The dynamics of overconfidence: Evidence from stock market forecasters

This version: October 7, 2005

Richard Deaves*

Michael G. DeGroote School of Business, McMaster University
&
Center for European Economic Research *(ZEW)*
deavesr@mcmaster.ca

Erik Lüders

Pinehurst Capital
&
Center for European Economic Research *(ZEW)*

Michael Schröder

Center for European Economic Research *(ZEW)*

The dynamics of overconfidence: Evidence from stock market forecasters

This version: October 7, 2005

ABSTRACT

As a group, market forecasters are egregiously overconfident. In conformity to the dynamic model of overconfidence of Gervais and Odean (2001), successful forecasters become more overconfident. What's more, more experienced forecasters have "learned to be overconfident," and hence are more susceptible to this behavioral flaw than their less experienced peers. It is not just individuals who are affected. Markets also become more overconfident when market returns have been high.

1. Introduction and literature review

Man is an overconfident breed. Overconfidence arises when knowledge perception exceeds its reality. While we should not be overly surprised if individuals operating outside their natural domains fall prey to this flaw, naturally it is to be hoped that, for an individual making his living in a certain field of expertise, the actual knowledge level is high, and the perception of this level accurate. Unfortunately the reality is that experts can be quite overconfident.[1]

Theory suggests that overconfidence is likely to be more prevalent if feedback is infrequent and ambiguous (Fischhoff, Slovic and Lichtenstein (1977)). Based on this insight, one might expect overconfidence to be less common for frequent forecasters of unambiguous events. That is to say, professional forecasters, whether they are meteorologists or sports handicappers, might be expected to develop a better sense of their knowledge and limitations through experience, and be less overconfident than neophytes. Another example of an expert prognosticator is an individual who earns her living in part by making stock market forecasts. One might expect that fairly frequent unambiguous feedback, coupled with the fact that poor performers are routinely weeded out in financial markets, should lead those surviving being good at their craft and developing a sense of their knowledge in line with its reality.

A careful examination of this issue requires some clarity on how overconfidence can best be measured, and why it is usually viewed as a negative attribute. Unfortunately universal agreement does not exist on what overconfidence is. The *better-than-average effect* (e.g., Svenson (1981)) stems from the observation that more than 50% of the population seems to think that they are more skilled at driving than average; their children are smarter than average; they can choose stocks better than the average trader; and so on. Those subject to *illusion of control* (Langer (1975)) believe that they can somehow exercise more control over

3

events such as avoiding accidents, winning lotteries, buying hot stocks or avoiding poorly performing stocks than can realistically be possible.

These two approaches however do not lend themselves to easy quantification of overconfidence though, so a third approach to its measurement which is more convenient in this regard, namely miscalibration, is preferable. In a calibration test (Lichtenstein *et al* (1982), Keren (1991) and McClelland and Bolger (1994)), when individuals are asked to construct *x%* confidence intervals for currently (or soon to be) known magnitudes such as the height of Mount Everest (or the level of the *Dow* in a month), a percentage of individuals usually markedly below *x%* produce intervals that bracket the true answer. Another reason for preferring the miscalibration approach is that the new theoretical models (e.g., Kyle and Wang (1997) and Odean (1998)), in which overconfidence plays center stage, utilize a definition of overconfidence that is most akin to miscalibration. Still, it must be acknowledged that there is no ideal overconfidence metric, a point which picks up resonance when it is observed that people's overconfidence scores using different metrics often have low correlations (Glaser and Weber (2005)).

What damage is wrought by overconfidence? A good example is Odean (1998), who has formulated a behavioral model that suggests that overconfident investors, believing that they possess greater precision on security valuations than is merited, trade too much and thereby lower their wealth (and expected utility). The intuition is clear: the more certain you are of your view, the less credence you will accord those of others, and the more likely you will be to transact at a price perceived favorable to your view. Barber and Odean (2000) demonstrate that investors who trade more suffer the consequences by receiving lower risk-adjusted returns. And Deaves, Lüders and Luo (2005) have produced experimental evidence supporting the *direct* correlation between overconfidence and trading activity.[2] In a corporate finance context, it has been demonstrated (e.g., Malmendier and Tate (2005)) that CEOs sometimes make sub-optimal investment decisions if they are prone to overconfidence. What about economy-wide forecasts, such as the level of the stock market? Needless to say, since both

investors and corporate decision-makers use forecasts for portfolio formation and planning purposes, a poor sense of the market's true volatility can inflict egregious harm.

The dynamics of overconfidence is clearly an important issue. It is logical to think that if we recall our successes and failures equally clearly, over time we should obtain an accurate view. Experience should make us wise. On the other hand, the prevalence and persistence of overconfidence suggest that forces able to eliminate it are weak. The reality is that we prefer to forget what did not go our way: this is called *cognitive dissonance.* Working in tandem, *self-attribution bias* leads us to remember our successes with great clarity, if not embellishment (Fischoff (1982), Langer and Roth (1975), Miller and Ross (1975) and Taylor and Brown (1988)). And *confirmatory bias*, the tendency to search out evidence consistent with one's prior beliefs and to ignore conflicting data, may also contribute (Forsythe, Nelson, Neumann and Wright (1992)). A strict efficient markets view of the world would seem to argue that those fooling themselves in this way will be driven from the marketplace, but some have called this into question (Hirshleifer and Luo (2001)).

