

Exchange rate forecasters' performance: evidence of skill?

Ronald MacDonald, University of Glasgow, UK,
Lukas Menkhoff, Leibniz Universität Hannover, Germany, and
Rafael R. Rebitzky, Leibniz Universität Hannover, Germany

Abstract

This paper sheds new light on a long-standing puzzle in the international finance literature, namely, that exchange rate expectations appear inaccurate and even irrational. We find for a comprehensive dataset that individual forecasters' performance is skill-based. 'Superior' forecasters show consistent ability as their forecasting success holds across currencies. They seem to possess knowledge on the role of fundamentals in explaining exchange rate behavior, as indicated by better interest rate forecasts. Superior forecasters are more experienced than the median forecaster and have fewer personnel responsibilities. Accordingly, foreign exchange markets may function in less puzzling and irrational ways than is often thought.

JEL classification: F31, G14

Keywords: Foreign exchange market; individual exchange rate forecasts; interest rate forecasts; forecaster experience

February 16, 2009

We thank participants at the European Economic Association Annual Conference in Milan and at several seminars for their helpful comments. Moreover, we thank the Centre for European Economic Research (ZEW) in Mannheim, Germany, for providing the data. Financial support by the German Research Foundation (Deutsche Forschungsgemeinschaft) is gratefully acknowledged.

Ronald MacDonald, University of Glasgow, Department of Economics, Adam Smith Building, Glasgow G12 8RT, United Kingdom; email: R.MacDonald@lbss.gla.ac.uk.

Lukas Menkhoff and Rafael R. Rebitzky, Leibniz Universität Hannover, Department of Economics, Königsworther Platz 1, D-30167 Hanover, Germany; email: menkhoff@gif.uni-hannover.de, rebitzky@gif.uni-hannover.de.

Exchange rate forecasters' performance: evidence of skill?

1 Introduction

Is it really convincing that professional exchange rate forecasters persistently form inaccurate expectations, implying clear inefficiency in foreign exchange markets? This view was established twenty years ago by Frankel and Froot (1987). Their result has been confirmed in many subsequent studies, covering consensus expectations as well as expectations of individual forecasters and has become something of a stylized fact in the international finance literature.¹ However, there is contradicting evidence as well. Elliott and Ito (1999) show for an 11 years period that 42 firms' yen/US dollar forecasts seem to be slightly profitable, a view being shared by a few other studies on firm performances, including Goodman (1979) and Pojarliev and Levich (2007). Thus, there is doubt whether the Frankel and Froot verdict is a truly general characterization of exchange rate forecasters' performance.

This inconclusive evidence of existing studies, however, raises new questions. Why do some studies indicate irrational and others rational forecasting behavior? Are these opposing results largely accidental and thus driven by small samples drawn from a universe of heterogeneous forecasters? There is indeed some evidence indicating heterogeneous performance among exchange rate forecasters. However, what drives this heterogeneity? Are some forecasters just better than others because they are lucky or are there any systematic determinants? This is an important issue because the potential answer indicates a dramatically different view of the foreign exchange market than the conventional wisdom suggests. If performance is only driven by chance and not at all by skill then the outcome of this market would be regarded in a skeptical way. If, however, heterogeneous forecasting performance can be traced to skill, then the foreign exchange market seems to function like other markets and outcomes of this market, i.e. exchange rates, would be regarded as trustworthy prices.

In order to find out whether exchange rate forecasters' performance can be related to their skill, we regress possible skill-based determinants on a cross-section of forecasters' performance, an approach which has previously been applied in stock markets. Analyzing the success of forecasters, however, puts two strong requirements on the underlying data. First, and distinguishing this research from all former related studies on foreign exchange, one needs truly individual data that can be linked to further characteristics of these same forecasters. Second, the data set has to be comprehensive in order to provide variation in the cross-

section (i.e. many forecasters), in the time-series (i.e. including exchange rate up- and downswings) and in the object of forecasting (i.e. several exchange rates). We employ such a data set with about 150 professional forecasters, forecasting three exchange rates over a maximum period of 14 years.

The guiding question of our research is whether forecasting performance is systematically determined by skill, comparable to stock markets where performance can be partially explained in this way. In particular, we examine three questions:

First, is superior exchange rate forecasting based on *ability* in understanding foreign exchange markets in general as success determinants are not only specific to one exchange rate? We test this by examining whether the performance in forecasting an exchange rate can be explained by the performance in forecasting another exchange rate.

Second, is superior exchange rate forecasting determined by *knowledge* about exchange rate fundamentals? In order to test this we rely on “our” forecasters’ performance in contemporaneously forecasting several fundamentals.

Third, is superior exchange rate forecasting determined by some *personal characteristics*? We test this by relating performance to a set of available personal characteristics.

Whatever determinants we find, the relevance of our findings rests on three foundations. First, is our data set reliable in the sense that earlier stylized facts can be reproduced? Second, is the level of forecasting performance sensible for a competitive market, i.e. not too good and not too bad? Third, is heterogeneity in forecasters’ performance so large that it is economically relevant to examine its determinants? Fortunately, determinants are revealed in a sample providing solid foundations: irrational forecasting behavior is shown but for some shorter time periods only; the level of forecasting success is slightly better than random forecasting² and performance heterogeneity is profound (and beyond accidental heterogeneity). Based on these foundations we test the above mentioned three hypotheses which receive support: first, forecasters show consistent ability in that their success holds across currencies. Second, forecasters seem to possess knowledge on the role of fundamentals in explaining exchange rate behavior, as indicated by better interest rate forecasts. Third, better forecasters are more experienced than the median forecaster and in addition have fewer personnel responsibilities.

¹ See Dominguez (1986), the surveys by Takagi (1991) and MacDonald (2000), and the recent confirmation by Bacchetta et al. (2009).

² The average performance we find is slightly better than a random forecast, but not consistently so, as we will show later. Thus, there remains some uncertainty about the degree of forecasters’ performance, which we think is unavoidable in efficient markets. Clearly, irrational or poor forecasts are a sign of inefficiency, as of course are significantly good forecasts.

The paper is structured in seven parts. Section 2 relates our research to the literature and Section 3 introduces our data. In Section 4 we examine heterogeneity in performances among our forecasters, while in Section 5 we relate exchange rate forecast performance to individuals' corresponding forecast performance of exchange rate fundamentals. Then Section 6 links forecast performance of the individuals to related personal characteristics. In order to test for robustness, the main analyses are repeated in Section 7 by using profits based upon a popular trading strategy. Section 8 concludes.

2 Related literature

Following on from the seminal paper by Meese and Rogoff (1983), it has become a stylized fact that foreign exchange rates are difficult to forecast at horizons of less than three years (Frankel and Rose, 1995, Obstfeld and Rogoff, 2000, Cheung et al., 2005).³ However, this view has been challenged in a number of papers (for example, MacDonald and Taylor, 1994, MacDonald, 1999), where it is demonstrated that the successful modeling of the underlying data generation process can produce successful out-of-sample forecasts at horizons as short as two months ahead. In particular the monetary model's fundamentals seem to be linked to exchange rates (see e.g. MacDonald, 1999, Rapach and Wohar, 2002, Kilian and Taylor, 2003, Sarno et al., 2004, and Abhyankar et al., 2005). However, stability of these findings and forecasting power are challenged by Rogoff and Stavrakeva (2008).

In addition, practitioners provide a challenging contrast to the academic pessimism as we observe a gigantic volume in currency transactions which easily dominates that of the largest stock exchanges. Professionals who are relevant in this market – indicated by the size of their salaries and by the volume of assets under management – put money behind their forecasts, which questions the conventional academic wisdom of enduring “systematic expectational errors” by professional exchange rate forecasters (Frankel and Froot, 1987, p.150).⁴

Therefore, it would seem wrong to preclude that some forecasters may be able to interpret fundamentals in a useful way, showing that some forecasters might be superior to others. There is, indeed, mixed evidence on exchange rate forecasting performance. The literature starts with two early studies covering a small group of financial services which produce ex-

³ Recent research argues that this disconnection may be caused either by a time-consuming aggregation process (Bacchetta and van Wincoop, 2006), or by a non-fundamental price determinant resulting from the use of technical analysis (De Grauwe and Grimaldi, 2006).

⁴ According to the Bank for International Settlements (2005), the average daily turnover in April 2004 measured in billions of US-dollar for the US-dollar/euro, the JP-yen/euro and the GB-pound/euro market was 501, 51 and 43 respectively.

change rate forecasts on a professional basis and show ability (Goodman, 1979, and extension by Levich, 1979).⁵

The first study analyzing a broader sample of individual exchange rate forecasts is Ito (1990). He examines the yen-dollar exchange rate expectations collected by the Japan Center for International Finance (JCIF) over the period May 1985 to June 1987. The data set has a panel dimension, due to 44 firms from six industries participating. Ito finds heterogeneity in expectation formation, indicated by significant individual effects in the panel analysis, and questions forecasters' rationality.⁶ However, false expectations may be due to small sample size, as Ito (1990, p.437) states with respect to the yen appreciation after September 1985.

Another source of forecasting data is provided by Consensus Economics which consists of monthly forecasts of about 150 institutions. The individual dimension of this data is examined by MacDonald and Marsh (1994), who analyze the complete response of 30 forecasters over two years. Again, forecasts are mostly inaccurate, irrational and unprofitable. Findings are extended and largely confirmed by MacDonald and Marsh (1996) who also interestingly note that top performers in one currency are not necessarily also among the best in another currency, although this examination is limited to just 22 forecasters due to incomplete responses.⁷

The first sample which was long enough to cover both up- and downswings of exchange rates is Elliott and Ito (1999, p.435), who extend the sample from Ito (1990) to an eleven-year period until May 1996 and find that "survey data can be used to obtain on average positive profits". Interestingly, there is evidence that profits are quite volatile over time and that they are related to a common measure of risk, i.e. the difference between the forward rate and the average expected rate (Frankel and Froot, 1987). However, profits are not related to another measure of risk, i.e. dispersion of forecasts.

Finally, Pojarliev and Levich (2007) analyze the performance of 34 currency funds over 17 years and find profitability, indicating forecasting ability. They emphasize, however, that this performance is largely the compensation for several risk factors.

⁵ Further evidence on small samples contributes to the impression of heterogeneity: Blake et al. (1986) do not find ability in the exchange rate forecasting performance of three British institutions over five exchange rates at two horizons during three to four years, whereas Allen and Taylor (1990) do find some ability in their tracking of technical analysts' exchange rate forecasts.

⁶ This is supported by Wakita's (1989) analysis of a slightly extended sample ranging from May 1985 to March 1988.

⁷ Marsh and Power (1996) examine these 22 forecasters and assess their implicit portfolios built on long or short positions in three exchange rates. Although some forecasters seem to give profitable advice, only one forecaster does consistently so by significantly beating the no change alternative.

We summarize four main findings of studies on individual (or firm-specific) exchange rate expectations: first, there is indeed heterogeneity which makes it important to focus on individual forecasts and their determinants. Second, in order to avoid distorted performance the use of longer samples seems important. Third, evidence indicates that inaccurate point forecasts may be nevertheless good directional and profitable forecasts. Fourth, profits may represent a compensation for risk.

