Professional forecasters' overconfidence and underconfidence: Evidence from foreign exchange

by Oliver Gloede and Lukas Menkhoff

Abstract

This paper examines professionals' over- and underconfidence in their forecasting performance. We use a dataset with unique comprehensive information in foreign exchange. The forecasters in our data set are indeed overconfident. However, there is no relation between their individual self-rating and effective performance, which seems surprising for professionals. So, bad forecasters are on average more overconfident. Interestingly, there are also underconfident professionals. In explaining over- and underconfidence, we find an easing effect from experience and market alignment on overconfidence but a negative effect from recent forecasting successes. The story is extended by plausible relations between job positions and over-/underconfidence.

JEL-Classification: F 31 (int'l finance), D 84 (expectations), G 1 (general fin markets)
Keywords: foreign exchange, forecasting, overconfidence, underconfidence, better-than-average, performance, self-rating

February 26, 2009

We thank participants of the seminar at the Leibniz University Hannover for their helpful comments. In addition, we thank the Centre for European Economic Research (ZEW) in Mannheim, Germany, for providing the data. Oliver Gloede gratefully acknowledges a Ph.D. scholarship of the Foundation of German Business (Stiftung der Deutschen Wirtschaft).

Oliver Gloede, Lukas Menkhoff; Department of Economics, Leibniz Universitaet Hannover, Koenigsworther Platz 1, D-30167 Hannover, Germany; e-mail: gloede@gif.uni-hannover.de, menkhoff@gif.uni-hannover.de; tel.: +49-511-762-19960, fax: +49-511-762-4796

Professional forecasters' overconfidence and underconfidence: Evidence from foreign exchange

1 Introduction

We know from a series of by now "classical" studies that most people are overconfident regarding their own abilities, such as their driving performance (Svenson, 1981). One reason may be that people do not get regular feedback on their true performance so that unrealistic believes about the own skills can easily survive (Fischhoff et al., 1977). However, there is also evidence that financial market professionals, who repeatedly and precisely learn their performance on short-term intervals, are overconfident (Ben-David et al. ,2007; Oberlechner and Osler, 2008). This is puzzling by itself and becomes even more puzzling if we consider that overconfidence is costly, for example due to its effect of increasing useless trading and lowering profits (Barber and Odean, 2000). Besides, on the individual level there also appears underconfidence which has been scarcely addressed, e.g. Kirchler and Maciejovsky (2002). But the negative effects for market efficiency of underconfidence could be even as severe as of overconfidence (Moore and Cain, 2007). Due to those arguments, overconfidence and underconfidence among financial professionals deserves further research.

We contribute towards a better understanding of overconfidence and underconfidence among professional forecasters by analyzing three aspects on individual level. First, we measure overconfidence more precisely than often possible. It seems important that overconfidence really captures the difference between self-rating and the same person's true performance. This precise measure is often not available, either because data cannot be analyzed on an individual basis or because there are not both items available, i.e. self-rating and performance. Second, we examine the relation between overconfidence and two related measures, i.e. self-rating and a more general misperception of one's own performance; we show differences between these three related measures. Third, we use forecasters' demographic characteristics, job characteristics and forecasting characteristics in explaining overconfidence and underconfidence.

In order to conduct these analyses we use a sample of 105 professional forecasters who are regularly contributing to the ZEW financial market survey. From this survey we take the monthly forecasts on the US-Dollar exchange rate (versus the Euro, respectively Deutsche Mark) over several years so that we can calculate a meaningful forecasting performance. We complement this performance information by data from additional surveys being conducted at the time of the regular surveys. These supplementary surveys inform about detailed demographic characteristics and the self-rating abilities of these professionals. We document that our sample of forecasters is quite “normal” in several dimensions, in particular as these persons regard themselves mostly better in their self-rating than their fellow forecasters and thus clearly show some overconfidence.

Our further analyses reveal the following: first, the mentioned better-than-average self-rating of professional forecasters is not significantly related to the same person’s (true) overconfidence. As overconfidence is just the difference between self-rating (being better than the average) and performance, our finding implies that – among its two components – overconfidence is indeed mainly driven by low performance. Second, as this effect applies to good and bad forecasters as well, we do not only find overconfidence, but interestingly also a large degree of underconfidence. Third, the difference between overconfidence and underconfidence is highlighted by the finding that the determinants for these two measures differ in their effect on the both: for the average forecaster we observe a similar effect of being a fund manager on overconfidence as well as underconfidence. In contrast the effects of experience and a positive development in recent forecasting success show large differences between over- and underconfidence.

The paper is structured as follows: Section 2 introduces into the data used, Section 3 shows relations between self-rating and performance as well as overconfidence and performance. Section 4 examines determinants of these imperfect self-rating measures and Section 5 concludes.

2 Data

2.1 The ZEW dataset

We use individual survey data of the Financial Market Survey (Finanzmarkttest), which is conducted by the ZEW, Mannheim. Aggregate statistics are derived from the individual data set and are published in financial media like Reuters or Bloomberg as wells as from the ZEW itself. Like comparable datasets, e.g. the Consensus Forecasts London, the majority of the participants are employed in the banking sector (75 %). Others work in the insurance sector (15 %) or in large industry enterprises (10 %). Overall there are 400 financial experts asked to participate monthly in the survey of which on average about 350 answer frequently.

The financial experts give individual forecasts for the exchange rate over a 6-month horizon of the USD/EUR, or of the USD/DEM exchange rate for the time period before December 1998 respectively. The observations range from December 1991 to October 2008. This gives us a maximum of almost 17 years of monthly data on individual expectations. The forecasts are qualitative and indicate whether the exchange rate is expected to appreciate, depreciate or not change. From these regular forecasts we retrieve average hit rates which approximate the true skill level of the forecasters (for the exact procedure see Section 3). To obtain a reliable and valid criterion for the true skill level we solely use observations from forecasters which participated in the survey for at least 36 times, i.e. for minimum of three years if they participate every single month.

In addition to this, several special surveys were conducted contemporaneously with the regular monthly surveys. Therewith we observe demographic and job characteristics, e.g. age, experience, job position, responsibility for staffs and/or output, etc. Since we use these personal characteristics in our analysis we want to make sure that each observation corresponds to one person. We follow all changes in the contact persons and employers and use solely data which refer to the very same person. This provides us with a highly consistent micro data set. The “price” that we have to pay for this is that we end up with a sample of 105 professional forecasters for which we have complete observations, i.e. enough forecasts as well as information about their self-rating as forecasters, demographic and job information.