The dynamic nature of overconfidence is stressed in the models of Daniel, Hirshleifer and Subrahmanyam (1998, 2001) and Gervais and Odean (2001). In the latter paper, which operates in a multi-period setting, past successes, through the mechanism of self-attribution bias, exacerbate overconfidence (while past failures tend to be ignored). The inference is that those who have had the good fortune of being successful in their fields might for a time be more overconfident than those who have just entered. Eventually however experience should reveal to people their true knowledge level.

While overconfidence can afflict *individuals*, it can also be endemic in *markets*. The most likely trigger is high past market returns. Statman, Thorley and Vorkink (2004) show that lagged market returns are correlated with increases in trading activity. This constitutes indirect evidence of an increase in

overconfidence through self-attribution bias, since behavioral models typically associate overconfidence with excessive trading. Glaser and Weber (2004) investigate whether individual investors adjust volume more in response to market returns or own-portfolio returns. They find that those who keep track of their own portfolios are more influenced by the latter, while those less aware are more likely to be influenced by market returns. While this line of research is suggestive, it relies on using volume as a proxy for overconfidence. Clearly it would be helpful to investigate whether lagged market movements and aggregate overconfidence are directly related.

The purpose of this paper is to examine both the statics and dynamics of overconfidence of stock market forecasters. The survey instrument employed is the *ZEW Finanzmarkttest*. The latter is a monthly survey of financial market practitioners in Germany. Respondents are asked for 90% confidence intervals for the level of the *DAX* six months ahead. While an analysis of forecasting accuracy is far from unusual, our purpose here is quite different. The availability of not just point estimates but also confidence intervals allows for a careful exploration of overconfidence in both its static and dynamic manifestations. Graham and Harvey (2003) make use of a somewhat similar dataset of *CFO* forecasts and confidence intervals for the U.S. stock market. Their focus is not on overconfidence, however, as they investigate what can be learned about *ex ante* equity premiums and the relationship between risk and return.

We first investigate whether the respondent group as a whole is overconfident. Next we explore whether people learn from past successes and failures, and from work experience in financial markets. Finally we consider whether the market as a whole becomes overconfident in response to high past returns. To preview, we first conclude that market forecasters are egregiously overconfident. Second, success, measured by correct prediction leads to increased overconfidence. Third, market experience, which is symptomatic of past success, is associated with higher levels of overconfidence. And, fourth, the market does learn to be overconfident through past collective success.

Section 2 describes the *ZEW Finanzmarkttest*. The next section specifies the hypotheses to be tested. Section 4 details the key empirical findings. And the final section concludes.

2. The *ZEW Finanzmarkttest* survey

The *ZEW Finanzmarkttest* is a monthly survey of about 350 financial market practitioners in Germany. Most of these individuals work for a commercial bank, investment bank, insurance company or investment department of a large German company. Each month, beginning in 1991, participants have been asked to predict a series of key macroeconomic and financial market variables for the key industrialized economies as of six months in the future. For example, participants are asked to predict the inflation rate, long-term and short-term interest rates, economic activity, and stock market levels for these countries. Until recently questions only asked for direction: that is, rise/fall/unchanged. This questionnaire forms the basis for the well-known *ZEW Finanzmarkt Indicator,* an indicator which, together with the *IFO* index, is one of the most important and most closely followed economic indicators in Germany.[3]

Of course in Germany one of the key magnitudes to forecast is the level of the *DAX*. This index is analogous to a German *Dow*. Specifically, the *DAX* is an index composed of the 30 largest and most important German companies traded on the *German Stock Exchange* in Frankfurt. Starting in February 2003, *ZEW* survey respondents were asked to provide, in additional to a directional forecast, a quantitative one for the *DAX*. Specifically, point estimates for the *DAX* six months in the future, as well as lower and upper bounds forming 90% confidence intervals, are now solicited. We restrict our analysis to *these* forecasts and intervals.

To gather more information on the participants, in September 2003 a demographic survey of participants was taken. Most of the regular *ZEW Finanzmarkttest* participants responded to this survey (about 250 out of 350).

Thus, as of September 2003, we know the gender, age, educational level, educational field and professional experience of respondents – as well as whether or not the forecasts are based on self-conducted research or whether they rely on others' research. Advantageously, the quantitative forecasts examined here are never more than 18 months away from when the demographic survey was conducted. This mitigates substantially the problem of respondent changes.[4]

Table 1 provides a demographic snapshot. It is apparent that this group is predominately male and well-educated,[5] tends to have had an educational specialization in economics or business, and does its own research. On average the group has good experience in financial markets. Since this will be a key variable in our analysis, we show the distribution of market experience in *Figure 1*. Evidently there is a very broad range and the distribution is skewed to the right.