As the value of forecasts is thus questionable and forecasters seem to be so different in performance (for stock fund managers see Kosowski et al., 2006) and approach, it would be interesting to have more information about the determinants of their behavior. One important question for exchange rates is whether good forecasts are related to knowledge about economic fundamentals. Loh and Mian (2006) find such a relation for stock market analysts but for the foreign exchange market we are unaware of any comparable study.

The literature has also identified some personal characteristics that might help to understand (forecasting) performance. Chevalier and Ellisson (1999) find that good education improves the performance of fund managers. Another important variable is professional experience. More experienced analysts underreact less to prior earnings information (Mikhail et al., 2004) and issue rather bold and accurate than herding forecasts (Clement and Tse, 2005). Again, we are unaware of any such study in foreign exchange.

3 Data

Our analysis is based upon a micro dataset comprising more than 14 years of individual forecasts from the Financial Market Survey (Finanzmarkttest) of the Centre for European Economic Research (ZEW) in Mannheim, Germany. The individual forecasts constitute the data from which monthly aggregate statistics, such as consensus forecasts, have been computed from December 1991 until the present and which are passed to the financial media, such as Bloomberg and Reuters. The ZEW's survey structure is conventional and e.g. similar to Consensus Forecasts (London), since around about 75% of the participants of the survey work in the banking or bank-related sectors (i.e. retail banks, investment banks, investment funds, etc.), whereas the others are either associated with the insurance or industrial sector. Participants in the data set have about 300 responses on average. The survey asks participants to judge whether the respective variable will go up, down or remain unchanged over the next six months. The survey raises information on a monthly census of financial market professionals.

We study the micro data from December 1991 until July 2006, which produces up to 176 months of data, and the forecasts studied are for the US-dollar/euro, GB-pound/euro and

JP-yen/euro (until the end of 1998 the D-mark was used instead of the euro), as well as related US, British and Japanese exchange rate fundamentals, i.e. forecasts of inflation, economic growths and interest rates.

Since we study individual forecast series, we have to conduct several adjustments concerning the dataset. The initial sample observations regarding the US-dollar/euro (DM), GB-pound and Japanese-yen/euro forecasts appear in the first row of [Table 1](#). The second row indicates that approximately 36% of all initial forecasts are dropped to ensure consistency. Sometimes participants are replaced by new colleagues because of, for example, a job change. In such cases, we consider the new participant as another forecaster and we also track the previous participant, if he continues participating in the survey. Moreover, we set the minimum rate of participation to 50%, i.e. we require participation in at least 88 of 176 months, to ensure the elimination of accidental forecast accuracy. This reduces the amount of usable forecasts to about 25%.⁸ Finally, we use personal information about the forecasters. This limits respective analyses to participants for whom such information is available, which leads to further reduction of the forecasts by about 13% (see the fourth row of [Table 1](#)).

Some further characteristics of the final dataset are reported in [Table 2](#). This information stems from a specific questionnaire conducted in October 2006. All in all, forecasters' characteristics do not differ meaningfully on average between the three exchange rates. Even though the number of women participating in the survey has in recent years risen continuously, most of them do not pass the minimum participation rate; so the share of women is fairly low. The average age of participants is 48, whereas experience in financial market practice averages 21 years. So our representative forecaster is middle-aged and fairly experienced. This corresponds to the fact that 41 percent of participants head at least a department. Moreover, 54% of participants have personnel responsibilities, 80% have operational responsibility and, finally, 67% of the participants have an academic degree.

Since we are interested in links between forecasting power in exchange rates and competence in fundamental analysis, we base our analysis on related expectations data. As already noted, the ZEW-survey asks about exchange rate forecasts as well as several other economic variables. Considering the monetary model as our reference model, we draw on related inflation, interest rate and economic growth forecasts for the US, Great Britain and Japan respectively. Regarding the exchange rate forecasts, we use daily exchange rates of the US-

⁸ The 50% participation threshold obviously excludes many forecasts and thus also forecasters. There may be several reasons why participation changes, one reason can be a lack of success. Thus our sam-

dollar/euro, GB-pound/euro and JP-yen/euro (-/D-mark respectively). With respect to our fundamental forecasts, we use the monthly data for the euro zone CPI-based inflation rates, for industrial production to calculate a measure of monthly income growth, and for 6-month Libor rates to represent interest rates.⁹

A graphical presentation of the aggregated data – consensus forecasts and corresponding exchange rate changes over time – is shown in [Figure 1](#) for the US-dollar, GB-pound and JP-yen.¹⁰ At first glance, one can see that the direction of consensus exchange rate expectations is rather sluggish compared to realized exchange rate changes. This implies that there may be periods of “wrong” expectations and, indeed, just looking at the US dollar figure shows that such wrong expectations occurred in the early 1990s and, in particular, between 1998 and 2001. Accordingly, we can reproduce the standard finding of the literature, i.e. inaccurate (biased) expectations over some periods of time. However, the sign of the bias does sometimes switch, which indicates, as mentioned in some prior studies (e.g. Frankel and Froot, 1987), that the sample length may be important. Indeed, if we consider the whole 14 year period, there is no significant bias for any of the three exchange rates. Moreover, the graphical impression that expectations for the yen may often have the same sign as realized exchange rate changes is confirmed by our calculations: the hit rate of *consensus* expectations is 64.2% for the yen, 58.3% for the dollar but only 48.1% for the pound, so tossing a coin for the pound would have yielded a better result in our sample ([Appendix 2](#)). On average, 56.2% of consensus expectations in the total sample are in the right direction, which is a strong indication that exchange rate expectations are not systematically distorted in general.

This seems a reasonable basis to further examine *individual* forecasts in more detail. In doing so different techniques are required than for the so far considered aggregate (consensus) forecasts. As we compare now for each forecaster her tri-variate expectations – i.e. up, down or no change – with realized exchange rate changes, we need a definition of a “no change” band. Fortunately we can rely on a special ZEW survey where individual thresholds have

ple includes some element of survivorship bias. However, this is not relevant for our analysis because we examine the relative performance and not much interested in the level of forecasting performance.

⁹ Daily exchange rate data are obtained from the Deutsche Bundesbank, whereas 6-month Libor rates are taken from EcoWin. Industrial productions and CPI inflation stem from IMF’s International Financial Statistics. German government bond yields are taken from the Deutsche Bundesbank and US, British as well as Japanese yields from the Federal Reserve, the Bank of England and the Bank of Japan respectively.

¹⁰ Properties of the individual exchange rate forecasts are shown in [Appendix 1](#) including consensus, dispersion and actual exchange rate data.

been asked for.¹¹ The response is given separately for all forecast series so that we have direct information on how much variables have to change to be regarded as a change (say, for example, by three percent). Accordingly, we use related median thresholds, separated for the different variables. Although we do not claim that this is a perfect way of proceeding, we regard it as being simple, appropriate and transparent.

Finally, it is worth noting that we know and also use the exact dates when participants gave their forecasts. This is not self-evident since these forecasts are collected each month in two-weekly periods, so that the forecasters underlie in principle different market conditions if one does not carefully consider the individual response.

4 Heterogeneous performance of forecasters

In this section, we show that exchange rate forecasters differ with respect to their performance and that these differences are statistically significant and consistent across currencies.

Forecasters in the ZEW survey make directional forecasts, i.e. they predict that an exchange rate will go up, down or remain unchanged over the next six months, and so these forecasts can be directly compared to realizations of the exchange rates over the same horizon. The ZEW polls its survey participants about the size of the no change band and this is three percent in both directions from the present exchange rate. Accordingly, we regard an exchange rate change by less than three percent over the next six months as consistent with a no-change prediction.

As we have three alternatives that a forecaster can choose from, there will only be one correct outcome and the forecaster can make two mistakes. In particular, when the exchange rate goes up or down, there is the possibility of a more or less severe forecasting mistake, depending on whether the forecasters predicted no change – which is a comparatively small mistake – or whether the forecaster predicted the wrong direction of change – which is a comparatively large mistake. In order to cover the severity of wrong forecasts we do not simply code right or wrong forecasts, but we code three possibilities, i.e. right, small and big mistake. The codes for these alternatives are either 2, 1 or 0, so that higher codes indicate better forecasts. Forecasting performance is thus measured as directional performance and the measure can be understood as a hit rate.

¹¹ Actually, several studies follow the other way round, using specific thresholds in order to quantify qualitative forecasts (see surveys by Nardo, 2003, and Pesaran and Weale, 2006).

Figure 2 gives the frequency distribution of hit rates that the 150 forecasters realized during the sample period for the three currencies. One can directly see from this that there is a lot of divergence in forecasting performance. Taking the first graph in Figure 2, for the US dollar, almost all forecasters realize an average hit rate of between 0.9 and 1.3, the median being about 1.1. The realized hit rates for the GB pound and JP yen show somewhat higher values than for the US dollar, as individual hit rates lie basically in the range between 1.0 and 1.4, with the pound results being comparatively higher.

These analyses are informative regarding divergence of performance which is the objective of this paper. However, in order to indicate the individuals' quality-level of forecasting performance, an appropriate benchmark would be the performance of an uninformed forecaster who makes exchange rate predictions just by chance. In this way, the realized hit rate of random forecasts is calculated. Therefore we take the exchange rate time series and allocate to each month either an "up", "down" or "no change", according to the preceding six month effective change. This results in a concrete distribution of changes for our period of time. We take this given distribution and simulate 10,000 random time series, then we calculate hit rates by comparing the random time series – expressing forecasting by chance – with effective exchange rate changes. Obviously, the average simulated random hit rate can differ according to the time series properties of the underlying exchange rate.¹² We find for our sample that the benchmark to be beaten in the case of the US dollar is a simulated hit rate of 1.0975, for the pound the value is 1.2530 and it is 1.0928 for the yen – these values, as well as the respective 95% confidence bands, are also shown in Figure 2.

The information about the random forecast performance and the individual forecast performances – as graphically presented in Figure 2 – can be used to explicitly show the best and worst forecasters in the sample. Appendix 4 gives respective information about hit rates realized for the US dollar. The best forecaster realizes an average hit rate of 1.31 on 101 participations, which at the one percent level is significantly different from a random forecast. Overall, there are six individuals significantly better than the random forecast at the 5% level. Interestingly, there are also 14 individuals who are significantly worse at the same level of significance. Considering individual forecasting performance of our two other exchange rates also gives an interesting result (see Appendices 5 and 6): among 146 pound forecasters, there are 24 who produce forecasts that are significantly better than a random forecast at the 5%

¹² In particular we do not assume that exchange rates change according to a distribution where one third of cases each is exchange rate up, down and no change. In fact, the simulated distribution of the

level of significance, but 37 produce forecasts which are worse. The picture is different again for the yen, where 58 are better and only 2 worse than the benchmark at the 5% level of significance.

Overall, we find statistically successful and unsuccessful forecasters for all three exchange rates. Although the mean performance is very different for the three exchange rates – i.e. comparatively the best for the yen, worst for the dollar and very heterogeneous for the pound – the coexistence of successful “hot hands” and unsuccessful “cold hands” is a new aspect in the foreign exchange literature.¹³ This finding raises the question of whether skill depends on the specific exchange rate or whether it depends on general ability. Fortunately, most individuals in our sample – i.e. 145 persons – give forecasts for all three exchange rates. We can thus relate their hit rates across exchange rates and consistently find positive correlations. [Table 3](#) shows that the hit rate of US dollar forecasts can be significantly “explained” by the quality of pound forecast but not by yen forecasts. Among the three relations between currencies, it is only the dollar-yen exchange rate where the spillover is not significant. Nevertheless, even this coefficient is positive and the others are significantly positive, which indicates that performance in forecasting exchange rates is also driven by a general ability and not just specific to certain exchange rates.