2.2 Sample characteristics

The descriptive statistics for the sample of 105 professionals are shown in Table 1. The average forecaster is male, in his mid age, and holds a university degree. He has been working for many years in the financial market and holds a senior position, i.e. with operative and personnel responsibilities. It is most likely that he works as fund manager. His forecasts are mainly the results of fundamental analyses and he has not been able to improve his forecasting performance for the last three years. Most of the time he stays in line with the dominant market expectation. Though the average forecaster is already in a senior position, which gives an impression of the characteristics of many of the forecasters, we observe the whole range for the attributes.

The forecasters show the same behavioral biases as other people which we demonstrate in two exercises. On the one hand, the three often suggested dimensions of overconfidence are not closely related. On the other hand, we observe the same better-than-average effect in the self-rating as other studies do.

Overconfidence refers to the overestimation of the own abilities. This can include the misperceived relative ability to others (better-than-average effect, Greenwald, 1980), the overestimated precision of the ability (miscalibration bias, Fischoff et al., 1977), or the belief of influence capability (illusion-of-control, Langer, 1975). To measure the better-than-average effect we asked the financial experts in two surveys (April 2007 and October 2008) the following question: „How do you evaluate your USD/EUR-forecast compared to the average forecasting hit rate of all participants of the ZEW financial market survey?“. The scale ranged from 1 to 21. The questions for miscalibration was posed in the survey of October 2008. Participants were asked for a 90-% confidence interval for the 6-month future USD/EUR exchange rate. Data for the illusion-of-control was collected in October 2008 and was extracted from the following question, which has been used for this purpose (Menkhoff et al., 2006): "Most of the published business news does not surprise me at all." Respondents answered on a scale ranging from 1 to 20. As Table 2 shows these criteria are loosely related. This is a standard finding in the literature (e.g. Klayman et al. 1999, Régner et al. 2004, Glaser et al. 2005, Glaser and Weber 2007a) and implicitly demonstrates that our sample behaves "well".

For the further analysis we use the better-than-average self rating. This is the best way to address the raised issue of true overconfidence: the relation between self-rating and performance. Our observation for the better-than-average rating is quite general and similar to those of other studies. When we asked participants to rate their forecasting performance in relation to the other participants of the ZEW survey, experience of other similarly designed studies shows that there is a tendency to overrate one's own performance. This is usually interpreted as overconfidence of the total group (Larrick et al., 2007). Indeed, we find this behavior in our sample as well, as Figure 1 shows. We therefore think that our data for better-than average is not so different from other samples.

However, this aggregate analysis of overconfidence cannot be taken to the level of individual data because some overconfident respondents may be truly better than their competitors and thus only appear overconfident. At the same time there may be participants who seem modest by rating themselves as average but who are in fact relatively bad forecasters and are thus truly overconfident. So, a measure of individual overconfidence profits from individual information about real forecasting performance. Fortunately, we have this information and can match it with the self-rating. As mentioned above we use the self-evaluation of the forecasting performance relative to the average forecaster.

3 Descriptive analysis

3.1 Self-rating and performance

Our measure of individual overconfidence is the difference between the self-rating and the effective performance. We receive the information about self-rating from the survey, but we have to calculate the measure of performance which requires some decision to make.

(1) We consider all forecasts of a person. (2) We exactly calculate whether the forecast was right or wrong. In this respect, the survey participants have a time window of about two weeks to submit their forecasts. To achieve a maximum of accuracy and consistency we use individual forecasting windows, i.e. we compare the forecasted change of the exchange rate to the realized exchange rate in exactly six months for each individual separately. This provides us with a precise measure of how often the individual forecast hits the actual exchange movement. (3) Due to the trivariate nature of forecasts (up, down, unchanged) we need to make a decision as to how much of change is compatible with an "unchanged" exchange rate. Fortunately, we can use information directly from the forecasters themselves. In a special survey in 2006 they state that on average a 3% change of the exchange rate over 6-months is considered to be no change. (4) Since the expectations are qualitative forecasts usual error measures on an individual basis, e.g. RMSE, are not computable. Instead we apply the often used measure of hit rates (e.g. Granger and Pesaran, 2000). For this purpose we convert the continuous exchange rate process into a discrete process which corresponds with the forecast categories of appreciation, depreciation and no change. (5) To incorporate that the experts can choose between three options, a hit rate is coded in three categories: large deviation, small deviation and no deviation of forecast from the true process. Large deviations are predictions which indicate the opposite direction of the actual movement, whereas small deviations are expectations which are neither a correct forecast nor a large deviation. The code values are 0-1-2 in a way that a higher hit rate is better. This performance measure has been used for ZEW exchange rate forecasts and reveals that forecasts are on average close to random forecasts (Menkhoff et al., 2008) and heterogeneous on the individual level (MacDonald et al., 2008).

In Figure 2 we plot the 105 forecasters' self-rating against their true performance which is again measured via hit rates. There is no obvious relation between self-rating and performance recognized This is quite surprising since overconfidence leads to excessive trading and significant lower profits (Odean, 1998; Barber and Odean, 2001; Biais et al., 2005). But such independence has been shown, see e.g. Ackerman et al. (2001) and for an overview Moore (2007). Glaser and Weber (2007b) argue that inexperienced private investors do not know their past portfolio performance and therefore cannot give an appropriate evaluation of their skills. For professionals the result remains surprising. Professionals in the financial sector do receive frequent feedback either from the market or the customer and moreover their salary is based on their performance in many cases. Hence, one might expect that professionals are better able to assess their own relative performance. However, we find that even professionals are unable to assess their relative forecasting performance.

3.2 Overconfidence and performance

The result from Section 3.1 provides already a clear expectation on the relation between self-rating and true overconfidence. As people cannot correctly asses their relative performance, overconfidence, which is measured as the difference between self-rating and effective performance, will be driven by the effective performance. This is indeed what the plot between overconfidence and forecasting performance in Figure 3 shows. The average level of overconfidence is 0.72, which is in line with our hypothesis of an overconfident sample and corresponds to former research which found that overconfidence appears on financial markets, e.g. Chuang and Lee, 2006. On contrast, it remains puzzling that professionals are overconfident, even though they receive frequent and clear feedback. We observe values ranging from -13 to 14, which means that on an individual basis professionals exhibit both, overconfidence as well as underconfidence. The good forecasters are rather underconfident, whereas the relatively overconfident forecasters tend to be in effect worse than the average.

In a sense, this may be what one expects in efficient markets. Under efficiency, the universe of forecasters will not be able to produce clearly valuable forecast. This is true here. Moreover, as it is difficult to systematically outperform the market, it will be difficult to assess any systematic component in one's own forecasting performance and thus self-ratings will be biased and largely accidental. This applies here as well.