3. Hypotheses

The first hypothesis tested is a static one:

Hypothesis 1: Market forecasters as a group are overconfident in their predictions.

Since overconfidence seems to be so prevalent in society, the finding that it exists in this realm should not surprise us. Still, we might expect only moderate levels of overconfidence since the respondents are able to fairly frequently assess how accurate their forecasts are. Moreover those performing poorly are unlikely to have long careers and many in our sample do have extensive experience in financial markets.

The next three hypotheses are dynamic in nature. Hypotheses 2 and 3 focus on the cross section, with the former conditioning on forecasting success:

Hypothesis 2: Success, measured by recent forecast accuracy, increases overconfidence, while failure (less strongly) mitigates overconfidence.

There is evidence that certain personal characteristics are correlated with overconfidence. For example, males tend to be more overconfident than females (Lundeberg, Fox and Punccohar (1994), Beyer (1990) and Barber and Odean (2001)).[6] And Bhandari and Deaves (2005) find that highly-educated well-paid males are especially susceptible. Past success is also argued (by Gervais and Odean (2001)) to increase overconfidence. The thrust here is that past success triggers self-attribution bias that exacerbates overconfidence, while failures are in part ignored.

Hypothesis 3: Success, measured by experience in financial markets, increases overconfidence.

Since survival in financial markets is a signal of persistent success, we would also expect a correlation between market experience and overconfidence. Experience however should also engender wisdom, that is, a better sense of one's limitations. It is an empirical question as to which dominates. Nevertheless, the model of Gervais and Odean (2001) suggests the latter is the weaker force, especially at low to moderate levels of experience.

Hypothesis 4 is in terms of market-level overconfidence:

Hypothesis 4: Success at the level of the market, proxied by high lagged market returns, increases the average level of overconfidence.

As mentioned earlier, Statman, Thorley and Vorkink (2004) show that lagged market returns are correlated with increases in aggregate trading activity. Volume however is only a proxy for overconfidence. It is preferable to test this hypothesis *directly* by seeking to establish a linkage between past returns and measured overconfidence.

4. Empirical results

i. Hypothesis 1

The respondents must produce 90% confidence intervals for the *DAX* six months in the future. The obvious approach is to ascertain how often these intervals contain the eventually correct level of the *DAX*. Unfortunately the direct route is not ideal. The forecasts, while made monthly, are for six months ahead. If we use overlapping observations there will be a substantial amount of non-independence. One surprise will have an impact on not just one forecast horizon but instead on six. A purely cross-sectional approach does not help. The problem is that a truly unexpected market event surprises not just one forecaster but everyone. An alternative is to wait six months to see if individuals are right while eschewing overlapping observations. Unfortunately, for a barely minimal number of independent observations we would have to possess 10 years of data (as opposed to just over two years).

Two indirect tests of the hypothesis that forecasters are overconfident were carried out. First, by examining a time series of the history of the *DAX,* we can ascertain how wide a 90% confidence interval *should* be. We then compare the width of respondent confidence intervals to this ideal width. If the typical width is too narrow relative to this ideal width, then we infer that forecasters are overconfident.

More specifically, we calculate, at each forecast date, historical (based on half-year returns) 10-year rolling window standard deviations. An approximate 90% confidence interval width consistent with *DAX* history should be *2 * 1.645 * Current DAX * SD of DAX returns*.[7] So, at each forecast point, a comparison can be made between the average confidence interval width of respondents and what is suggested by the history of the *DAX*. *Figure 2* documents that the average interval of respondents is substantially narrower than what it should be as suggested by *DAX* history.[8] Oddly, the divergence between the curves increases

10

over the first year or so. The interval based on history increases because the *DAX* increases dramatically over this period (e.g., 60% for the year ending March 2004), necessitating a commensurate interval expansion. At the same time, the average interval among respondents declines over the first year. This in fact offers preliminary support for Hypothesis 4 which calls for lagged returns and average overconfidence to be positively correlated. We will return to this hypothesis later in the paper.

Another approach for ascertaining overconfidence begins with inferring one-month forecast confidence intervals from six-month intervals based on a simple reasonable assumption, namely that forecasters believe that the growth rate in the *DAX* will be constant over the next six months. More specifically, suppose a given forecaster provides L6, F6 and U6 for the 6-month interval lower bound, forecast point estimate and interval upper bound respectively. From F6 the one-month forecast point estimate (F1) is calculated assuming constant growth in the *DAX:*

$$(1) \qquad F1 = (\frac{F6}{Current\ DAX})^{1/6} * Current\ DAX$$

On the assumption of *i.i.d. DAX* 1-month returns, the standard deviation of 1-month returns is $1/\sqrt{6}$ times the 6-month standard deviation. Respondents have chosen their intervals to reflect what they perceive to be the correct number of standard deviations on each side of their point estimate, so all we have to do is multiply U6-F6 and F6-L6 by $1/\sqrt{6}$ to arrive at U1-F1 and F1-L1. And since we have F1 it is straightforward to calculate L1 and U1. Finally, using L1 and U1, we note whether or not the eventual realization of the *DAX* falls between the bounds. Assuming our procedure is sound and respondents are well-calibrated, this should occur 90% of the time.