This provides a strong motivation to further investigate the possible determinants of forecasting performance and, in particular, to investigate a possible role of knowledge.

5 Forecasting the exchange rate and forecasting its fundamentals

In this section we examine the issue of whether individual heterogeneity in forecasting performance can be traced to the performances in other economic variables. Principally, we seek to determine if any forecasting success in exchange rates is related to the accuracy of fundamental research, i.e. the analysis of exchange rate fundamentals. Since we have access to individuals’ exchange rate forecasts and also to their forecasts of fundamental variables – all formed at the same time – we can investigate whether such relationships do indeed exist.

There are several reasons why such an analysis may prove insightful. First, it is demonstrated in the empirical literature that exchange rates are connected to fundamentals in the long run (see recent studies e.g. Rapach and Wohar, 2002, Kilian and Taylor, 2003, and Sarno et al., 2009). Second, longer term exchange rate expectations underlie strong fundamental

GB-pound/euro hit rates differs from the two other simulated distributions, because relatively many actual GB-pound/euro changes have occurred in the indeterminate interval, i.e. “no change”.

considerations (see Cheung and Wong, 2000 and MacDonald, 2000). Third, we already showed in Section 3 that the majority of participants of the Financial Market Survey belong to groups which follow longer-term considerations, such as analysts and fund managers. Fourth, a related study on the stock market by Loh and Mian (2006) reveals such a fundamental relationship. In fact, they show a fundamental link between the profitability of analysts' stock recommendations and the accuracy of corresponding earnings forecasts. Taken together, this motivates us to analyze whether the quality in fundamental analysis coincides with forecast success in exchange rates. A priori, we thus expect average hit rates of exchange rate fundamentals to be positively correlated with exchange rate hit rates. With respect to the relevant fundamental variable set, we choose the monetary model as a baseline and thus consider inflation, economic growth and the interest rate for Euroland, the US, Great Britain and Japan, respectively.

In fact, using the exchange rate fundamentals of the monetary model for all three exchange rates, we find a significant positive relationship between superior exchange rates forecasts and interest rates forecasts. Additionally, GB-pound forecasts prove to be significantly better on average than US-dollar as well as JP-yen forecasts.

Using a rule-of-thumb calculation, we note that a neutral hit rate would generate a value of 1.1. Since most of the average hit rates lie right of 1.1, we at least acknowledge that the majority of the forecasters do not form poor forecasts. Although we do not want to overemphasize these results, since we have said nothing so far about statistical significance, we nevertheless believe this finding serves as an indicator concerning the overall accuracy of the individual exchange rate forecasts.

In line with our procedure in Section 4, we have to generate qualitative forecast errors for all respective fundamentals. However, since these variables have different underlying statistical characteristics from the exchange rate, individual threshold values have to be considered. Again, we use the threshold values which have been revealed by the Centre for Economic European Research amongst the survey's participants, via specific questionnaires. Subsequent reported threshold values are measured as percentages of the respective variables, and they each indicate if a specific participant classifies a variable to go up or down; a respective threshold value of 4.5 percent has been incorporated into our calculations.

¹³ Of course, already the early studies of Goodman (1979) and Levich (1979) mention this fact but for a very small group only, i.e. 9 and 10 participants respectively.

Table 4 represents the corresponding results of the relations in forecasting performances and fundamentals for the US-dollar/euro, GB-pound/euro and the JP-yen/euro.¹⁴ Regarding the US-dollar, superior domestic interest rate forecasts significantly correlate with superior exchange rate forecasts. In addition, US inflation hit rates exercise a negative influence on exchange rate hit rates; however, the latter relationship is not significant at the 5 percent level. Turning to the GB-pound, we find that foreign inflation positively influences forecast success in exchange rates but only at the ten percent level. However, higher hit rates in both domestic and GB-interest rates are significantly related to higher hit rates in exchange rates. Finally, for the JP-yen, only hit rates in Japanese interest rates correlate positively and significantly with those of the exchange rate. So in respect to all three exchange rate equations, we find only significant connections with one or both (domestic and foreign) interest rate hit rates, but consistent with our prior, the signs prove to be positive.

In order to extend these findings, we now analyze all exchange rate equations simultaneously. Table 5 shows corresponding results of the pooled data, where the JP-yen is set as the basic currency. Different forecast success rates between the US-dollar, the GB-pound and the JP-yen exchange rate can be attributed to different complexities of the respective exchange rates. In order to control for these constant country effects, we use a US-dollar dummy and a GB-pound dummy. It turns out that the only variables which are significant at the five percent level are: domestic and foreign interest rate hits as well as the GB-pound dummy. Furthermore, domestic inflation and the US-dollar dummy are significant at the 10 percent level. Restricting the model to significant variables via a top down approach, we find that only the two interest rates and the GB-pound dummy remain in the model.

In summary, superior analyses in domestic and foreign interest rates are strongly related to greater forecasting success in exchange rates. The significant influence arising from the GB-pound dummy confirms the former observation with respect to Figure 2, where exchange rate forecasts of the GB-pound appear better than those of the other exchange rates. However, the latter finding reveals a constant effect, independent of the performance of fundamental analysis.

¹⁴ Since the minima and maxima of each exchange rate hit rate series are far from their natural bounds zero and plus two, there are no extreme aggregate expectations in the sample (see also Appendix 3).

6 Personal characteristics of superior exchange rate forecasters

Due to available personal characteristics of exchange rate forecasters we find that there are two significant influences on performance: professional experience in financial markets seems to help, whereas responsibility for personnel seems to harm performance.

The difficulty in analyzing a possible influence from personal characteristics is data availability. Indeed, this data are not regularly compiled but stems partly from an additional survey that the ZEW sometimes conducts among participants of the ZEW financial market survey. As a consequence, the response to this survey provides another restriction to our sample, further reducing the number of respondents to 93 individuals. However, we consider the forecasts of each exchange rate independently, and thus analyze 279 cases.

In particular, we examine the possible impact of the seven items documented in Table 2 on forecasting performance. As the average forecasting performance differs much between currencies (see Section 4), we include exchange rate dummies in the regression. Moreover, we consider findings in the preceding Section 5 and include fundamental forecast performance in the regression because personal characteristics may be related to forecasting ability in fundamentals too. All this raises the stakes for personal characteristics to become significant.

Table 6 shows the role of personal characteristics in explaining forecasting performance. Specification (1) gives the result when all variables are considered. It shows that beyond the constant, some fundamentals and exchange rate dummies only one personal characteristic seems to be important, i.e. a detrimental influence from personnel responsibilities. In another specification we reduce the variables considered to those that remain significant. This leads to specification (2) where a second personal characteristic of importance appears, i.e. a positive influence from experience.

It is reassuring that regression results on the role of personal characteristics also hold when we exclude the fundamental variables (not shown to save space). Moreover, we note that the major influences found in Table 5 – i.e. considering the impact of fundamentals forecast performance but not of personal characteristics – remain unchanged, although the number of observations is reduced from 442 to 279. We regard this as another sign of robustness and it also shows that personal characteristics are important over and above fundamental forecasts, although the latter are comparatively more important than personal characteristics.

7 The relation between directional forecasting success and profits

As a robustness test, we examine whether directional forecasting success translates into profits. We find, indeed, and in accordance with earlier studies, that our results hold in a

qualitative sense when the directional performance measures used so far are substituted by profits.

Professionals participating in the ZEW survey give their directional expectations and it seems reasonable to form a trading strategy based on these expectations. Accordingly, we assume that the expectation of an “up” of the US dollar is translated into a US dollar long position, a “down” into a short and a “no change” into a neutral position. Although the expectations are formed for six-month horizons, the possibility of new information should be considered, so that effectively, a position may be changed whenever the expectation changes, which may occur once per month. Then the return for each month is calculated, assuming that the asset value is always the same; in particular, this procedure excludes cumulating profits. Finally, these (gross) profits resulting purely from exchange rate changes are complemented by net profits where the interest rate difference against the euro is taken into account.¹⁵

Based on this procedure, the graphical distribution of returns for about 150 individuals, separated for the three exchange rates is obviously very similar to the frequency distribution of hit rates introduced above (Figure 2). If we take the zero profit position as a benchmark, we also recognize that trading strategies based on forecasters’ expectations would yield some profits in most cases (see [Appendix 7](#)).¹⁶ Overall, forecasters’ performance measured by profits produces the same qualitative picture as the hit rate measure: professionals’ expectations are more often accurate than inaccurate.

Clearly, the same qualitative result does not necessarily imply that it is driven by the same persons. Thus, we examine the relations – person by person – between the performance measured by hit rates and that measured by profits, for the three exchange rates individually. [Table 7](#) shows that there is a highly significant positive relation. This relation holds independent of the method applied, i.e. either an OLS regression or (non)parametric correlations (see also [Figure 3](#)).

Finally, we follow the direction of Elliott and Ito (1999) and test whether forecasting returns may be a compensation for risk. We do not find many significant relations and in the cases which are, they often have the “wrong” sign (see [Table 8](#)). This indicates that forecast-

¹⁵ The translation of expectations in trading strategies requires making some assumptions and is thus open to criticism. Therefore, we regard the profit calculations as a robustness exercise only and prefer to stick as closely as possible to the available information, i.e. directional expectations, and assess them via the hit rate measure.

¹⁶ Accordingly, the average yearly returns for the best and the worst 25 forecasters in the US-dollar/euro, which would have been realized by following corresponding forecasts month-by-month, are qualitatively similar to the result based on hit rates (see [Appendix 8](#)).

ers' performance holds after risk-adjustment, a finding which is largely consistent with the results in Elliott and Ito (1999).

8 Conclusion

The research reported in this paper seeks to shed new light on a long-standing puzzle in international finance, namely the finding of inaccurate, and even irrational, consensus exchange rate expectations of professional forecasters. In order to gain more insight into this issue we employ a new dataset which has three advantages over previous studies: first, it contains truly individual (and not firm-specific) data; second, it spans a long time period (more than 14 years) and a broad cross section (we rely on about 150 persons); third, and perhaps most interestingly, individuals' exchange rate expectations can be linked to other information, such as individual expectations on exchange rate fundamentals and forecasters' personal characteristics.

We use this comprehensive dataset to reveal that skill is an important determinant of forecasting performance in three ways: first, professional exchange rate forecasters forming superior directional expectations on one exchange rate are likely to forecast better on other exchange rates as well; second, superior forecasters make use of their knowledge about relevant fundamentals; and, third, their performance is influenced by experience (positively) and by personnel responsibilities (negatively).

These findings contribute to an assessment of foreign exchange markets that is similar to that of other financial markets: professionals in this market do not behave irrationally in that they consistently form unnecessarily bad expectations but their expectations reflect their differential forecasting skill. These results are not due to the use of a particular dataset. Translating the directional expectations into consensus point forecasts for shorter samples – i.e. repeating the exercises of Frankel and Froot (1987) and others – reproduces their finding as these consensus expectations are often inaccurate and irrational. However, this does not apply to the full sample.