4 Determinants of overconfidence and underconfidence

4.1 Methodology

We have shown that the professional forecasters are overconfident as well as underconfident, which we want to explain in the following by a set of demographic, job, and forecasting characteristics. We estimate an ordinal regression model for over- and underconfidence (Greene, 2008; Long, 1997). Therefore we define the level of over-/underconfidence as a piecewise-defined function of the difference between self-rating and hit rate:

$$OVC_i = \begin{cases} 1\ if\ SR_i - HR_i > 0 \\ 0\ if\ SR_i - HR_i = 0 \\ -1 if\ SR_i - HR_i < 0 \end{cases} \tag{1}$$

with SR_i as self-rating and HR_i representing hit-rate, each for forecaster i. So, the criterion distinguishes overconfident and underconfident forecasters as well as forecasters who are neither the former nor the ladder. We observe 36 % underconfident and 51 % overconfident forecasters and 12 % who are neither the former nor the ladder.

Using this estimation procedure gives us several advantages compared to the usual ordinary least square analysis. First, as argued above it is important to address overconfidence properly, i.e. consider both, the own self-rating as well as the true skill level. We take this into account by deriving the overconfidence measure from the difference between self-rating and true performance. Second, just using just the directional information of the difference puts the less restrictive assumptions on the data as the cardinal criterion. Third most importantly, estimating the difference by linear estimation procedures does not distinguish between the effect on over- and underconfidence of the regressors which is hypothesized to be asymmetric, i.e. nonlinear. To distinguish between the effect on overconfidence and underconfidence we have to use non linear estimation techniques. We use an ordered logit approach. This indeed allows non-linear us to look for asymmetric effects on over- and underconfidence. To control for heteroscedasticity we use robust variance estimators for all estimations.

4.2 Results

Possible determinants may stem from three sources: first, there may be personal reasons approximated by demographic characteristics (see Chevalier and Ellison, 1999, Edwards and Caglayan, 2001). A possible hypothesis in this respect would be that higher educated persons (here with an academic education) are better able to rate their performance properly. A second group of determinants stems from the job that people occupy. A hypothesis in this respect might be that more responsibility indicates career success and thus ability which might lead to more accurate self-rating. However, an influence may also work in the opposite way as more responsibilities may distract attention away from proper forecasting and its assessment. A third group of determinants is formed on the basis of participants' own forecasting characteristics. A hypothesis here is that a positive trend in forecasting performance may foster confidence and even generate overconfidence. In the end, the validity of such arguments is an empirical issue and the results for ordered logit estimation are shown in Table 3. For interpretation we present effects of marginal changes, and discrete changes for dummy variables respectively, in Table 4.

Our findings are quite robust. Significant results remain significant for the different specifications. For interpretation we use the most restricted model which leaves just the significant results. The robustness of our estimation is underlined by the fact that the estimation of the ordered probit results in the same significant variables, see Appendix 1. The coefficients differ between the two procedures by the usual factor of 1.7.

In the theoretical literature overconfidence is modeled as a process of learning due to biased by self-attribution (e.g. Daniel et al., 1998): recent successes take relatively too large weight for self-evaluation (Bem, 1965 Langer and Roth, 1975 as well as Miller and Ross, 1975). We find evidence for self-attribution bias. We observe positive as well as negative trends in the forecasting performance of the recent past. If self-attribution bias is a reason for overconfidence the positive trend should be significant and the negative should not be significant. Indeed, this is the case. The dummy variable for a positive trend in the hit rate is continuously significant for all model specifications. The analysis for the reference case of an average forecaster shows that the trend has a larger effect on overconfidence than for underconfidence. Becoming a better forecaster results in a 26 % higher chance to be overconfident and in a 19 % lower chance to be underconfident. This finding supports the asymmetry of the effect of successes on over- and underconfidence.

This finding also raises the stakes for the hypothesis that the most recent successes and failures are most important for a high self-evaluation. To support this finding we estimate the model again using the mean hit rate of the last 6 months rather than using the dummy variables for positive and negative trend in the hit rate, see Appendix 2. The effect is also significantly positive which supports the hypothesis further.

The effect of experience is not as obvious as it looks on the first sight. Basically, one would probably assume that while becoming more and more experienced misalignments would decrease. Starting her career a forecaster might not know her true forecasting abilities. But during her career a forecaster collects a lot of experience about successes and failures. Using this information allows her to form rational expectation about her own forecasting performance even when the true skill was unknown in the beginning. But already Brehmer (1980) puts the effect of experience into question and discusses the reasons which might prevent us from learning, like confirmatory evidence and causality assumptions and neglection of negative information. On the other hand, Gervais and Odean (2001) present an environment where experience might lead to lower overconfidence. They use the effect of self-attribution bias to model the effect of experience on overconfidence. On the short run self-attribution bias leads to a larger degree of overconfidence with increasing experience. For the long run it depends on the level of self-attribution bias whether experience decreases overconfidence or remains on an equilibrium level or even explodes for longer time horizons. The forecasters seem to show reasonable levels of self-attribution so experience can have an effect. The effect of working experience in the financial sector is significant on the 5 % level anyway if we do or do not control for age. So, it is important to mention that the experience effect does not stem from just getting older. I our reference situation of an average forecaster ten more years of experience lead to a 14 % lower probability to be overconfident and to a 9 % higher probability to be underconfident.

A related effect to experience is forecasting knowledge which has been shown to have a decreasing impact on another behavioral flaw (Feng and Seasholes, 2005). Two measures of our data set are related to the level of knowledge: Academic education and the use of fundamental analysis. The first could enable forecasters to better know appropriate and more sophisticated models of forecasting. Since they might better understand the full properties of the models and their forecasting power they should be neither overconfident nor underconfident. In fact, we cannot show any significant impact of academic education on any of the two behavioral biases. This is in line with the hypothesis that higher education and hereby forecasting knowledge result in well calibration. We interpret the second variable as a measure for the extent to which one uses complex analysis methods rather than simple technical rules or relying just on good luck. In all specifications the coefficient for fundamental analysis is significant. For the reference case an increased use of fundamental analysis of 10 % corresponds with a 4 % smaller chance to be overconfident and a 3.4 % higher likelihood to be underconfident. Nevertheless the effect of the variable is not robust for alternative measures, see Appendix 2. In general, the level of information might not be useful. In an experimental setting Huber et al. (2009) show that only the best informed traders outperform other investors. We find that the level of significance of the marginal effect increases ceteris paribus for more intense use of fundamental analysis, which corresponds to this hypothesis.