Before providing the results, it should be noted that there was one complication that had to be addressed. Respondents typically have about two weeks to make

11

their 6-month ahead forecasts. Since they are not told to do otherwise, logically respondents would be making their forecasts for *exactly* six months in the future. Using our 1-month interpolation procedure, our values are then appropriate for exactly one month in the future. But one month in the future occurs on different dates for different people. So what we in fact have to do is use respondent-specific "current" *DAX* values and compare intervals to respondent-specific "eventual" *DAX* values.

What do the data tell us? While respondents should be right 90% of the time, *Figure 3* shows that in actual fact this occurs far less than that.[9] Note that at best about 80% of respondents are "right," and quite often less than 50% are right. The inference once again is that our respondents are egregiously overconfident.

It is also instructive to view overconfidence cross-sectionally. In *Figure 4,* we show a frequency distribution of the percentage of times individuals are right over the full sample. The mode suggests that many are right between 40%-70% of the time even though they are asked to be right 90% of the time. Note the spikes on the endpoints. These two groups are dominated by people who answered the survey infrequently. For example, the 22 people who had 100% accuracy on average responded to the survey only twice (vs. an overall sample average of 16.7 times), suggesting that many of the apparently underconfident were in fact lucky.

To sum up, using several (albeit imperfect) approaches, the survey data indicate that the market forecasters in our sample are egregiously (though perhaps not universally) overconfident.

ii. Hypothesis 2

The dynamics of individual overconfidence is approached in two ways. To begin, we explore whether a successful forecast (in the sense that the confidence interval contains the eventual *DAX* value) in the most recent period leads to a narrowing

of a respondent's confidence interval in the following period. Since overconfidence is the norm, a narrowing of intervals suggests *increased* overconfidence. Using the technique described above, one-month intervals are inferred from 6-month intervals. Of course it is a leap in faith to suggest that respondents will "check" in a meaningful sense how well their implicit one-month forecasts performed and adjust behavior in response. For this reason the associated noise might make it difficult to detect any innate dynamics.

More specifically, regressions of the following form are run:

$$(2) \qquad \Delta CIRW_{i,t} = a + b * RIGHT_{i,t-1}$$

where $\Delta CIRW_{i,t}$ is the change in the width of the confidence interval (relative to the cross-sectional average width) at time t for respondent $i;$ and $RIGHT_{i,t}$ is an indicator variable equaling one when the respondent's interval contains the eventual *DAX* value and zero otherwise. Note that a equals the change based on wrong answers; and $a+b$ equals the change for right answers.

Panel A of *Table 2* provides full-sample (panel) regression results. Additionally, the full sample is partitioned into roughly equal subsamples. For the full sample as well as for the first subperiod, there is extremely strong evidence that forecasters adjust the width of their confidence intervals based on how well their intervals performed in the previous month.[10] Typically when respondents get it wrong, they expand their intervals; and when they get it right they contract their intervals. Even for the second subperiod the evidence is fairly solid, in that the slope coefficient is still significant in the neighborhood of the 5% level.

A few other salient points can be made. First, these results are quite robust. Cross-sectional forecast-by-forecast regressions are run and the coefficient estimates are shown in *Figure 5*. The intercept is positive 23 out of 25 times, and the slope is negative all but once. In all but four cases, the absolute value of b is

greater than that of a, which means that intervals contract after success (a+b<0) but expand after failure (a>0).

Second, it is not possible to conclude that success has a greater impact than failure: interval contraction after success is about the same in magnitude as interval expansion after failure (since *2a* is approximately equal to *|b|* for the full sample*).*[11] Strictly speaking, this is inconsistent with self-attribution bias which argues that overconfidence increases with success but does not decline as much with failure.

Third, despite the highly significant slope coefficients, the R-squared's are quite low in these regressions, which of course means that relative confidence interval widths are changing for a variety of other (unknown) reasons. The fact that we had to impute 1-month intervals from 6-month obviously exacerbates the problem. Nevertheless, quite reliably, success, or its absence, is an important force.

Fourth, the Bayesian reader may argue that the direction of confidence interval changes is a logical response to new information and thus is not necessarily indicative of increases in overconfidence. A signal (a "right" answer) that one is a good forecaster would seem to imply that confidence intervals *should* contract. There is a problem, however, with this view, as revealed by *Table 3,* which provides prior (over all previous months) and subsequent-month success rates for those who were right or wrong in the current month.[12] Trivially, in a month where someone gets it right, she could be said to be "underconfident" since one was asked to be right 90%, not all the time (one of one). So it is possible to argue that the "right" group intervals are being contracted to move things towards proper 90% calibration. Focus on the third and sixth columns though. In the months prior to the successful forecast (column 3) nowhere near 90% of these people got it right; and their lack of success holds up in the subsequent month (column 6).[13] This suggests of course that if they contracted their intervals they

were unwise to do so – unless (oddly) the impact accorded the last observation in the Bayesian updating is excessive.

iii. Hypothesis 3

Next we look at the dynamics of individual overconfidence in a different way. One problem with the previous dynamic approach is that it relies on respondents "checking" their forecasts. The variant of the second hypothesis focusing on financial market experience does not however require this. As mentioned earlier, a one-time demographic survey was conducted in September 2003 of the *ZEW Finanzmarkttest* survey participants, with about 70% of survey respondents participating in this "extra" task. Conveniently, one of the questions was years of experience working in financial markets. Recall that the model of Gervais and Odean (2001) implies that experience, at least up to a point, will exacerbate overconfidence.