References

- Abhyankar, Abhay, Lucio Sarno and Giorgio Valente (2005), Exchange rates and fundamentals: evidence on the economic value of predictability, *Journal of International Economics*, 66: 325-348.
- Allen, Helen and Mark P. Taylor (1990), Charts, noise and fundamentals in the London foreign exchange market, *Economic Journal*, 100: 49-59.
- Bacchetta, Philippe and Eric van Wincoop (2006), Can information heterogeneity explain the exchange rate determination puzzle?, *American Economic Review*, 96: 552-76.
- Bacchetta, Philippe, Elmar Mertens and Eric van Wincoop (2009), Predictability in financial markets: what do survey expectations tell us?, *Journal of International Money and Finance*, forthcoming.
- Bank for International Settlements (2005), Triennial central bank survey of foreign exchange and derivatives market activity in 2004, March 2005.
- Blake, David, Michael Beenstock and Valerie Brasse (1986), The performance of UK exchange rate forecasters, *Economic Journal*, 96: 986-99.
- Carlson, John A. and Michael Parkin (1975), Inflation expectations, *Economica*, 42: 123-38.
- Cheung, Yin-Wong, Menzie D. Chinn and Antonio Garcia Pascual (2005), Empirical exchange rate models of the nineties: are any fit to survive?, *Journal of International Money and Finance*, 24: 1150-75.
- Cheung, Yin-Wong and Clement Y.-P. Wong (2000), A survey of market practitioners' views on exchange rate dynamics, *Journal of International Economics*, 51: 401-19.
- Chevalier, Judith and Glenn Ellison (1999), Are some mutual fund managers better than others? Cross-sectional pattern in behavior and performance, *Journal of Finance*, 3: 875-99.
- Clement, Michael B. and Senyo Y. Tse (2005), Financial analyst characteristics and herding behavior in forecasting, *Journal of Finance*, 60: 307-41.
- De Grauwe, Paul and Marianna Grimaldi (2006), Exchange rate puzzles: a tale of switching attractors, *European Economic Review*, 50: 1-33.
- Dominguez, Kathryn M. (1986), Are foreign exchange forecasts rational? New evidence from survey data, *Economics Letters*, 21: 277-281.
- Elliott, Graham and Takatoshi Ito (1999), Heterogeneous expectations and tests of efficiency in the yen/dollar forward exchange rate market, *Journal of Monetary Economics*, 43: 435-456.

- Frankel, Jeffrey A. and Kenneth A. Froot (1987), Using survey data to test standard propositions regarding exchange rate expectations, *American Economic Review*, 77: 133-153.
- Frankel, Jeffrey A. and Andrew K. Rose (1995), Empirical research on nominal exchange rates, in: G. Grossmann and K. Rogoff (eds.), *Handbook of International Economics*, Vol. III (Amsterdam et al.: North-Holland): 1689-729.
- Goodman, Stephen H. (1979), Foreign exchange rate forecasting techniques: implications for business and policy, *Journal of Finance*, 34(2): 415-27.
- Ito, Takatoshi (1990), Foreign exchange rate expectations: micro survey data, *American Economic Review*, 80(3): 434-49.
- Kilian, Lutz and Mark P. Taylor (2003), Why is it so difficult to beat the random walk forecast of exchange rates?, *Journal of International Economics*, 60: 85-107.
- Kosowski, Robert, Allan Timmermann, Russ Wermers and Hal White (2006), Can mutual fund “stars” really pick stocks? New evidence from a bootstrap analysis, *Journal of Finance*, 111: 2551-95.
- Leitch, Gordon and J. Ernest Tanner (1991), Economic forecast evaluation: profits versus the conventional error measures, *American Economic Review*, 81: 580-90.
- Levich, Richard M. (1979), Analyzing the accuracy of foreign exchange advisory services: theory and evidence, NBER working paper 336.
- Loh, Roger K. and G. Mujtaba Mian (2006), Do accurate earnings forecasts facilitate superior investment recommendations?, *Journal of Financial Economics*, 80: 455-83.
- MacDonald, Ronald (1999), Exchange rates: do fundamentals matter?, *Economic Journal*, 109: 673-91.
- MacDonald, Ronald (2000), Expectations formation and risk in three financial markets: Surveying what the surveys say, *Journal of Economic Surveys*, 14(1): 69-100.
- MacDonald, Ronald and Ian W. Marsh (1994), Combining exchange rate forecasts: What is the optimal consensus measure?, *Journal of Forecasting*, 13: 313-332.
- MacDonald, Ronald and Ian W. Marsh (1996), Currency forecasters are heterogeneous: confirmation and consequences, *Journal of International Money and Finance*, 15: 665-685.
- MacDonald, Ronald and Mark P. Taylor (1994), The monetary model of the exchange rate: long-run relationships, short-run dynamics and how to beat a random walk, *Journal of International Money and Finance*, 13(3): 276-290.
- Marsh, Ian and David M. Power (1996), A note on the performance of exchange rate forecasters in a portfolio framework, *Journal of Banking and Finance*, 20: 605-13.

- Meese, Richard A. and Kenneth Rogoff (1983), Empirical exchange rate models of the seventies: do they fit out of sample?, *Journal of International Economics*, 14: 345-73.
- Menkhoff, Lukas and Mark P. Taylor (2007), The obstinate passion of foreign exchange professionals: technical analysis, *Journal of Economic Literature*, 45:936-972.
- Mikhail, Michael B., Beverly R. Walther and Richard H. Willis (2004), Do security analysts exhibit persistent differences in stock picking ability?, *Journal of Financial Economics*, 74: 67-91.
- Nardo, Michela (2003), The quantification of qualitative survey data: a critical assessment, *Journal of Economic Surveys*, 17: 645-668.
- Obstfeld, Maurice and Kenneth S. Rogoff (2000), The six major puzzles in international macroeconomics: is there a common cause?, *NBER Macroeconomics Annual*, 339-390.
- Pesaran, M. Hashem and Martin Weale (2006), Survey expectations, in: *Handbook of Economic Forecasting*, Clive Granger and Alan Timmermann (eds.), Amsterdam, 715-776.
- Pojarliev, Momtchil and Richard M. Levich (2007), Do professional currency managers beat the benchmark?, NBER working paper 13714.
- Rapach, David E. and Mark E. Wohar (2002), Testing the monetary model of exchange rate determination: new evidence from a century of data, *Journal of International Economics*, 58: 359-385.
- Rogoff, Kenneth S. and Vania Stavrageva (2008), The continuing puzzle of short horizon exchange rate forecasting, NBER working paper 14071.
- Sarno, Lucio, Pasquale Della Corte and Ilias Tsiakas (2009), An economic evaluation of empirical exchange rate models, *Review of Financial Studies*, forthcoming.
- Sarno, Lucio, Giorgio Valente and Mark E. Wohar (2004), Monetary fundamentals and exchange rate dynamics under different nominal regimes, *Economic Inquiry*, 42: 179-193.
- Takagi, Shinji (1991), Exchange rate expectations: A survey of survey studies, *IMF Staff Papers*, 38: 156-183.
- Wakita, Shigeru (1989), Are survey forecasts trusted? American trade account deficit and Yen/Dollar rate, *Economics Letters*, 29: 339-44.

TABLE 1

Sample selection criteria and related observations

	sum of fore- casts	number of forecasts		
		US-dollar	GB-pound	JP-yen
raw sample	156,040	53,750	51,384	50,906
consistent sample	100,315	34,482	33,192	32,641
50% participation sample	60,540	20,965	19,834	19,741
final sample	40,344	13,859	13,280	13,205

Notes: The raw sample has to be corrected since we focus on identified persons, i.e. consistency criteria. Since participants of the survey change occasionally their professions, sometimes it happens that someone adopts the identity number (ID) of the former colleague. This would obviously bias our forecast series, therefore, we track on new individuals lodging them internally under new ID's and if possible, tracking changing participants on their new job under their ID. Furthermore, adopting the participation criteria, we set a minimum participation rate of 50%, meaning that only these persons are considered, who participate at least 88 times at the survey (in total, we deal with 176 months). Finally, for the final sample, we use personal information, which requires a selection via the personal criteria.

TABLE 2

Personal characteristics of the forecasters

	US-dollar	GB-pound	JP-yen
share of women	2%	1%	1%
average age in years	48	48	48
average experience in years	21	21	21
share of participants with academic background	68%	67%	67%
share of participants holding a chief position	40%	41%	42%
share of participants with operational responsibilities	80%	80%	80%
share of participants with personnel responsibilities	53%	54%	55%

Notes: These characteristics are based upon the final sample after adopting the personal criteria, see Table 1. This information follows from a specific questionnaire conducted in October 2006. The share of women according to the questionnaire response is higher than displayed above, because the share of female participants was low in the beginning years of the survey and has risen continuously until today. The characteristic experience is explicitly asked for experience in financial markets. Participants are considered having an academic background if they hold a bachelor, master or doctor degree. Furthermore, we judge participants as capturing a chief position, if they head at least a department. Operational responsibilities correspond to an employee who operates on its own authority, whereas personnel responsibilities is defined the way that someone can take disciplinary actions against junior staff – nevertheless, that must not automatically relate to a chief position. Finally, about 75% of the participants of the survey work in the banking sector (i.e. retail banks, investment banks, investment funds, etc.), whereas the others are either associated to the insurance or industrial sector. So, the survey's composition turns out being very similar to others, e.g. Consensus Forecasts London with 75% working in financial institutions.

TABLE 3 Interrelations between hit rate performances

$$fx_i^{hr} = \alpha + \beta \cdot fx_j^{hr} + e_i \quad \text{with } j \neq i$$

	US-dollar		GB-pound		JP-yen	
const.	0.82 ^{***}	[0.000]	0.49 ^{***}	[0.004]	0.83 ^{***}	[0.000]
US-dollar	-	[n. a.]	0.22 ^{**}	[0.042]	0.08	[0.299]
GB-pound	0.13 ^{**}	[0.042]	-	[n. a.]	0.21 ^{***}	[0.000]
JP-yen	0.09	[0.299]	0.42 ^{***}	[0.000]	-	[n. a.]
N	145		145		145	
R ²	0.05		0.13		0.05	
adj. R ²	0.04		0.12		0.09	
F-statistic	3.73 ^{**}	[0.027]	10.47 ^{***}	[0.000]	2.57 [*]	[0.049]

Notes: We assume a 6-month forecast horizons for the exchange rates and set the expectation threshold – according to these forecasters’ statement – at 3%. The hit rates range from 0 to 2, with 2 representing a correct qualitative forecast, 1 a simple forecast error (e.g. no change forecast but actually the respective rate falls) and 0 a double forecast error (e.g. forecast of a rise but actually the respective rate falls). We run OLS estimations, each exchange rate as the dependent variable, respectively. Asterisks refer to the regressors’ level of significance in the short-term relations: *, **, *** to ten, five and one percent. Single correlations between the exchange rates show up as following: US-dollar and GB-pound 0.206, US-dollar and JP-yen 0.147, GB-pound and JP-yen 0.320.