We measure how much each forecaster agrees with the market in her forecasts and interpret this as herding behavior. We can observe that the more (less) the forecaster aligns her forecast on the market the less (more) overconfident (underconfident) she is. In the baseline case of an average forecaster an increase in the market alignment of 10 % makes it more likely to be underconfident by 5 % and lowers the likelihood for overconfidence by 7.4 %. We propose two explanations for this evidence: (1) common risk-aversion for both measures, and (2) biased information processing. (1) A herding forecaster tends to rate herself quite conservatively because of the same reason why she is herding, due to high risk-aversion. Theoretical studies show that due to reputation effects lower risk taking and more herding go hand-in-hand, e.g. Diamond (1991) and Hirshleifer and Thakor (1992). Empirical evidence for this relationship is vast, for example Graham (1999) and Hong et al. (2000). Overconfident traders on the other side, underestimate the riskiness of projects, which was modeled by De Long et al. (1991) whose methodology was extended by Hirshleifer and Luo (2001). So forecasters who herd are unlikely to give a high self-rating On the whole forecasters who herd are less (more) overconfident (underconfident). (2) Another approach to explain this result might be an information procession bias which comes from the excessive trading literature (Barber and Odean, 2001), where the causality runs the other way round. Overconfident investors overweight the accuracy of their own information relatively to the precision of the information of other investors. Thus, overconfident investors do not pay as much as attention to the beliefs of others believes than to their own valuations (Odean, 1998). This intensifies differences in opinion (Harris and Raviv, 1993).

The literature often finds evidence that women are less affected of the better-than-average bias than men are (Barber and Odean, 2001). To cut the evidence short: women seem to be equally confident as men when they were truly correct, but nevertheless men match the stereotype when the tasks are typical masculine, e.g. Lundeberg et al. (1994), Beyer and Bowden (1998), for a review Barber and Odean (2001). Even though financial markets could be said to be masculine domains (for example remember that our sample encompasses 8 % women) some studies find also for the financial market domain no difference in overconfidence for gender: in an experimental financial market Deaves et al. (2003) cannot report any significant effect of gender – neither on trading volume nor on overconfidence. Oberlechner and Osler (2008) survey US traders but also do not report gender differences. For our sample we are not able to confirm these findings, neither for the group of overconfident nor the group of underconfident. But the large fraction men does not allow relying too much on these results. One exception is the small proportion of women that are included in our sample, which is about 8 %. But this is not unique for comparable data sets of financial market professionals, since women are still underrepresented in the financial sector business.

In order to control for effects of job positions held by the forecasters we include dummies for advisor and researcher and fund manager. Where researcher tend to be more overconfident, but not robustly, we find quite robust evidence for lower overconfidence of fund managers. For the average forecaster being a fund manager increases the likelihood to be underconfident by 32 % and reduces the likelihood to be overconfident by 32 %. This is a plausible finding when we assume that for our observed groups fund managers are the financial market participants with most clear and the most direct feedback. Since their salary is usually linked to their performance they are supposed to know their performance quite correctly.

5 Conclusion

This study examines overconfidence and underconfidence among professional forecasters. We contribute to the literature in that we combine "hard" performance information with self-rated performance and complement this with a comprehensive set of demographic and job characteristics. There are not many studies with a comparable depth of information and with respect to foreign exchange forecasters this study may be unique.

We reproduce standard findings on overconfidence and others to demonstrate the usefulness of our sample. Then, we find that there is no relation between the self-rating and effective performance which is surprising for professionals. Consequently, overconfidence is driven by performance in that bad forecasters are on average rather overconfident and the reverse. Interestingly, there are also underconfident professionals who have been largely neglected in earlier research.

In explaining these measures of over- and underrating one's own performance, we find a positive influence from market alignment and experience in the financial sector. In addition we show that fund managers are less overconfident than their fellow forecasters. We can provide evidence for the theoretical assumption that self-attribution leads to larger overconfidence.

References

Ackerman, P.L., Beier, M.E. and K.R. Bowen (2002), What We Really Know About Our Abilities and Our Knowledge, Personality and Individual Differences, 33, 587-605.

Barber, B. and T. Odean (2000), Trading is Hazardous to your Wealth: Common Stock Investment Performance of Individual Investors, Journal of Finance, 55, 773-806.

Barber, B. and T. Odean (2001), Boys will be Boys: Gender, Overconfidence, and Common Stock Investment, Quarterly Journal of Economics, 116, 261-292.

Bem, D.J. (1965), An Experimental Analysis of Self-Persuasion, Journal of Experimental Social Psychology, 1, 199-218.

Ben-David, I., Graham, J.R. and C.R. Harvey (2007), Managerial Overconfidence and Corporate Policies, NBER Working Paper 13711.

Beyer, S. and E. Bowden (1998), Gender Differences in Self-Perceptions: Convergent Evidence from Three Measures of Accuracy and Bias, Personality and Social Psychology Bulletin, 23, 157-172.

Biais, B., D. Hilton, K. Mazurier, and S. Pouget (2005), Judgemental Overconfidence, Self-Monitoring, and Trading Performance in an Experimental Financial Market, Review of Economic Studies, 72 , 287–312.

Brehmer, B. (1980), In One Word: Not from Experience, Acta Psychologica, 45, 223-241.

Chevalier, J. and G. Ellison (1999), Are Some Mutual Fund Managers Better than Others?, Journal of Finance, 54, 875-899.

Chuang, W.-I. and B.-S. Lee (2006), An Empirical Evaluation of the Overconfidence Hypothesis, Journal of Banking and Finance, 30, 2489-2515.

Daniel, K., Hirshleifer, D. and A. Subrahmanyam (1998), Investor Psychology and Security Market under- and overreactions, Journal of Finance, 53, 1839-1885.

Deaves, R., Lueders, E. and G.Y. Luo (2003), An Experimental Test of the Impact of Overconfidence and Gender on Trading Activity, discussion paper.

De Long, J.B., Shleifer, A., Summers, L.H. and R.J. Waldmann (1991), The Survival of Noise Traders in Financial Markets, Journal of Business, 64, 1-19.

Diamond, D. (1991), Reputation Acquisition in Debt Markets, Journal of Political Economy, 97, 828–862.

Diebold, F.X., Hickman, A., Inoue, A. and T. Schuermann (1998), Scale Models, Risk, 11, 104-107.

Edwards, F.R. and M. Caglayan (2001), Hedge Fund Performance and Manager Skill, Journal of Futures Markets, 21, 1003-1028.

Feng, L. and M.S. Seasholes (2005), Do Investor Sophistication and Trading Experience Eliminate Behavioral Biases in Financial Markets?, Review of Finance, 9, 305-351.

Fischhoff, B., Slovic, P. and S. Lichtenstein (1977), Knowing with Certainty: The Appropriateness of Extreme Confidence, Journal of Experimental Psychology, 3, 552-564.

Gervais, S. and T. Odean (2001), Learning to be Overconfident, Review of Financial Studies, 14, 1-27.

Glaser, M., T. Langer, and M. Weber (2005), Overconfidence of Professionals and Lay Men: Individual Differences Within and Between Tasks?, Working Paper, University of Mannheim.