We next explore this, but at the same time we can broaden our focus. First, it is possible to partition overconfidence into certainty (that is, confidence as opposed to overconfidence) and knowledge.[14] Only if the former *exceeds* the latter is one overconfident. Second, given our demographic data, we are also able to investigate whether other demographic factors have an impact on overconfidence (as well as on certainty and knowledge individually). The other demographic factors of interest are age, gender, educational level, area of study and whether or not one generated the forecast oneself.

In *Table 4* we report the results of cross-sectional regressions on experience and age.[15] To more specifically define our dependent variables, certainty is expressed in terms of an individual's confidence interval width relative to the cross-sectional average at each point in time. We then take a time series average to obtain an individual-specific average relative confidence interval width. The greater is this value, the *less* confident (or certain) is the individual. Thus the following metric for confidence is appropriate:

$$(3) \quad Certainty_i = \frac{Max(Avg\ rel\ conf\ range\ width\ ) - Avg\ rel\ conf\ range\ width_i}{Max(Avg\ rel\ conf\ range\ width\ ) - Min(Avg\ rel\ conf\ range\ width\ )}$$

Knowledge is based on average absolute forecast errors scaled by *DAX* realizations. It is defined as follows:

$$(4) \quad Knowledge_i = \frac{Max(Avg\ abs\ scaled\ fore\ error\ ) - Avg\ abs\ scaled\ fore\ error_i}{Max(Avg\ abs\ scaled\ fore\ error\ ) - Min(Avg\ abs\ scaled\ fore\ error\ )}$$

Note that, for both formulas above, the metrics are bounded between zero and unity, and the *max* and *min* are over the cross section.

Finally, overconfidence is one minus the percentage of the time that an individual's interval contains the eventual *DAX* realization. A value of .1 would reflect neither underconfidence nor overconfidence, but values above this threshold reflect overconfidence.

Focusing first on the simple regressions where years of experience is the sole independent variable, it is apparent that, consistent with the model of Gervais and Odean (2001), experience is detrimental in that it increases overconfidence. To interpret, three additional years of experience increase overconfidence by 1.2% (that is, one is likely to have 1.2% fewer intervals right). Is it deteriorating knowledge or a tendency to increase certainty that leads to additional overconfidence? Note that self-attribution bias would be more consistent with the latter, but our results are otherwise. The only significant coefficient is on knowledge, with the interpretation here being that five more years of experience drops someone down 1% in the knowledge ranking. How can this be explained? Perhaps job experience is a double-edged sword. On the one hand, additional years on the job should lead to greater wisdom. On the other hand, human nature is such that effort declines with maturity. Our evidence suggests that the second effect outweighs the first, causing more experienced people to display lower levels of knowledge (accuracy).

One might question whether it is job experience or age that drives the results. Not surprisingly the two are highly correlated (.85). To resolve this issue, *Table 4* also provides regression results where age – orthogonalized so that only its component independent of job experience can have an impact – is included as a second regressor. Clearly age adds nothing.

Finally (not shown in *Table 4*) we also perform multivariate regressions of certainty, knowledge and overconfidence on all the demographic variables of interest (not just the two shown in the table). In all cases, the additional regressors are safely insignificant, while the impact of financial market experience is easily retained.[16]

iv. Hypothesis 4

Finally, we examine the dynamics of market-wide overconfidence. Specifically, we investigate whether high recent market returns make the market more overconfident. First, begin with *Panel B* of *Table 2* which investigates individual-level changes in overconfidence as a function of individual forecasting success or failure (as in *Panel A*) *and* lagged market returns. The p-values for the latter are never close to indicating significance.

A decomposition of the variance of the full panel of imputed one-month confidence intervals indicates that most of the variability is cross-sectional. Only 4% of the overall variance is accounted for by variability in the time series average. This suggests that we are more likely to find an impact on aggregate overconfidence if our dependent variable is the change in the *average* confidence interval width.

Figure 5 indicates that this is indeed the case. Begin with the first regression which investigates whether average overconfidence increases with the percentage of respondents who are right. This is analogous to *Panel A* of *Table 2* but with averaging over all individuals. There is no discernible impact. The second

regression includes as an additional regressor the most recent one-month *DAX* return. Despite the very sparse number of time series observations, one can comfortably conclude that high past market returns push the entire market towards greater overconfidence. Interestingly, further regressions show that the impact of history is short-lived. It should be stressed that, while this result is consistent with the work of Statman, Thorley and Vorkink (2004) and Glaser and Weber (2004), it constitutes *direct* evidence in contrast to their more *indirect* evidence.

