel data. *American Economic Review*, 80(4):714–735, September 1990.

Alan P. Kirman. Whom or what does the representative individual represent? *Journal of Economic Perspectives*, 6(2):117–136, Spring 1992.

J. F. Kiviet. *Testing Linear Econometric Models*. University of Amsterdam, 1987.

Kon Lai. An evaluation of survey exchange rate forecasts. *Economics Letters*, 32:61–65, 1990.

Owen Lamont. Macroeconomic forecass and microeconomic forecasters. *Journal of Economic Behavior and Organization*, 48(3):265–280, 2002.

David Laster, Paul Bennett, and In Sun Geoum. Rational bias in macroeconomic forecasts. *Quarterly Journal of Economics*, pages 293–318, February 1999.

Blake LeBaron. Technical trading profitability in foreign exchange markets in the 1990's. Technical report, Brandeis U., July 2000.

Gordon Leitch and J. Ernest Tanner. Economic forecast evaluation: Profits versus the conventional error measures. *The American Economic Review*, 81(3):580–590, June 1991.

Peter Liu and G. S. Maddala. Rationality of survey data and tests for market efficiency in the foreign exchange markets. *Journal of International Money and Finance*, 11:366–381, 1992.

Richard K. Lyons. Foreign exchange: Macro puzzles, micro tools. *Economic Review, Federal Reserve Bank of San Francisco*, 2002.

Ronald MacDonald. Exchange rate survey data: A disaggregated G-7 perspective. *The Manchester School Vol LX Supplement*, pages 47–62, June 1992.

Alex Maynard and Peter C. B. Phillips. Rethinking and old empirical puzzle: Econometric evidence on the forward discount anomaly. *Journal of Applied Econometrics*, 16(6):671–708, 2001.

J. Mincer and V. Zarnowitz. The evaluation of economic forecasts. In *Economic Forecasts and Expectations*. New York: National Bureau of Economic Research, 1969.

John Muth. Rational expectations and the theory of price movements. *Econometrica*, 29: 315–335, 1961.

W. Newey and K. West. A simple positive semi-definite heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55:703–708, May 1987.

William P. Osterberg. New results on the rationality of survey measures of exchange-rate expectations. *Federal Reserve Bank of Cleveland Economic Reveiw*, 36(1):14–21, 2000.

Hashem M. Pesaran. General diagnostic tests for cross section dependence in panels. URL `http://www.econ.cam.ac.uk/faculty/pesaran/wp.htm#gdt`. no note, June 2004.

Hashem M. Pesaran and Martin Weale. Survey expectations. In G. Elliott C. W. J. Granger and A. Timmermann, editors, *Handbook of Economic Forecasting*. North-Holland, 2006 (forthcoming).

Peter Phillips and Bruce E. Hansen. Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies*, 57:99–125, 1990.

Peter C. B. Phillips and Alex Maynard. Rethinking an old empirical puzzle: Econometric evidence on the forward discount anomaly. *Journal of Applied Econometrics*, 16(6):671–708, 2001.

Keith Pilbeam. Exchange rate models and exchange rate expectations: an empirical investigation. *Applied Economics*, 27:1009–1015, 1995.

Pentti Saikkonen. Asymptotically efficient estimation of cointegration regressions. *Econometric Theory*, 7:1–21, 1991.

Alan C. Stockman. Economic theory and exchange rate forecasts. *International Journal of Forecasting*, 3:3–15, 1987.

Henri Theil. *Applied Economic Forecasting*. North-Holland, Amsterdam, 1966.

Nikolaos Zacharatos and Charles Sutcliffe. Is the forward rate for the greek drachma uni-based? a VECM analysis with both overlapping and non-overlapping data. *Journal of Financial Management and Analysis*, 15(1):27–37, 2002.

Victor Zarnowitz. Rational expectations and macroeconomic forecasts. *Journal of Business & Economic Statistics*, 3:293–311, 1985.

Arnold Zellner. Biased predictors, rationality and the evaluation of forecasts. *Economics Letters*, 21:45–48, 1986.

Eric Zivot. Cointegration and forward and spot exchange rate regressions. Working paper, Department of Economics, University of Washington, September 1998.

Appendix 1: Testing Micro-homogeneity with Survey Forecasts

The null hypothesis of micro-homogeneity is that the slope and intercept coefficients in the equation of interest are equal across individuals. This paper considers the case of individual unbiasedness regressions such as equation (2) in the text, repeated here for convenience.

$$s_{t+h} - s_t = \alpha_{i,h} + \beta_{i,h}(s^e_{i,t,h} - s_t) + \varepsilon_{i,t,h} \tag{A1.1}$$

and tests $H_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_N$, and $\beta_1 = \beta_2 = \ldots = \beta_N$.