Glaser, M. and M. Weber (2007a), Overconfidence and Trading Volume, Geneva Risk Insurance Review, 32, 1-36.

Glaser, M. and M. Weber (2007b), Why Inexperienced Investors Do Not Learn: They Do Not Know Their Past Portfolio Performance, Finance Research Letters, 4, 203-216.

Graham, J.R. (1999), Herding Among Investment Newsletters: Theory and Evidence, Journal of Finance, 54, 237-268.

Granger, C.W.J. and M.H. Pesaran (2000), Economic and Statistical Measures of Forecast Accuracy, Journal of Forecasting, 19, 537-560.

Greene, W.H. (2008), Econometric Analysis, 6th rev. ed., New Jersey: Prentice Hall.

Greenwald, A. (1980), The Totalitarian Ego: Fabrication and Revision of Personal History, American Psychologist, 35, 603-618.

Harris, M. and A. Raviv (1993), Differences in Opinion make a Horse Race, Review of Financial Studies, 6, 473-506.

Hirshleifer, D. and G.Y. Luo (2001), On the Survival of Overconfident Traders in a Competitive Securities Market, Journal of Financial Markets, 4, 73-84.

Hirshleifer, D. and A.V. Thakor (1992), Managerial Conservatism, Project Choice, and Debt, Review of Financial Studies, 5, 437-470.

Hong, H., Kubik, J.D. and A. Solomon (2000), Security Analysts' Career Concerns and Herding of Earnings Forecasts, Rand Journal of Economics, 31, 121-144.

Huber, J., M. Kirchler and M. Sutter (2008), Is More Information Always Better? Experimental Financial Markets with Cumulative Information, Journal of Economic Behavior and Organization, 65, 86-104.

Kirchler, E. and B. Maciejovsky (2002), Simultaneous Over- and Underconfidence: Evidence from Experimental Asset Markets, Journal of Risk and Uncertainty, 25, 65-85.

Klayman, J., J.B. Soll, C. Gonzáles-Vallejo, and S. Barlas (1999), Overconfidence: It Depends on How, What, and Whom You Ask, Organizational Behavior and Human Decision Processes, 79, 216-247.

Larrick, R.P., Burson, K.A. and J.B. Soll (2007), Social Comparison and Confidence: When Thinking You're Better than average Predicts Overconfidence (and when it does not), Organizational Behavior and Human Decision Processes, 102, 76-94.

Langer, E.J. (1975), The Illusion of Control, Journal of Personality and Social Psychology, 32, 311-328.

Langer, E.J. and J. Roth (1975), Heads I Win, Tails It's Chance: The Illusion of Control as a Function of the Sequence of Outcomes in a Purely Chance Task, Journal of Personality and Social Psychology, 32, 951-955.

Long, (1997), Regression Models for Categorical and Limited Dependent Variables, Sage.

Lundeberg, M., Fox, P. and J. LeCount (1994), Highly Confident but Wrong: Gender Differences and Similarities in Confidence Judgements, Journal of Educational Psychology, 86, 114-121.

MacDonald, R., Menkhoff. L. and R.R. Rebitzky (2008), Heterogeneity in Exchange Rate Expectations: Evidence on the Chartist-Fundamentalist Approach, mimeo.

Menkhoff, L., Rebitzky, R.R. and M. Schröder (2008), Do Dollar Forecasters Believe Too much in PPP?, Applied Economics, 40, 261-270.

Menkhoff, L., U. Schmidt and T. Brozynski (2006), The Impact of Experience on Risk Taking, Overconfidence and Herding of Fund Managers: Complementary Survey Evidence, European Economic Review, 50, 1753-1766.

Miller, D. T. and M. Ross (1975), Self-Serving Bias in Attribution of Causality: Factor Fiction?, Psychological Bulletin, 82, 213-225.

Moore, D.A. (2007), Not So Above Average After All: When People Believe They Are Worse Than Average and Its Implications for Theories of Bias in Social Comparison, Organizational Behavior and Human Decision Processes, 102, 42-58.

Moore, D.A. and D.M. Cain (2007), Overconfidence and Underconfidence: When and Why People Underestimate (and Overestimate) the Competition, Organizational Behavior and Human Decision Processes, 103, 197-213.

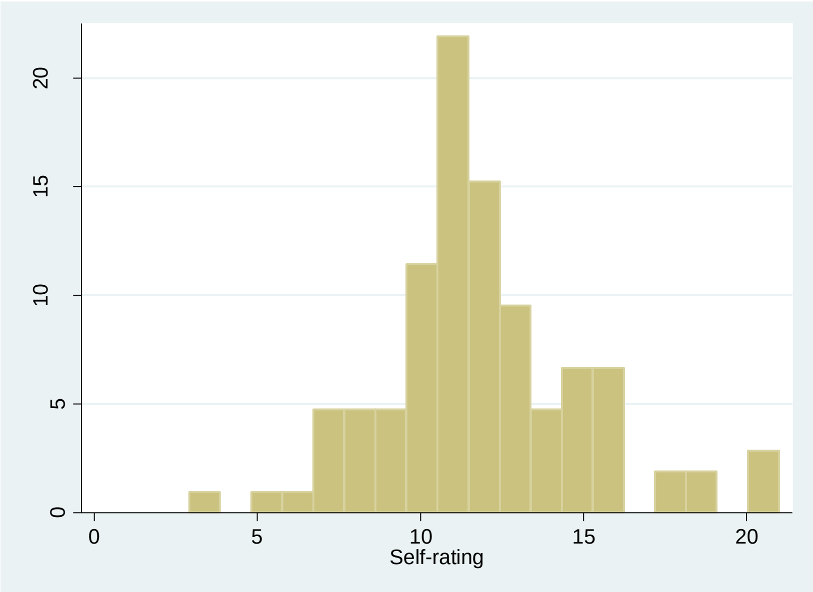
Oberlechner, T. and C. Osler (2008), Overconfidence in Currency Markets, Working Paper.

Odean, T. (1998), Volume, Volatility, and Profit When All Traders Are above Average, Journal of Finance, 53, 1887-1934.

Régner, I., Hilton, D., Cabantous, L. and S. Vautier (2004), Overconfidence, Miscalibration and Positive Illusions, mimeo.