5. Conclusion

The purpose of this paper was to explore the statics and dynamics of overconfidence. Overconfidence is proxied by miscalibration. We find that the *Finanzmarkttest* respondent group as a whole is egregiously overconfident. More interestingly, they do not seem learn from past successes and failures, or from experience. This flaw is also present at the level of the market as we see that high past returns induce increased aggregate overconfidence. All of this suggests that overconfidence is not only a pervasive phenomenon, but it is also one that is exceedingly difficult to eliminate.

While respondents alter their intervals dependably based on past success or failure, the low level of explanatory power suggests that there are other forces inducing changes in behavior. Indeed, an exploration into the nature of these additional factors would be a worthwhile avenue for future research.

Table 1: Demographic characteristics of survey respondents

	Mean	*Median*	*Min*	*Max*	*SD*	*N*
Age	42	42	24	65	8.7	236
Gender (male=1)	.95	1	0	1	.21	244
Years in financial markets	15	13	1	45	9.6	241
Own research? (yes=1)	.82	1	0	1	.39	240
Education (university equivalent=1)	.58	1	0	1	.49	244
Field (economics/business =1)	.95	1	0	1	.22	199

Table 2: Panel regressions of changes in relative confidence interval width on previous forecasting success and previous DAX returns

$$\Delta CIRW_{i,t} = a + b * RIGHT_{i,t-1} + c * DAX_{t-1}$$

	obs	*a*	*b*	*c*	*Adjusted R^2*
			Panel A		
Full sample	4966	0.046 (0.001)	-0.090 (0.000)		0.004
2003:03- 2004:03	2737	0.048 (0.000)	-0.103 (0.000)		0.016
2004:04- 2005:03	2229	0.043 (0.000)	-0.075 (0.052)		0.001
			Panel B		
Full sample	4966	0.047 (0.001)	-0.090 (0.000)	-0.049 (0.772)	0.004
2003:03- 2004:03	2737	0.049 (0.000)	-0.103 (0.000)	-0.041 (0.712)	0.016
2004:04- 2005:03	2229	0.043 (0.133)	-0.075 (0.053)	0.011 (0.985)	0.001

Note: P-values are in brackets beside co-efficient estimates.

Table 3: Prior/subsequent success rates of right/ wrong

Month	Prior month forecast accuracy			Subsequent month forecast accuracy		
	Wrong group (current month)	Right group (current month)	Diff.	Wrong group (current month)	Right group (current month)	Diff.
2003:03	22.2%	42.7%	20.4%	29.1%	39.2%	10.2%
2003:04	34.1%	30.8%	-3.3%	62.2%	83.1%	20.9%
2003:05	21.8%	37.8%	15.9%	25.8%	50.3%	24.6%
2003:06	36.1%	48.8%	12.7%	49.2%	82.1%	32.9%
2003:07	38.9%	42.4%	3.4%	46.8%	66.4%	19.6%
2003:08	37.5%	48.9%	11.5%	49.4%	47.7%	-1.7%
2003:09	45.0%	49.9%	5.0%	25.9%	24.7%	-1.2%
2003:10	44.2%	53.3%	9.1%	62.5%	82.0%	19.5%
2003:11	37.6%	48.0%	10.4%	35.1%	61.0%	25.9%
2003:12	43.1%	49.9%	6.8%	78.1%	71.8%	-6.3%
2004:01	44.2%	50.0%	5.8%	24.5%	49.4%	24.8%
2004:02	50.6%	52.3%	1.7%	56.6%	50.0%	-6.6%
2004:03	46.5%	50.7%	4.2%	32.7%	32.8%	0.0%
2004:04	49.1%	52.5%	3.4%	67.6%	61.8%	-5.7%
2004:05	40.2%	52.1%	11.9%	44.4%	42.9%	-1.6%
2004:06	46.7%	54.0%	7.3%	2.8%	20.5%	17.7%
2004:07	48.7%	65.2%	16.5%	50.0%	57.1%	7.1%
2004:08	41.0%	51.6%	10.6%	55.1%	72.1%	17.0%
2004:09	40.2%	50.6%	10.4%	65.8%	89.1%	23.2%
2004:10	41.1%	50.3%	9.3%	32.5%	54.2%	21.7%
2004:11	44.4%	54.9%	10.5%	42.9%	91.4%	48.5%
2004:12	37.0%	54.7%	17.7%	43.3%	94.2%	50.9%
2005:01	37.1%	55.0%	17.9%	56.3%	81.4%	25.2%
2005:02	43.7%	54.0%	10.3%	61.4%	81.5%	20.0%

Note: Percentages indicate how often the relevant group at the relevant time got a forecast right. Whether or not one belongs to the "right" or "wrong" group is a function of the forecast only in the current month.