Stack all N individual regressions into the Seemingly Unrelated Regression system

$$S = \mathbf{F}\theta + \varepsilon \tag{A1.2}$$

where S is the $NT \times 1$ stacked vector of realizations, s_{t+h}, and $\mathbf{F}$ is an $NT \times 2N$ block diagonal data matrix

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_1 & & \\ & \ddots & \\ & & \mathbf{F}_N \end{bmatrix}. \tag{A1.3}$$

Each $\mathbf{F}_i = [\iota\, s^e_{i,t,h}]$ is a $T \times 2$ matrix of ones and individual i's forecasts, $\theta = [\alpha_1\, \beta_1 \ldots \alpha_N\, \beta_N]'$, and ε is an $NT \times 1$ vector of stacked residuals. The vector of restrictions, $R\theta = r$, corresponding to the null hypothesis of micro-homogeneity is normally distributed, with $R\theta - r \sim N[0, R(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\Omega\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}R']$, where R is the $2(N-1) \times 2N$ matrix

$$R = \begin{bmatrix} 1 & 0 & -1 & 0 & \ldots & 0 \\ 0 & 1 & 0 & -1 & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & 1 & 0 & -1 \end{bmatrix}, \tag{A1.4}$$

and r is a $2(N-1) \times 1$ vector of zeros. The corresponding Wald test statistic, $(R\hat{\theta} - r)'[R(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\hat{\Omega}\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}R'](R\hat{\theta} - r)$, is asymptotically distributed as a chi-square random variable with degrees of freedom equal to the number of restrictions, $2(N-1)$.

For most surveys, there are a large number of missing observations. Keane and Runkle (1990), Davies and Lahiri (1995), Bonham and Cohen (1995, 2001), and to the best of our knowledge all other papers which make use of pooled regressions in tests of the REH have dealt with the missing observations using the same approach. The pooled or individual regression is estimated by eliminating the missing data points in both the forecasts and the realization. The regression residuals are then padded with zeros in place of missing observations to allow for the calculation of own and cross-covariances. As a result, many individual variances and cross-covariances are calculated with relatively few pairs of residuals. These individual cross-covariances are then averaged. In Keane and Runkle (1990) and Bonham and Cohen (1995, 2001) the assumption of $2(k+1)$ second moments, which are common to all forecasters, is made for analytical tractability and for increased reliability. In contrast to the forecasts from the Survey of Professional Forecasters used in Keane and Runkle (1990) and Bonham and Cohen (1995, 2001), the JCIF data set contains virtually no missing observations. As a result, it is possible to estimate each individual's variance covariance matrix (and cross-covariance matrix) rather than average over all individual variances and cross-covariance pairs as in the aforementioned papers.

We assume that for each forecast group i,

$$\begin{aligned} E[\varepsilon_{i,t,h}\,\varepsilon_{i,t,h}] &= \sigma^2_{i,0} \quad \text{for all } i,t, \\ E[\varepsilon_{i,t,h}\,\varepsilon_{i,t+k}] &= \sigma^2_{i,k} \quad \text{for all } i,t,k \text{ such that } 0 < k \leq h, \\ E[\varepsilon_{i,t,h}\,\varepsilon_{i,t+k}] &= 0 \quad \text{for all } i,t,k \text{ such that } k > h, \end{aligned} \tag{A1.5}$$

Similarly, for each pair of forecasters i and j we assume

$$\begin{aligned} E[\varepsilon_{i,t,h}\,\varepsilon_{j,t}] &= \delta_{i,j}(0) \quad \forall i,j,t, \\ E[\varepsilon_{i,t,h}\,\varepsilon_{j,t+k}] &= \delta_{i,j}(k) \quad \forall i,j,t,k \text{ such that } k \neq 0, \text{ and } -h \leq k \leq h. \\ E[\varepsilon_{i,t,h}\,\varepsilon_{j,t+k}] &= 0 \quad \forall i,j,t,k \text{ such that } k > |h|. \end{aligned} \tag{A1.6}$$

Thus, each pair of forecasters has a different $T \times T$ cross-covariance matrix,

$$\mathbf{P}_{i,j} = \begin{bmatrix} \delta_{i,j}(0) & \delta_{i,j}(-1) & \dots & \delta_{i,j}(-h) & 0 \\ \delta_{i,j}(1) & \delta_{i,j}(0) & \delta_{i,j}(-1) & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ & \dots & \delta_{i,j}(1) & \delta_{i,j}(0) & \delta_{i,j}(-1) \\ 0 & \delta_{i,j}(h) & \dots & \delta_{i,j}(1) & \delta_{i,j}(0) \end{bmatrix}, \tag{A1.7}$$

Finally, note that $P_{i,j} \neq P_{j,i}$, rather $P'_{i,j} = P_{j,i}$. The complete variance-covariance matrix, denoted Ω, has dimension $NT \times NT$, with matrices $\mathbf{Q}_i$ on the main diagonal and $\mathbf{P}_{i,j}$ off the diagonal.

The individual Q_i, variance covariances matrices are calculated using the Newey and West (1987) heteroscedasticity-consistent, MA(j) corrected form. The $P_{i,j}$ matrices are estimated in an analogous manner,