TABLE 4 Hit rates performance of FX and fundamental forecasts

$$fx_i^{hr} = \alpha + \theta \cdot X_i^{hr} + e_i$$

	US-dollar		GB-pound		JP-yen	
const.	1.04 ^{***}	[0.000]	0.36	[0.110]	0.92 ^{***}	[0.000]
Inflation	-0.05	[0.482]	-0.17 [*]	[0.092]	-0.03	[0.743]
Inflation*	-0.12 [*]	[0.093]	0.16	[0.104]	-0.01	[0.618]
Growth	-0.11	[0.106]	-0.05	[0.591]	-0.07	[0.345]
growth*	-0.01	[0.920]	0.09	[0.424]	0.10	[0.311]
Interest	0.29 ^{***}	[0.001]	0.24 ^{**}	[0.021]	0.02	[0.794]
interest*	-0.00	[0.993]	0.35 ^{***}	[0.000]	0.15 ^{***}	[0.008]
N	152		145		145	
R ²	0.11		0.30		0.07	
adj. R ²	0.07		0.27		0.03	
F-statistic	2.87 ^{**}	[0.011]	9.82 ^{***}	[0.000]	1.81	[0.101]

Notes: The results are based upon 176 monthly observations from December 1991 until July 2006. We assume 6-month forecast horizons for the exchange rates and the fundamentals as well. We set expectation thresholds at different rates: Exchange rates 3%, inflation, growth and interest rates at 4.5%, respectively. The hit rates range from 0 to 2, with 2 representing a correct qualitative forecast, 1 a simple forecast error (e.g. no change forecast but actually the respective rate falls) and 0 a double forecast error (e.g. forecast of a rise but actually the respective rate falls). We run regressions separated for the exchange rates using OLS. The minimum participation is set at 50%, so we consider only individuals who participated at least at 88 months. Following mean hit rates result, when we set the latter restriction: US-dollar 1.10, GB-pound 1.22 and JP-yen 1.20. Asterisks refer to the regressors’ level of significance in the short-term relations: *, **, *** to ten, five and one percent.

TABLE 5

Hit rate performance of FX and fundamental forecasts

$$fx_i^{hr} = \alpha + \theta \cdot X_i^{hr} + \kappa \cdot D_j + e_i$$

	full model		restricted model	
const.	0.71 ^{***}	[0.000]	0.55 ^{***}	[0.000]
Inflation	-0.10 [*]	[0.059]	-	[n. a.]
Inflation*	-0.01	[0.562]	-	[n. a.]
Growth	-0.07	[0.152]	-	[n. a.]
growth*	0.06	[0.334]	-	[n. a.]
Interest	0.19 ^{***}	[0.001]	0.13 ^{***}	[0.008]
interest*	0.22 ^{***}	[0.000]	0.27 ^{***}	[0.000]
D ₁ (US-dollar)	-0.03 [*]	[0.086]	-	[n. a.]
D ₂ (GB-pound)	0.08 ^{***}	[0.000]	0.09 ^{***}	[0.000]
N	442		442	
R ²	0.37		0.35	
adj. R ²	0.35		0.35	
F-statistic	31.13 ^{***}	[0.000]	79.90 ^{***}	[0.000]

Notes: The sample contains 176 monthly observations from December 1991 until July 2006. We run a combined regression for all three exchange rates using OLS. We assume a 6-month forecast horizons for the exchange rates and the fundamentals as well and set the expectation threshold at different rates: Exchange rates 3%, inflation, growth and interest rates at 4.5%, respectively. The hit rates range from 0 to 2, with 2 representing a correct qualitative forecast, 1 a simple forecast error (e.g. no change forecast but actually the respective rate falls) and 0 a double forecast error (e.g. forecast of a rise but actually the respective rate falls). We add two country dummies in order to control for country effects with dummy 1 represents the US and dummy 2 GB. The minimum participation is set at 50%, so we consider only individuals who participated at least at 88 months. Following mean hit rate results, when we set the latter restriction: 1.1713. Asterisks refer to the regressors' level of significance in the short-term relations: *, **, *** to ten, five and one percent.

TABLE 6

Hit rate performance of FX, fundamental forecasts and personal information

$$f\hat{x}_i^{hr} = \alpha + \theta \cdot X_i^{hr} + \kappa \cdot D_j + \omega \cdot Y_i + e_i$$

	full model		restricted model	
const.	0.62 ^{***}	[0.000]	0.57 ^{***}	[0.000]
Inflation	-0.16 ^{**}	[0.032]	-	[n. a.]
Inflation*	0.01	[0.806]	-	[n. a.]
Growth	-0.08	[0.184]	-	[n. a.]
growth*	0.14 [*]	[0.062]	-	[n. a.]
Interest	0.20 ^{***}	[0.008]	0.13 [*]	[0.052]
interest*	0.22 ^{***}	[0.000]	0.25 ^{***}	[0.000]
D ₁ (US-dollar)	-0.02	[0.362]	-	[n. a.]
D ₂ (GB-pound)	0.09 ^{***}	[0.000]	0.09 ^{***}	[0.000]
Sex	-0.03	[0.488]	-	[n. a.]
age	0.00	[0.749]	-	[n. a.]
experience	0.01	[0.251]	0.01 ^{**}	[0.014]
academic	0.01	[0.341]	-	[n. a.]
chief	0.01	[0.272]	-	[n. a.]
operations	-0.01	[0.725]	-	[n. a.]
personnel	-0.04 ^{***}	[0.005]	-0.03 ^{***}	[0.003]
N	279		279	
R ²	0.38		0.35	
adj. R ²	0.35		0.34	
F-statistic	10.81 ^{***}	[0.000]	30.03 ^{***}	[0.000]

Notes: The sample contains 176 monthly observations from December 1991 until July 2006. We run a combined regression for all three exchange rates using OLS. We assume a 6-month forecast horizons for the exchange rates and the fundamentals as well. We set the expectation thresholds at different rates: Exchange rates 3%, inflation, growth and interest rates at 4.5%, respectively. The hit rates range from 0 to 2, with 2 representing a correct qualitative forecast, 1 a simple forecast error (e.g. no change forecast but actually the respective rate falls) and 0 a double forecast error (e.g. forecast of a rise but actually the respective rate falls). The minimum participation is set at 50%, so we consider only individuals who participated at least at 88 months. Following mean hit rate results, when we set the latter restriction: 1.1756. We add two country dummies in order to control for country effects with dummy 1 represents the US and dummy 2 GB. Moreover we control for the following personal characteristics of the participants: Sex (1 male, 0 female), age, experience (in years), academic (1 university degree, 0 not), chief (1 yes, 0 no), operations (1 operational responsibilities, 0 not) and personnel (1 personnel responsibilities, 0 not). Asterisks refer to the regressors' level of significance in the short-term relations: *, **, *** to ten, five and one percent.

TABLE 7

Relation between hit rates returns in exchange rate forecasting

$$fx_i^{re} = \alpha + \beta \cdot fx_i^{hr} + e_i$$

	US-dollar		GB-pound		JP-yen	
const.	-1.01 ^{***}	[0.000]	-0.64 ^{***}	[0.000]	-0.79 ^{***}	[0.001]
hit rate	0.98 ^{***}	[0.000]	0.52 ^{***}	[0.000]	0.80 ^{***}	[0.000]
N	152		146		146	
R ²	0.1793		0.1637		0.1024	
adj. R ²	0.1738		0.1579		0.0961	
F-statistic	32.77 ^{***}	[0.0000]	28.20 ^{***}	[0.0000]	16.42 ^{***}	[0.0001]

Notes: The hit rates and returns are taken from calculations mentioned above. We run OLS estimations, each exchange rate as the dependent variable, respectively. Asterisks refer to the regressors' level of significance in the short-term relations: *, **, *** to ten, five and one percent. Single correlations between the two performance measurements show up as following: US-dollar 0.4234, GB-pound 0.4047 and JP-yen 0.3200. Further, rank-correlations add up to: US-dollar 0.4645, GB-pound 0.3488 and JP-yen 0.2730.

TABLE 8

Relation between forecasting returns and different risk measures

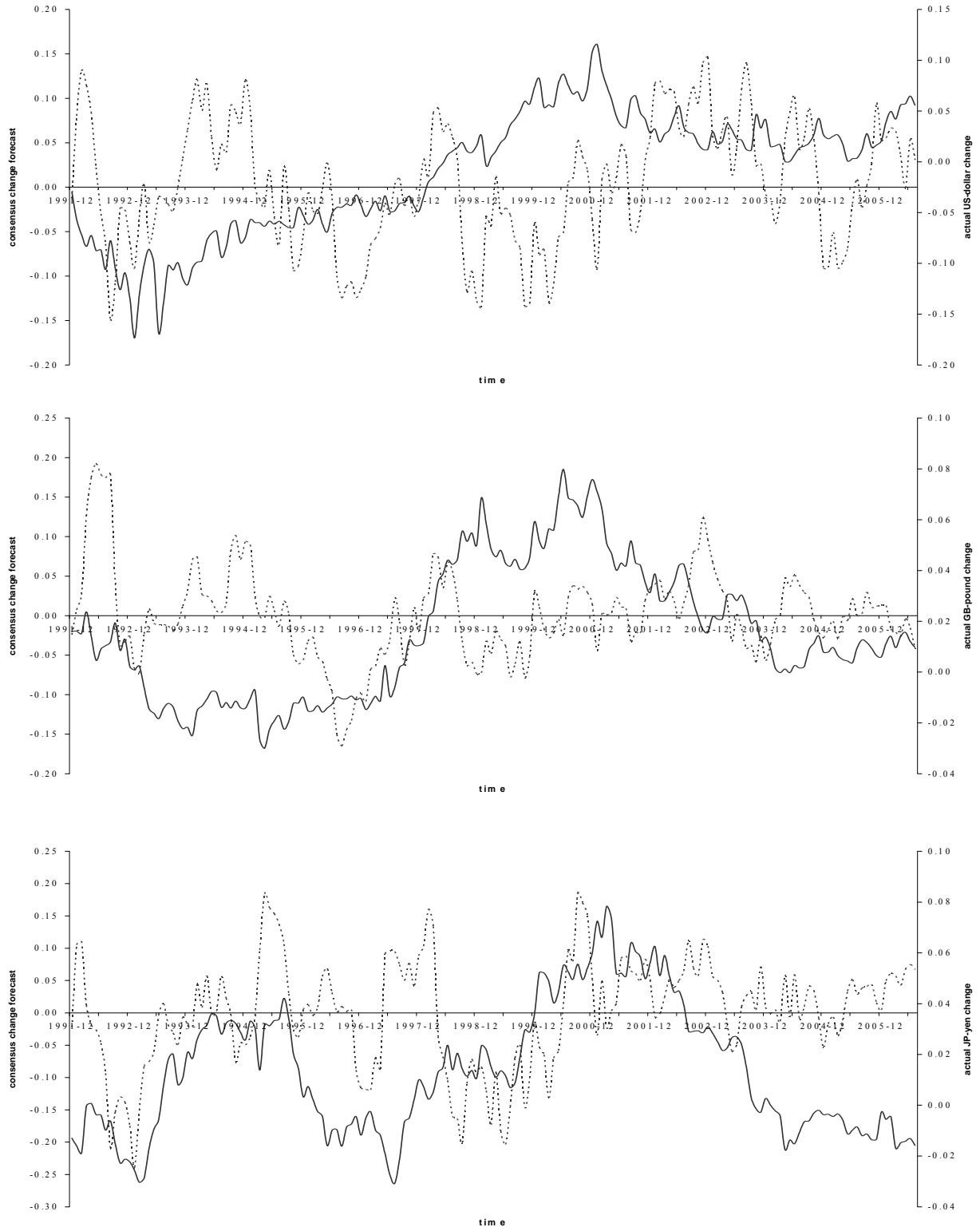
$$fx_{i,t}^{re} = \alpha + \beta \cdot risk_t + e_t$$