Table 4: Certainty, knowledge and overconfidence vs. experience and age

$$DV_i = a + b * Years \ in \ markets_i + c * Age_i$$

Dependent variable (DV)	a (p-value)	b (p-value)	c (p-value)	R^2 Adjusted R^2	Observations
Certainty (or certainty)	3.031 (.000)	.0008 (.280)	---	.004 (.001)	239
	2.989 (.000)	-.0004 (.827)	.0015 (.410)	.007 (-.002)	231
Knowledge	-0.105 (.000)	-.0019 (.034)	---	.050 (.046)	240
	-0.096 (.000)	-.0017 (.095)	-.0002 (.806)	.052 (.044)	232
Overconfidence	0.431 (.000)	.0041 (.002)	---	.040 (.036)	239
	0.495 (.000)	.0058 (.027)	-.0022 (.451)	.042 (.034)	231

Note: For the knowledge regressions above, it was not possible to reject heteroscedasticity at conventional significance levels (based on the White test), so reported standard errors are consistent in the presence of heteroscedasticity.

Table 5: Time series regressions of changes in average confidence interval width on past DAX returns and previous average forecasting success

$$\Delta ACIW_t = a + b * PRIGHT_{t-1} + c * DAX_{t-1} + d * DAX_{t-2} + e * DAX_{t-3}$$

Sample	a	b	c	d	e	Adj. R^2
2003:03-2005:03	-0.005 (0. 273)	0.002 (0.784)				-0.040
2003:03-2005:03	-0.004 (0.350)	0.002 (0.751)	-0.057 (0.028)			0.132
2003:03-2005:03	-0.003 (0.450)	0.001 (0.880)	-0.057 (0.031)	0.007 (0.798)		0.093
2003:03-2005:03	-0.003 (0.465)	0.001 (0.898)	-0.058 (0.034)	0.006 (0.822)	0.006 (0.777)	0.052

Note: P-values are in brackets beside co-efficient estimates.

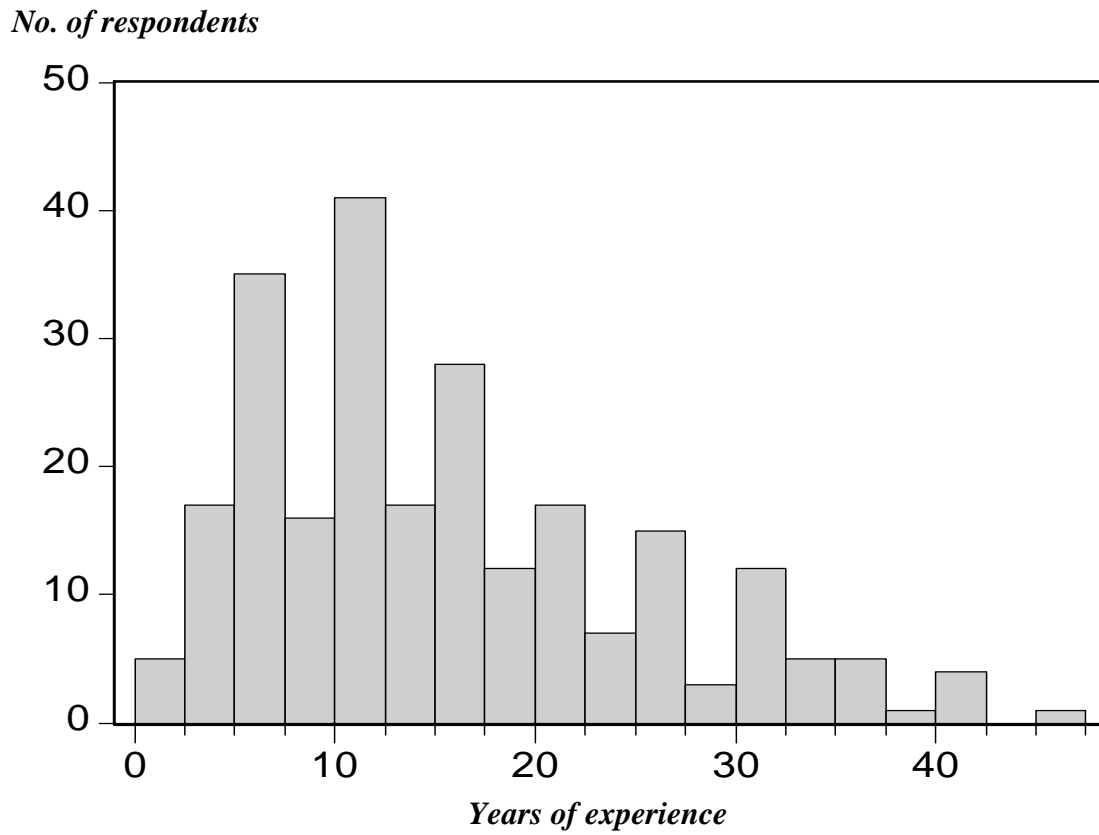**Figure 1: Frequency distribution for years of experience in financial markets**