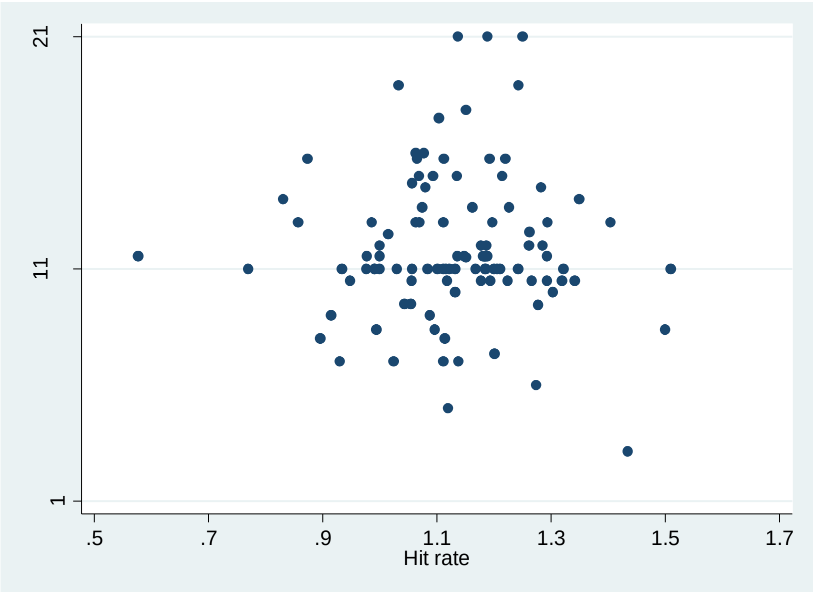
Svenson, 0. (1981), Are We All Less Risky and More Skillful Than Our Fellow Drivers?, Acta Psychologica, 47, 143-148.

Figure 1: Distribution of self-rating



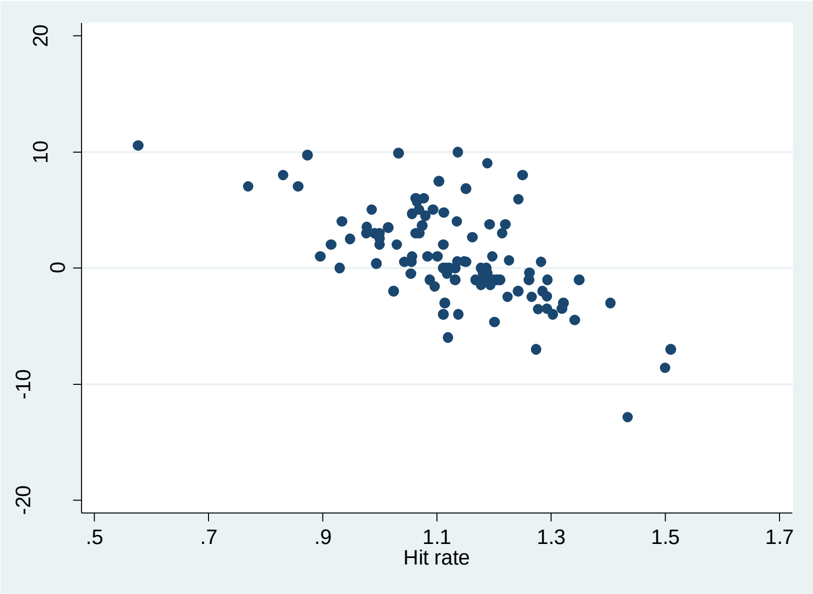
Note: We asked the financial experts in two surveys (April 2007 and October 2008) the following question: „How do you evaluate your USD/EUR-forecast compared to the average forecasting hit rate of all participants of the ZEW financial market survey?". The scale ranges from 1 to 21.

Figure 2: Scatter plot of self-rating and hit rate



Note: The self-rating is a survey item ranging from 1 to 21 and indicates whether someone believes to be above (21) or below (1) the average hit rate. Hit rate is extracted from the survey forecasts over 36 monthly observations at minimum. To incorporate that the experts can choose between three options a hit rate is coded in three categories, large deviation (0), small deviation (1) and no deviation (2) of forecast from the true process. Large deviations are predictions which indicate the opposite direction of the actual movement, whereas small deviations are expectations which are neither a correct forecast nor a large deviation.

Figure 3: Scatter plot of over-/underconfidence and hit rate



Note: Overconfidence is measured as the difference between the self-rating and hit rate. To have both on the equal scale the ladder is recoded to the same scale as the self-rating, which again ranges from 1 to 21. The lower (upper) boundary for a hit rate which corresponds to a considerably lower (better) self-rating than average is 0.548 (1.727). This corresponds to an interval of 4 standard deviations around the median. According to Chebyshev's inequality we capture hereby 94% of the distribution at minimum. Assuming normal distribution we collect even more than 99.99% of the theoretical probability mass. Our actual sample ranges from 0.577 to 1.510. So, all observations are included in the transformed hit rate measure.

Table 1: Descriptive statistics of financial experts' characteristics

	Obs.	Mean	Std. Dev.	Min.	Max.
Overconfidence characteristics					
Better-than-average[1]	105	11.91	3.15	3	21
Miscalibration[2]	74	0.14	0.07	0.04	0.33
Illusion-of-control[3]	73	13.47	3.75	4	20
Demographic characteristics					
Male[†]	105	0.92	0.27	0	1
Academic education[†]	105	0.76	0.43	0	1
Age[4]	105	44.56	8.11	29	65
Working experience in fin. market[4]	105	17.69	8.78	3	43
Job characteristics					
Personnel responsibilities[†]	105	0.50	0.50	0	1
Operative responsibilities[†]	105	0.78	0.42	0	1
Advisor[†]	105	0.18	0.39	0	1
Fund manager[†]	105	0.30	0.46	0	1
Researcher[†]	105	0.23	0.42	0	1
Forecasting characteristics					
Hit rate[6]	105	1.13	0.14	0.58	1.51
Recent success[7]	105	1.18	0.43	0	2
Positive trend in hit rate[†,8]	105	0.15	0.36	0	1
Negative trend in hit rate[†,8]	105	0.07	0.25	0	1
Fundamental analysis[9]	105	55.05	22.41	0	100
Forecast market alignment[10]	105	51.42	18.54	1.02	92.08

Note: Dummy variables are denoted by "[†]".

[1] We asked the financial experts in two surveys (04/2007 and 10/2008) the following question: „How do you evaluate your USD/EUR-forecast compared to the average forecasting hit rate of all participants of the ZEW financial market survey?". The scale ranged from 1 to 21.

[2] The question for miscalibration was an item in the survey of October 2008. Respondents gave a 90-% confidence interval for the 6-month future USD/EUR exchange rate.

[3] Illusion-of-control was surveyed in October 2008. The information was extracted from the following question: "Most of the published business news does not surprise me at all." Respondents answered on a scale ranging from 1 to 20.

[4] Age and working experience are given in years.

[6] Hit rate measures the individual average hit rate over the observation period, where the individual hit rate at a time point codes the forecasting performance in no deviation (2), small deviation (1), and large deviation (0).