	US-dollar		GB-pound		JP-yen	
	t (H ₀ = 0)		t (H ₀ = 0)		t (H ₀ = 0)	
Risk premium						
max.	0.1080***	2.6192	0.0973**	2.1456	0.1879***	3.4994
min.	-0.1576***	-3.3103	-0.1342***	-2.6424	-0.1081	-1.5644
Average	-0.0023	-0.1091	-0.0093***	-0.2121	0.0408	0.5734
(+) sig. 5% level	1		1		6	
(-) sig. 5% level	7		8		0	
Dispersion						
max.	0.3353***	3.8069	0.2808***	2.8952	0.5671***	3.1819
min.	-0.1843	-1.5935	-0.3453**	-2.5681	-0.7124***	-3.1410
Average	0.0894	-0.1091	0.0184	0.1016	0.0145	0.1069
(+) sig. 5% level	17		4		5	
(-) sig. 5% level	0		3		4	
Lagged FX						
max.	0.1753***	3.1565	0.1561**	2.5479	0.1682***	2.7100
min.	-0.1753***	-2.8210	-0.1230***	-2.8294	-0.1770***	-2.6557
Average	-0.0121	-0.1916	-0.0286	-0.4700	-0.0051	-0.0922
(+) sig. 5% level	4		2		4	
(-) sig. 5% level	8		14		6	
Forward Prem.						
max.	6.6624***	3.5010	2.3297**	2.4488	10.3996***	4.0609
min.	-1.8349**	-2.3892	-4.9604***	-2.7322	-7.4488**	-2.1439
Average	0.7425	0.4998	0.0326	0.0129	2.4710	0.6639
(+) sig. 5% level	10		1		10	
(-) sig. 5% level	1		5		1	

Notes: For each forecaster we run Newey-West regressions using 5 lags separated for the US-dollar/-, GB-pound/- and JP-yen/euro (in total 153, 146 and 146, respectively). We regress respective individual profits on five different risk proxies, i.e. the risk premium, dispersion, 1-month lagged exchange rate change and the forward premium, respectively. The risk premium is the difference between the forward rate and the consensus, i.e. the average expected exchange rate (see Frankel and Froot, 1987), whereas dispersion, i.e. expectation heterogeneity, is the cross section variability in the individual forecasts. Asterisks refer to the regressors' level of significance in the short-term relations: *, **, *** to ten, five and one percent.

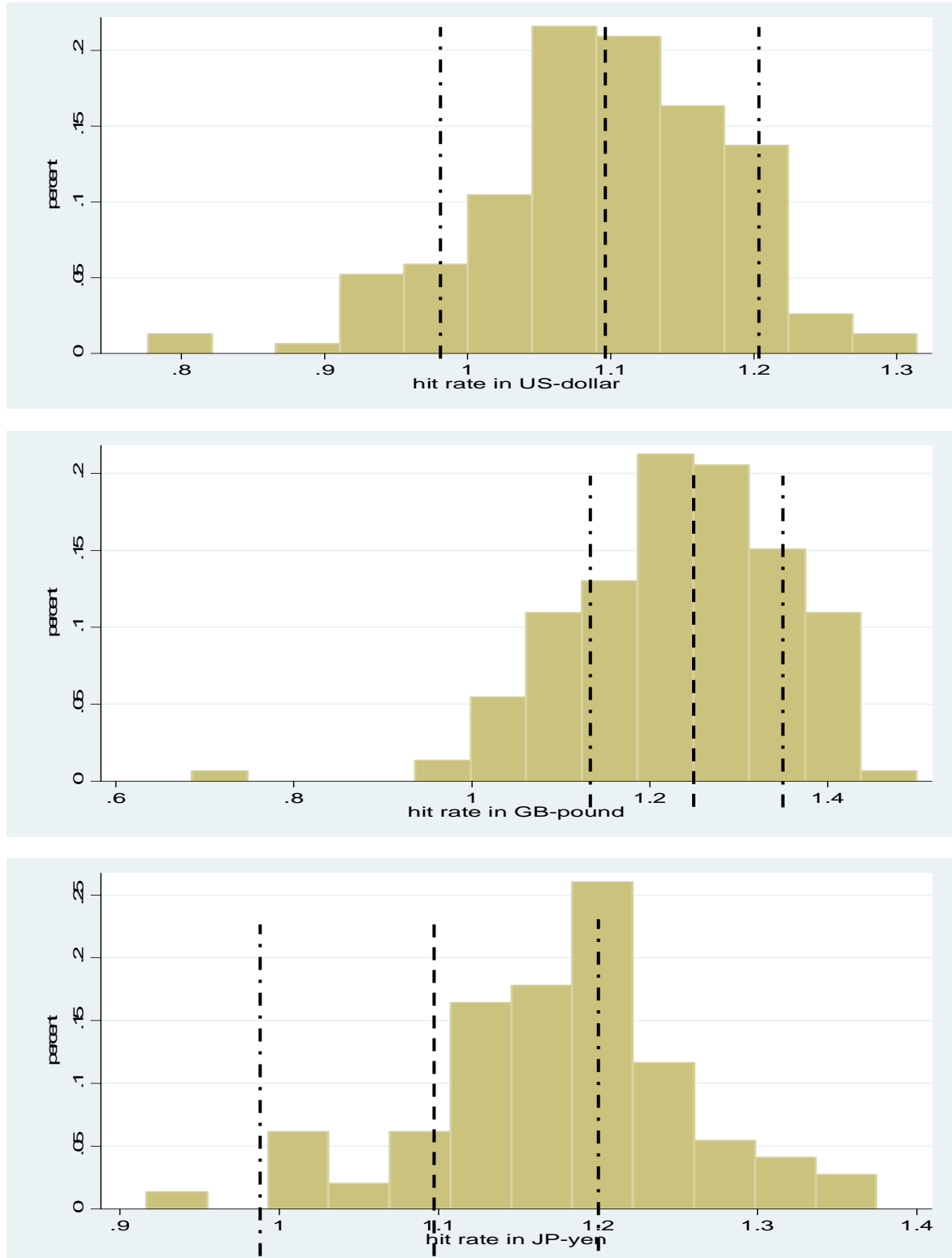
Figures

F 1 Consensus forecasts and corresponding 6-month exchange rate changes



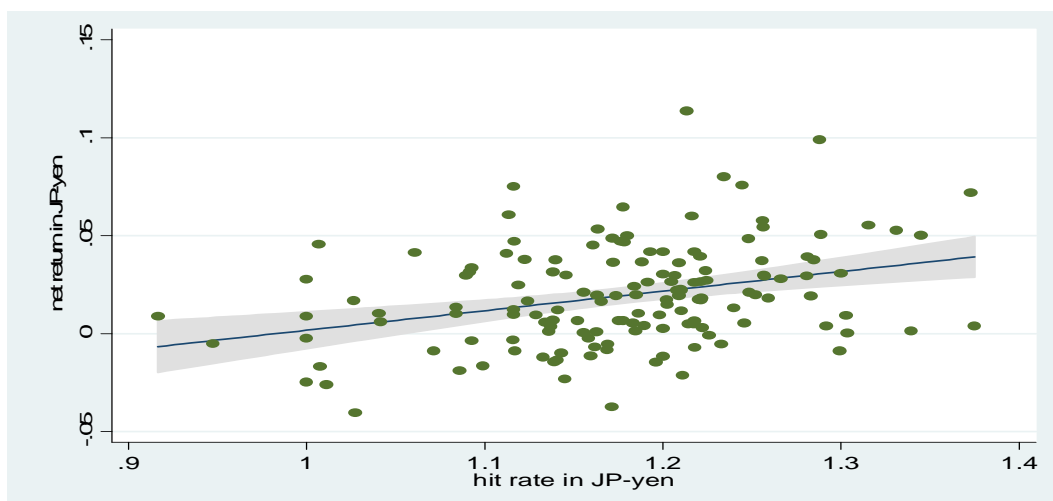
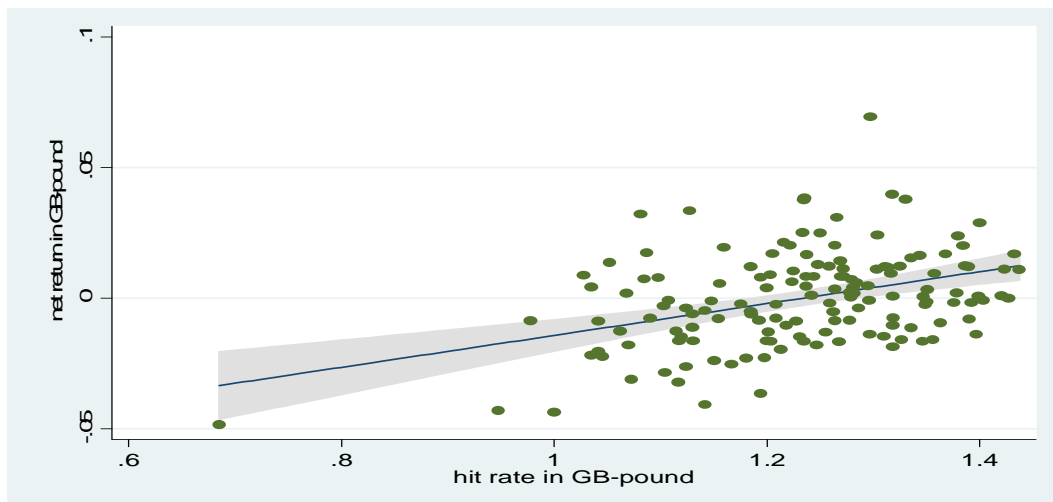
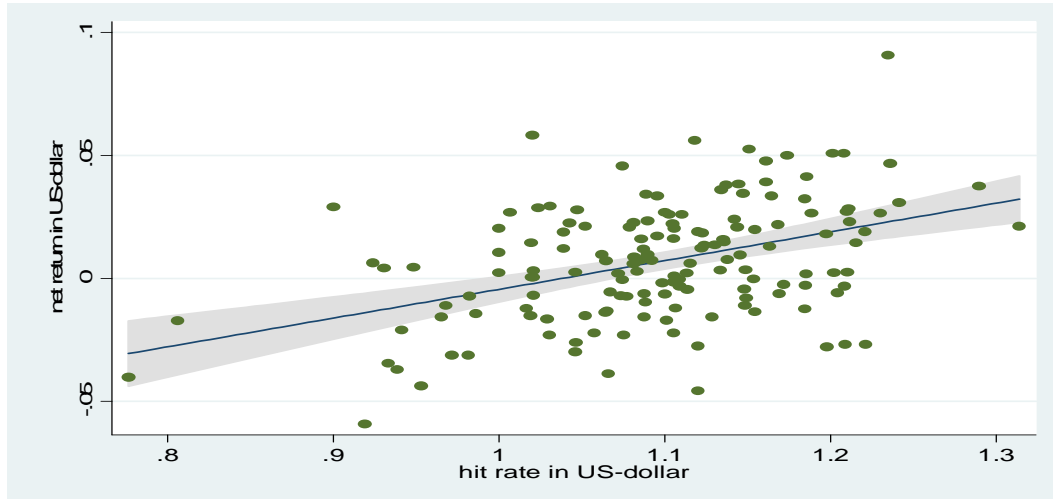
Notes: The sample contains 176 monthly observations from December 1991 until July 2006. The graphs show the time series of consensus (solid lines) and related spot rate changes (dashed lines) separately – moving from top to bottom – in the US-dollar/-, GB-pound/- and JP-yen/euro (until 1998, /D-Mark).

F 2 Histograms of hit rates on exchange rate forecasts



Notes: The hit rates are based upon the individual 6-month US-dollar/euro expectations from Dec. 1991 until July 2006. The hit rates range from 0 to 2, with 2 representing a correct qualitative forecast, 1 a simple forecast error (e.g. forecast of no change but actually the exchange rate falls) and 0 a double forecast error (e.g. forecast of a rise but actually the exchange rate falls). This histograms show the distributions of the hit rates, when moving from top to bottom, in the US-dollar, GB-pound and JP-yen (each versus euro, with the D-Mark until 1998 respectively), considering only participants with at least 50% participation (see further information in Table 1). The dashed lines represent the mean (the middle one) as well as the borders of the 95%-confidence interval (the left and the right ones) of corresponding bootstrap results for the respective exchange rates, respectively.

F 3 Scatter plots of (yearly) net returns and corresponding hit rates



Notes: The (yearly) net returns are based upon the individual 6-month US-dollar/euro expectations from Dec. 1991 until July 2006. The scatter-plots show the net returns and corresponding hit rates, when moving from top to bottom, in the euro/US-dollar, euro/GB-pound and euro/JP-yen (each with the D-Mark/- respectively), considering only participants with at least 50% participation (see further information in Table 1). The returns range for the US-dollar from -0.0592 to 0.0908, for the GB-pound from -0.0484 to 0.0695 and for the JP-yen from -0.0404 to 0.1138.

Appendix

A1

Descriptive statistics of aggregated forecasts and actual exchange rates

	US-dollar/euro	GB-pound/euro	JP-yen/euro
6-m. Δ FX	0.0031	0.0003	-0.0009
std. 6-m. Δ FX	0.0701	0.0581	0.0838
min. 6-m. Δ FX	-0.1488	-0.1649	-0.2392
max. 6-m. Δ FX	0.1459	0.1922	0.1847
Δ consensus	-0.0093	0.0153	0.0130
std. Δ consensus	0.0602	0.0260	0.0254
min. Δ consensus	-0.1736	-0.0300	-0.0308
max. Δ consensus	0.1155	0.0800	0.0783
Δ dispersion	0.0698	0.0412	0.0431
std. Δ dispersion	0.0161	0.0108	0.0089
min. Δ dispersion	0.0431	0.0234	0.0278
max. Δ dispersion	0.1323	0.0860	0.0795

Notes: The sample contains 176 monthly observations from December 1991 until July 2006. All variables are measured in changes and therefore marked with the symbol Δ . Using Carlson and Parkin's (1975) quantification procedure and setting a 3% threshold, we generate consensus forecasts. Asterisks refer to the significance level: *, **, *** to ten, five and one percent.

A 2

Bias and hit rate in consensus exchange rate forecasts

$$\Delta error_t = \alpha + \varepsilon_t$$

$$\text{with } \Delta error_t = \Delta fx_t - \Delta fx_{t,t-6}^e$$

	US-dollar/euro	GB-pound/euro	JP-yen/euro
sub periods			
(1) bias (6-months) [prob. value] [#]	0.0751*** [0.000]	0.0120 [0.556]	-0.0191 [0.304]
(2) bias (6-months) [prob. value] [#]	-0.0487*** [0.021]	-0.0518*** [0.000]	-0.0417 [0.118]
(3) bias (6-months) [prob. value] [#]	0.1667 [0.298]	-0.0062 [0.522]	0.0210** [0.036]
total period			
bias (6-months) [prob. value] [#]	0.0146 [0.318]	-0.0154 [0.121]	-0.0143 [0.248]
hit rate	0.5833* [0.086]	0.4808* [0.099]	0.6418* [0.083]

Notes: The sample contains 176 monthly observations from December 1991 until July 2006. The sub periods are defined as follows: (1) 1992-1995, (2) 1996-2000 and (3) 2001-2006. The errors underlying the consensus bias are based upon the difference between the actual exchange rate change (1-month and 6-month, respectively) and the related consensus forecast. A positive (negative) sign means that the foreign currency-strength is on average overrated (underrated). Using Carlson and Parkin's (1975) quantification procedure and setting a 3% threshold, we generate consensus forecasts. All regressions are estimated by Newey-West standard-errors. The hit rate shows the share of accurate direction forecasts and is evaluated via a χ^2 -test, which is based under the null hypothesis that the forecast accuracy is purely chance, i.e. a hit rate of 0.5. Asterisks refer to the significance level: *, **, *** to ten, five and one percent.

A 3

Exchange rate hit rate statistics

	mean	median	25%-q.	75%-q.	max.	min.	std. dev.
US-dollar	1.0966	1.1062	1.0443	1.1534	1.2566	0.7979	0.0773
GB-pound	1.2231	1.2339	1.1411	1.3016	1.4494	0.7191	0.1169
JP-yen	1.1960	1.2114	1.1465	1.2473	1.3775	0.9368	0.0833

Notes: This table shows descriptive statistics for the exchange rate hit rate series employed in the empirical analysis separately for the US-dollar/-, GB-pound/- and JP-yen/euro (-/D-mark respectively).

A 4

Best and worst 25 forecasters in the US-dollar (hit rates)

strong hands	[probability]	participation	weak hands	[probability]	participation
1.31 ^{***}	[0.001]	101	0.76 ^{***}	[0.000]	93
1.28 ^{***}	[0.002]	151	0.81 ^{***}	[0.000]	98
1.24 ^{**}	[0.021]	97	0.91 ^{***}	[0.001]	149
1.24 ^{**}	[0.021]	169	0.92 ^{***}	[0.003]	135
1.23 ^{**}	[0.032]	88	0.92 ^{***}	[0.004]	119
1.22 ^{**}	[0.036]	125	0.93 ^{***}	[0.005]	144
1.22 [*]	[0.051]	153	0.93 ^{***}	[0.005]	101
1.21 [*]	[0.053]	112	0.93 ^{***}	[0.005]	146
1.21 [*]	[0.059]	118	0.93 ^{***}	[0.006]	153
1.21 [*]	[0.061]	166	0.94 ^{***}	[0.009]	134
1.21 [*]	[0.062]	157	0.95 ^{**}	[0.012]	169
1.21 [*]	[0.067]	96	0.97 ^{**}	[0.030]	94
1.21 [*]	[0.074]	141	0.97 ^{**}	[0.035]	140
1.20 [*]	[0.078]	142	0.97 ^{**}	[0.036]	143
1.20 [*]	[0.079]	103	0.98 [*]	[0.052]	108
1.20 [*]	[0.079]	108	0.98 [*]	[0.053]	113
1.20 [*]	[0.080]	172	0.99 [*]	[0.062]	143
1.20 [*]	[0.097]	91	1.00	[0.102]	108
1.20 [*]	[0.100]	142	1.00	[0.102]	147
1.20	[0.104]	148	1.01	[0.127]	153
1.19	[0.113]	93	1.01	[0.127]	151
1.19	[0.137]	96	1.01	[0.156]	154
1.18	[0.149]	119	1.01	[0.159]	147
1.18	[0.151]	130	1.01	[0.159]	146
1.18	[0.158]	153	1.01	[0.161]	142

Notes: The results are based upon 176 monthly observations from December 1991 until July 2006. We assume a 6-month forecast horizon and set the expectation threshold at 3%. The hit rates range from 0 to 2, with 2 representing a correct qualitative forecast, 1 a simple forecast error (e.g. forecast of no change but actually the exchange rate falls) and 0 a double forecast error (e.g. forecast of a rise but actually the exchange rate falls). The minimum participation is set at 50%, so we consider only individuals who participated at least at 88 months with sample mean participation of 133, minimum of 88 and maximum of 173 months. Though, 152 individuals remain with a mean hit rate of 1.09. Corresponding p-values of the realized hit rates are calculated via bootstrap technique. Asterisks refer to the regressors' level of significance in the short-term relations: *, **, *** to ten, five and one percent.

A 5

Best and worst 25 forecasters in the GB-pound (hit rates)

strong hands	[probability]	participation	weak hands	[probability]	participation
1.45 ^{***}	[0.000]	150	0.95 ^{***}	[0.000]	89
1.43 ^{***}	[0.000]	104	1.00 ^{***}	[0.000]	96
1.43 ^{***}	[0.000]	88	1.03 ^{***}	[0.000]	90
1.42 ^{***}	[0.001]	111	1.03 ^{***}	[0.000]	93
1.42 ^{***}	[0.001]	137	1.04 ^{***}	[0.000]	109
1.41 ^{***}	[0.002]	118	1.04 ^{***}	[0.000]	119
1.40 ^{***}	[0.002]	89	1.05 ^{***}	[0.000]	115
1.40 ^{***}	[0.003]	140	1.06 ^{***}	[0.000]	171
1.40 ^{***}	[0.003]	153	1.07 ^{***}	[0.000]	97
1.39 ^{***}	[0.004]	152	1.07 ^{***}	[0.000]	110
1.39 ^{***}	[0.005]	145	1.07 ^{***}	[0.000]	151
1.39 ^{***}	[0.006]	118	1.07 ^{***}	[0.000]	161
1.39 ^{***}	[0.006]	144	1.09 ^{***}	[0.000]	146
1.39 ^{***}	[0.007]	168	1.09 ^{***}	[0.000]	156
1.38 ^{***}	[0.009]	107	1.09 ^{***}	[0.000]	96
1.38 ^{**}	[0.011]	132	1.10 ^{***}	[0.001]	94
1.38 ^{**}	[0.011]	103	1.10 ^{***}	[0.001]	125
1.37 ^{**}	[0.017]	113	1.11 ^{***}	[0.001]	100
1.37 ^{**}	[0.024]	156	1.11 ^{***}	[0.001]	110
1.36 ^{**}	[0.031]	150	1.12 ^{***}	[0.002]	91
1.35 ^{**}	[0.045]	173	1.12 ^{***}	[0.003]	136
1.35 ^{**}	[0.046]	159	1.12 ^{***}	[0.003]	143
1.35 ^{**}	[0.046]	125	1.12 ^{***}	[0.003]	102
1.35 ^{**}	[0.047]	145	1.13 ^{***}	[0.005]	141
1.35 [*]	[0.054]	172	1.13 ^{***}	[0.005]	113

Notes: The results are based upon 176 monthly observations from December 1991 until July 2006. We assume a 6-month forecast horizon and set the expectation threshold at 3%. The hit rates range from 0 to 2, with 2 representing a correct qualitative forecast, 1 a simple forecast error (e.g. forecast of no change but actually the exchange rate falls) and 0 a double forecast error (e.g. forecast of a rise but actually the exchange rate falls). The minimum participation (part.) is set at 50%, so we consider only individuals who participated at least at 88 months with sample mean participation of 132, minimum of 88 and maximum of 173 months. Though, 146 individuals remain with a mean hit rate of 1.23. Corresponding p-values of the realized hit rates are calculated via bootstrap technique. Asterisks refer to the regressors' level of significance in the short-term relations: *, **, *** to ten, five and one percent.