[7] Recent success is the individual average hit rate of 1-month forecasts over the last 6 months. The minimum observation number is 3. The boundaries of the no change category are adjusted via the square root formula (Diebold et al., 1998). So, since the 6-month boundary (b_{t+6}) is 3% we compute for the 1-month boundary $b_{t+1} = b_{t+6}(30/180)^{0.5} = 1.22$ %. The variable is robust for variations in the underlying number of months (5, 7, and 9).

[8] Positive (negative) trend in hit rate is a dummy variable for a significant (on the 10 % level) positive (negative) trend in the forecasting performance over the last three years. We used spearman's rank correlation. Kendall's tau gives nearly identical results since we use just the directional information.

[9] Fundamental analysis is a self-evaluation. The value measures the degree how much fundamental analysis (in %) is used for creating the exchange rate expectation.

[10] Forecast market alignment measures how often (in %) a forecaster expects the exchange rate change in the same direction as the market does. We refer to market by using the mode of all participating forecasters and set a minimum participation of 30 to have a robust estimate of the market opinion. The individual minimum participation rate for which we calculated market alignment is 12.

Table 2: Correlations between better-than-average, miscalibration and illusion-of-control

VARIABLES	Better-than-average bias[1]	Miscalibration bias[2]	Hindsight Bias[3]
Better-than-average [1]	1.0000		
	71		
	.		
Miscalibration [2]	-0.0444	1.0000	
	71	71	
	(0.7131)	.	
Illusion-of-control [3]	**0.3111*****	-0.1398	1.0000
	71	71	71
	(0.0080)	(0.2450)	.

Note: The table provides Spearman's rank correlation coefficients and the corresponding p-values (in parenthesis) and the number of observations. The level of significance are denoted by *** p<0.01, ** p<0.05, * p<0.1.

[1] We asked the financial experts in two surveys (April 2007 and October 2008) the following question: „How do you evaluate your USD/EUR-forecast compared to the average forecasting hit rate of all participants of the ZEW financial market survey?". The scale ranged from 1 to 21.

[2] The questions for miscalibration was posed in the survey of October 2008. Participants were asked for a 90-% confidence interval for the 6-month future USD/EUR exchange rate.

[3] Data for illusion-of-control was collected in October 2008 and was extracted from the following question: "Most of the published business news does not surprise me at all." Respondents answered on a scale ranging from 1 to 20.

Table 3: Ordered logit estimation results for over-/underconfidence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Male	0.436	0.444	0.460	0.482					
	(0.668)	(0.660)	(0.647)	(0.632)					
Academic education	-0.249	-0.245	-0.256						
	(0.682)	(0.683)	(0.667)						
Age	0.0389	0.0390	0.0379	0.0305	0.0319				
	(0.494)	(0.496)	(0.514)	(0.582)	(0.563)				
Working exp. in fin. sector	**-0.0880***	**-0.0877***	-0.0864	-0.0769	-0.0760	**-0.0502***	**-0.0496****	**-0.0543****	**-0.0578****
	(0.100)	**(0.0957)**	(0.106)	(0.103)	(0.112)	**(0.051)**	**(0.049)**	**(0.035)**	**(0.019)**
Personnel responsibility	-0.476	-0.476	-0.485	-0.516	-0.535	-0.523	-0.516		
	(0.271)	(0.271)	(0.264)	(0.237)	(0.213)	(0.221)	(0.229)		
Operative responsibility	0.0546								
	(0.940)								
Advisor	0.127	0.121							
	(0.852)	(0.856)							
Fund Manager	**-1.240****	**-1.233****	**-1.277****	**-1.286****	**-1.277****	**-1.247****	**-1.259****	**-1.168****	**-1.384*****
	(0.034)	**(0.032)**	**(0.016)**	**(0.016)**	**(0.017)**	**(0.018)**	**(0.017)**	**(0.017)**	**(0.002)**
Researcher	1.023	0.996	0.954	0.886	0.828	0.872	0.786	0.855	
	(0.167)	(0.115)	(0.117)	(0.145)	(0.155)	(0.132)	(0.174)	(0.146)	
Positive trend in hit rate	**1.318***	**1.314***	**1.321***	**1.285***	**1.221***	**1.264***	**1.326***	**1.310***	**1.352***
	(0.097)	**(0.095)**	**(0.092)**	**(0.088)**	**(0.089)**	**(0.075)**	**(0.059)**	**(0.069)**	**(0.055)**
Negative trend in hit rate	-0.763	-0.774	-0.775	-0.801	-0.756	-0.732			
	(0.419)	(0.394)	(0.385)	(0.372)	(0.397)	(0.412)			
Fundamental analysis	**-0.0197***	**-0.0196***	**-0.0196***	**-0.0189***	**-0.0185***	**-0.0187***	**-0.0185***	**-0.0187***	**-0.0165***
	(0.052)	**(0.055)**	**(0.056)**	**(0.065)**	**(0.064)**	**(0.064)**	**(0.061)**	**(0.061)**	**(0.096)**
Forecast market alignment	**-0.0394*****	**-0.0394*****	**-0.0391*****	**-0.0386*****	**-0.0388*****	**-0.0376*****	**-0.0380*****	**-0.0349****	**-0.0305****
	(0.003)	**(0.004)**	**(0.003)**	**(0.004)**	**(0.004)**	**(0.005)**	**(0.006)**	**(0.011)**	**(0.018)**
Cut 1	**-3.611***	**-3.633***	**-3.685***	**-3.598***	**-3.987***	**-4.856*****	**-4.816*****	**-4.439*****	**-4.399*****
	(0.098)	**(0.087)**	**(0.083)**	**(0.086)**	**(0.051)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
Cut 2	-2.969	-2.990	-3.042	-2.957	**-3.347***	**-4.217*****	**-4.181*****	**-3.813*****	**-3.785*****
	(0.168)	(0.154)	(0.147)	(0.153)	**(0.098)**	**(0.001)**	**(0.001)**	**(0.001)**	**(0.001)**
Pseudo-R2	0.134	0.134	0.134	0.133	0.132	0.130	0.126	0.119	0.108
Observations	105	105	105	105	105	105	105	105	105

Note: The level of significance are denoted by *** p<0.01, ** p<0.05, * p<0.1, robust p-values in parentheses. For description of the variables see annotations of Table 1. The present table shows estimations of the full model, which includes all available control variables. We deselect insignificant variables via backward selection procedure. Forward selection yields the same final result.

Table 4: Effects of a marginal/discrete change in the ordered logit regression model

	UNC	NN	OVC
Working experience in financial sector	0.0119	0.0023	-0.0141
Fund Manager	0.3296†	-0.0085†	-0.3212†
Positive trend in hit rate	-0.1935†	-0.0723†	0.2658†
Fundamental analysis	0.0034	0.0007	-0.0040
Forecast market alignment	0.0063	0.0012	-0.0075
P(y | x)	28.85	13.96	57.19
P(y)	36.19	12.38	51.43

Note: UNC (OVC) corresponds to forecasters who are underconfident (overconfident) and NN to forecasters who are neither nor. The marginal effects are calculated for the reference case of a forecaster who shows average working experience (18 years) and use of fundamental analysis (55 %) and market alignment (51 %) and who is neither a fund manager nor has a positive trend in the hit rate. The effect for the dummy variables is a discrete change from 0 to 1, denoted by "†".