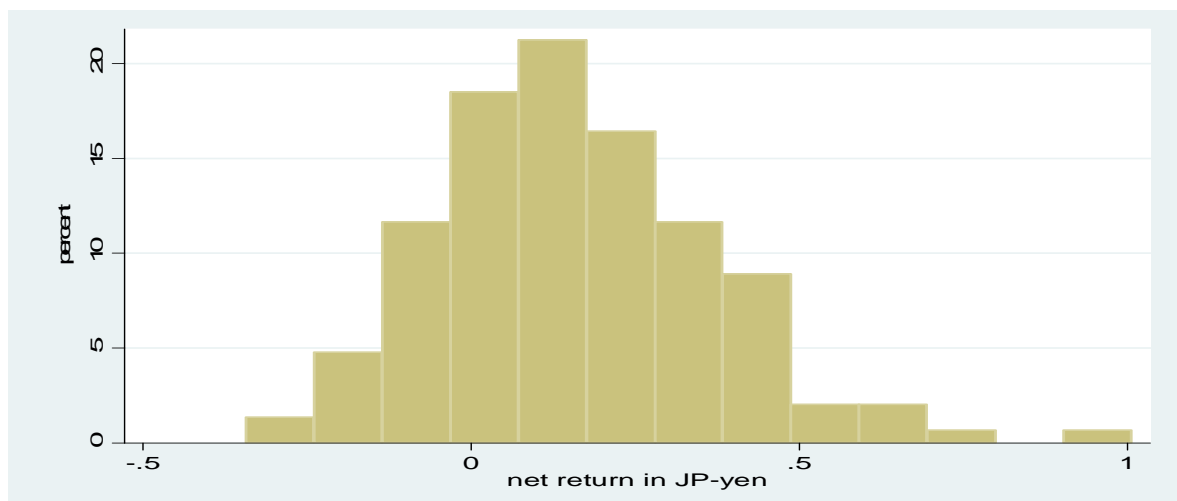
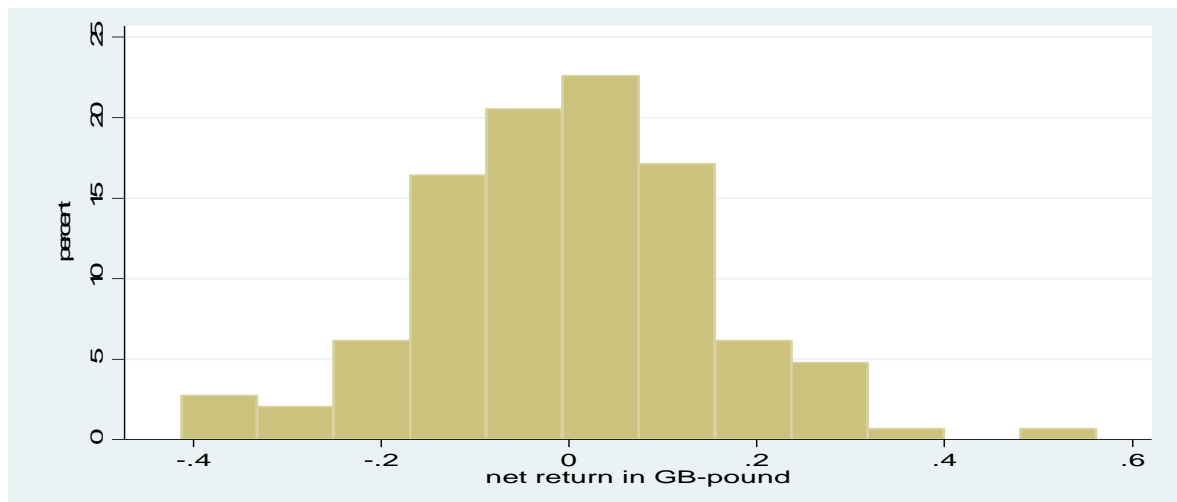
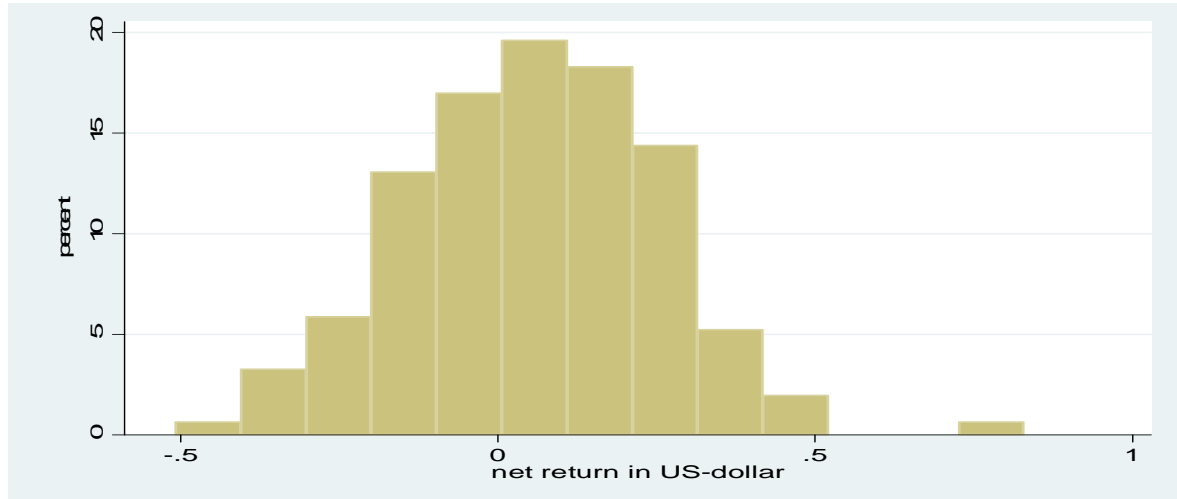
A 6

Best and worst 25 forecasters in the JP-yen (hit rates)

strong hands	[probability]	participation	weak hands	[probability]	participation
1.37***	[0.000]	151	0.91***	[0.001]	107
1.37***	[0.000]	125	0.95**	[0.011]	95
1.35***	[0.000]	144	1.00	[0.105]	97
1.34***	[0.000]	161	1.00	[0.105]	102
1.33***	[0.000]	141	1.00	[0.105]	130
1.32***	[0.000]	88	1.01	[0.129]	169
1.32***	[0.000]	130	1.01	[0.153]	90
1.30***	[0.000]	151	1.01	[0.165]	150
1.30***	[0.000]	159	1.01	[0.173]	135
1.30***	[0.000]	156	1.03	[0.250]	111
1.29***	[0.000]	143	1.03	[0.299]	150
1.29***	[0.001]	167	1.03	[0.301]	119
1.29***	[0.001]	142	1.05	[0.446]	122
1.29***	[0.001]	118	1.06	[0.574]	132
1.29***	[0.001]	105	1.07	[0.709]	98
1.28***	[0.001]	152	1.09	[0.958]	156
1.28***	[0.002]	145	1.09	[0.966]	166
1.27***	[0.004]	168	1.09	[0.982]	153
1.26***	[0.004]	108	1.09	[0.983]	142
1.26***	[0.004]	139	1.09	[0.997]	162
1.26***	[0.004]	147	1.09	[0.997]	108
1.26***	[0.004]	89	1.09	[0.997]	108
1.26***	[0.004]	117	1.10	[0.899]	90
1.26***	[0.004]	121	1.11	[0.789]	111
1.25***	[0.006]	172	1.11	[0.717]	141

Notes: The results are based upon 176 monthly observations from December 1991 until July 2006. We assume a 6-month forecast horizon and set the expectation threshold at 3%. The hit rates range from 0 to 2, with 2 representing a correct qualitative forecast, 1 a simple forecast error (e.g. forecast of no change but actually the exchange rate falls) and 0 a double forecast error (e.g. forecast of a rise but actually the exchange rate falls). The minimum participation (part.) is set at 50%, so we consider only individuals who participated at least at 88 months with sample mean participation of 132, minimum of 88 and maximum of 172 months. Though, 146 individuals remain with a mean hit rate of 1.18. Corresponding p-values of the realized hit rates are calculated via bootstrap technique. Asterisks refer to the regressors' level of significance in the short-term relations: *, **, *** to ten, five and one percent.

A 7 Histograms of average 1-month returns on exchange rate forecasts



Notes: The average 1-month returns are based upon the individual 6-month US-dollar/euro expectations from Dec. 1991 until July 2006. These returns are net of respective interest rates, so taking into account related interest rate spreads. This histograms show the distributions of the returns, when moving from top to bottom, in the euro/US-dollar, euro/GB-pound and euro/JP-yen (each with the D-Mark/- respectively), considering only participants with at least 50% participation (see further information in Table 1). The returns range for the US-dollar from -0.51 to 0.73, for the GB-pound from -0.41 to 0.56 and for the JP-yen from -0.34 to 0.90.

A 8

Best and worst 25 forecasters in the US-dollar (annualized average returns)

winning hands	[probability]	participation	loosing hands	[probability]	participation
0.0908***	0.0000	98	-0.0592***	0.0039	136
0.0583***	0.0078	148	-0.0456**	0.0267	133
0.0561**	0.0103	110	-0.0436**	0.0343	170
0.0525**	0.0164	106	-0.0402*	0.0514	94
0.0509**	0.0199	173	-0.0388*	0.0604	91
0.0508**	0.0202	149	-0.0371*	0.0724	147
0.0500**	0.0221	161	-0.0346*	0.0943	120
0.0478**	0.0286	149	-0.0311	0.1320	141
0.0468**	0.0322	89	-0.0311	0.1325	108
0.0458**	0.0358	134	-0.0298	0.1488	109
0.0413*	0.0582	97	-0.0277	0.1800	91
0.0392*	0.0724	118	-0.0274	0.1855	158
0.0384*	0.0779	90	-0.0267	0.1964	158
0.0380*	0.0810	153	-0.0267	0.1968	113
0.0375*	0.0850	152	-0.0261	0.2062	151
0.0362*	0.0965	112	-0.0231	0.2639	93
0.0346	0.1117	136	-0.0230	0.2656	163
0.0342	0.1159	124	-0.0222	0.2832	104
0.0336	0.1227	157	-0.0221	0.2838	95
0.0335	0.1236	170	-0.0211	0.3074	154
0.0324	0.1362	168	-0.0172	0.4039	98
0.0309	0.1562	170	-0.0169	0.4121	148
0.0294	0.1769	130	-0.0164	0.4259	103
0.0290	0.1826	150	-0.0156	0.4492	144
0.0287	0.1869	126	-0.0156	0.4497	148

Notes: The results are based upon 176 monthly data sets from December 1991 until July 2006. Above captioned numbers are expressed in average annual returns, i.e. the best individual forecaster shows an average return of 9.08% per year. The minimum participation is set at 50%, so we consider only individuals who participated at least at 88 months. This leads to a reduced sample of 153 individuals with a mean participation of 133 as well as minimum and maximum of 88 and 174 months, respectively. The related mean return amounts to 0.66% per year. Corresponding p-values are calculated via bootstrap technique. Asterisks refer to the regressors' level of significance in the short-term relations: *, **, *** to ten, five and one percent.