Appendix 1: Ordered probit regression results for over-/ underconfidence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Male	0.280	0.281	0.294	0.308					
	(0.617)	(0.614)	(0.597)	(0.581)					
Academic education	-0.110	-0.109	-0.117						
	(0.759)	(0.758)	(0.740)						
Age	0.0237	0.0238	0.0228	0.0195	0.0192				
	(0.462)	(0.467)	(0.490)	(0.537)	(0.550)				
Working exp. in fin: sector	-0.0502	-0.0501	-0.0490	-0.0448	-0.0435	**-0.0277***	**-0.0276***	**-0.0311****	**-0.0332****
	(0.103)	(0.100)	(0.115)	(0.107)	(0.124)	**(0.063)**	**(0.061)**	**(0.039)**	**(0.024)**
Personnel responsibility	-0.266	-0.266	-0.275	-0.291	-0.296	-0.293	-0.281		
	(0.295)	(0.294)	(0.279)	(0.253)	(0.243)	(0.245)	(0.268)		
Operative responsibility	0.0105								
	(0.979)								
Advisor	0.0918	0.0909							
	(0.815)	(0.816)							
Fund Manager	**-0.725****	**-0.724****	**-0.758****	**-0.767****	**-0.755****	**-0.741****	**-0.735****	**-0.704****	**-0.841*****
	(0.032)	**(0.030)**	**(0.012)**	**(0.012)**	**(0.014)**	**(0.015)**	**(0.017)**	**(0.017)**	**(0.002)**
Researcher	0.657	**0.651***	**0.619***	**0.584***	0.552	**0.576***	0.523	0.544	
	(0.121)	**(0.082)**	**(0.083)**	**(0.097)**	-0.106	**(0.091)**	(0.116)	(0.104)	
Positive trend in hit rate	**0.740***	**0.740***	**0.739***	**0.732***	**0.699***	**0.719***	**0.760***	**0.755***	**0.773***
	(0.093)	**(0.091)**	**(0.094)**	**(0.090)**	**(0.091)**	**(0.081)**	**(0.063)**	**(0.066)**	**(0.060)**
Negative trend in hit rate	-0.507	-0.508	-0.505	-0.519	-0.489	-0.473			
	(0.329)	(0.317)	(0.314)	(0.303)	(0.331)	(0.346)			
Fundamental analysis	**-0.0121****	**-0.0120****	**-0.0120****	**-0.0117****	**-0.0114****	**-0.0114***	**-0.0113****	**-0.0114****	**-0.00995***
	(0.039)	**(0.041)**	**(0.043)**	**(0.049)**	**(0.049)**	**(0.051)**	**(0.049)**	**(0.049)**	**(0.083)**
Forecast market alignment	**-0.0237*****	**-0.0237*****	**-0.0235*****	**-0.0232*****	**-0.0232*****	**-0.0225*****	**-0.0224*****	**-0.0207*****	**-0.0181****
	(0.002)	**(0.002)**	**(0.002)**	**(0.002)**	**(0.003)**	**(0.003)**	**(0.004)**	**(0.007)**	**(0.015)**
Cut 1	-2.062	**-2.066***	**-2.111***	**-2.068***	**-2.341****	**-2.858*****	**-2.807*****	**-2.633*****	**-2.611*****
	(0.101)	**(0.092)**	**(0.082)**	**(0.086)**	**(0.042)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
Cut 2	-1.679	-1.683	-1.728	-1.685	**-1.959***	**-2.477*****	**-2.428*****	**-2.257*****	**-2.244*****
	(0.178)	(0.166)	(0.152)	(0.159)	**(0.087)**	**(0.000)**	**(0.001)**	**(0.001)**	**(0.001)**
Pseudo-R2	0.133	0.133	0.132	0.132	0.130	0.128	0.123	0.118	0.105
Observations	105	105	105	105	105	105	105	105	105

Note: The level of significance are denoted by *** p<0.01, ** p<0.05, * p<0.1, robust p-values in parentheses. For description of the variables see annotations of Table 1. The present table shows estimations of the full model, which includes all available control variables. We deselect insignificant variables via backward selection procedure. Forward selection yields the same final result.

Appendix 2: Ordered logit/ ordered probit estimation for different regressors

	(1) OLogit	(2) OLogit	(3) OProbit	(4) OProbit
Male	0.194		0.200	
	(0.847)		(0.726)	
Academic education	-0.153		-0.120	
	(0.787)		(0.724)	
Age	0.0464		0.0280	
	(0.391)		(0.371)	
Working exp. in fin. sector	**-0.0903***	**-0.0546****	**-0.0520***	**-0.0305****
	(0.073)	**(0.029)**	**(0.078)**	**(0.034)**
Personnel responsibility	-0.438		-0.224	
	(0.343)		(0.399)	
Operative responsibility	-0.158		-0.0948	
	(0.829)		(0.813)	
Advisor	0.0999		0.0127	
	(0.890)		(0.975)	
Fund Manager	**-1.156****	**-1.097*****	**-0.686****	**-0.662*****
	(0.033)	**(0.008)**	**(0.033)**	**(0.009)**
Researcher	0.588		0.419	
	(0.468)		(0.348)	
Recent success	0.814	**0.874***	0.475	**0.508***
	(0.127)	**(0.052)**	(0.122)	**(0.064)**
Fundamental analysis	**-0.0152***		**-0.00992***	
	(0.080)		**(0.062)**	
Forecast market alignment	**-0.0406*****	**-0.0314****	**-0.0236*****	**-0.0180****
	(0.004)	**(0.013)**	**(0.002)**	**(0.014)**
Cut 1	-2.667	**-2.554****	-1.472	**-1.448****
	(0.185)	**(0.013)**	(0.223)	**(0.017)**
Cut 2	-2.035	**-1.950****	-1.095	**-1.088***
	(0.306)	**(0.049)**	(0.361)	**(0.065)**
Pseudo-R2	0.121	0.0935	0.119	0.0894
Observations	105	105	105	105

Note: The level of significance are denoted by *** p<0.01, ** p<0.05, * p<0.1, robust p-values in parentheses. For description of the variables see annotations of Table 1. The present table shows estimations of the full model (1) and (3), as well as the restricted models of a backward selection procedure (with threshold .1), see equation (2) and (4). Forward selection yields the same final result.