No. of respondents

Years of experience

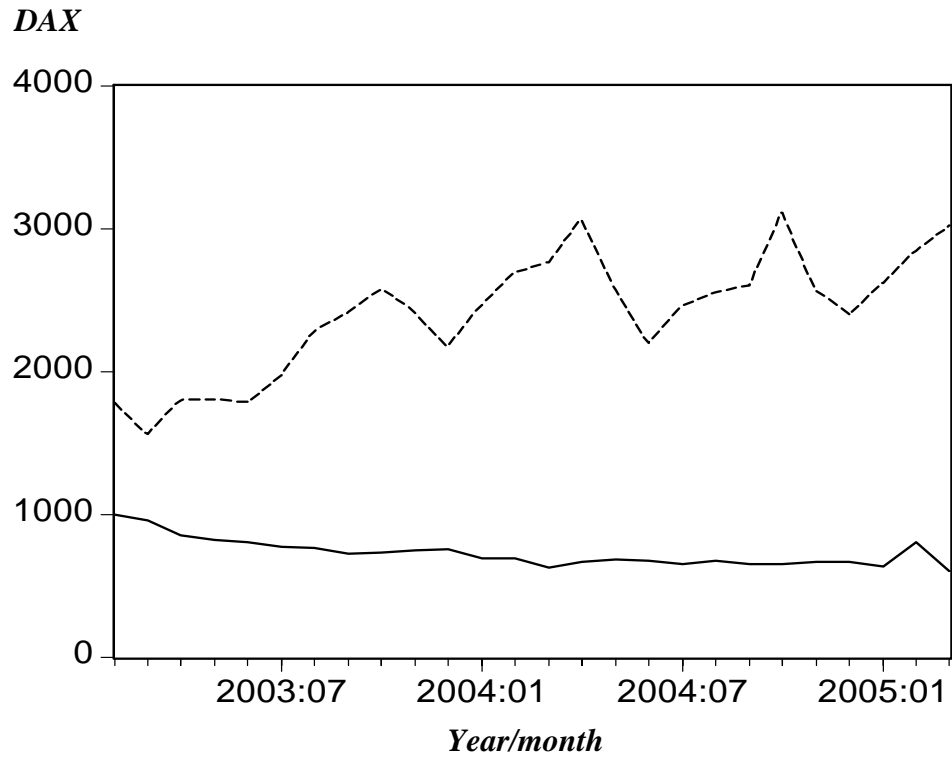**Figure 2: Average respondent confidence interval (solid line) vs. interval suggested by DAX history (broken line)**

DAX

Year/month

Figure 3: Percentage of time that respondents' intervals contain eventual value of DAX (time series)
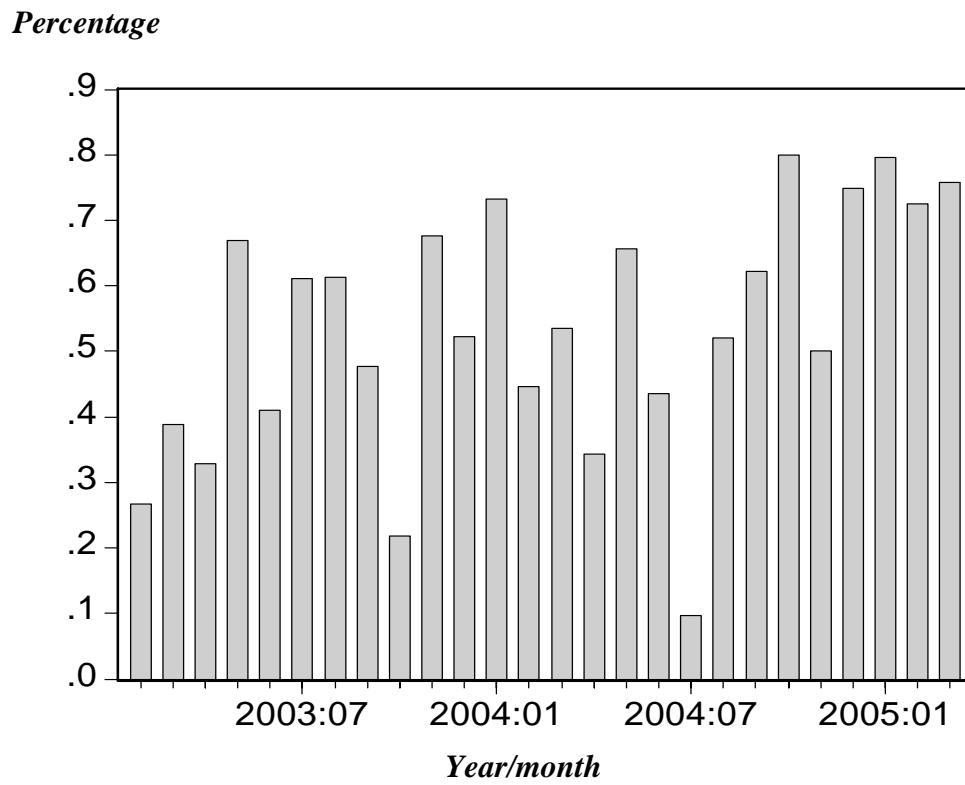
Percentage

Year/month

Figure 4: Percentage of time that respondents' intervals contain eventual value of DAX (cross-sectional)

Percentage of respondents
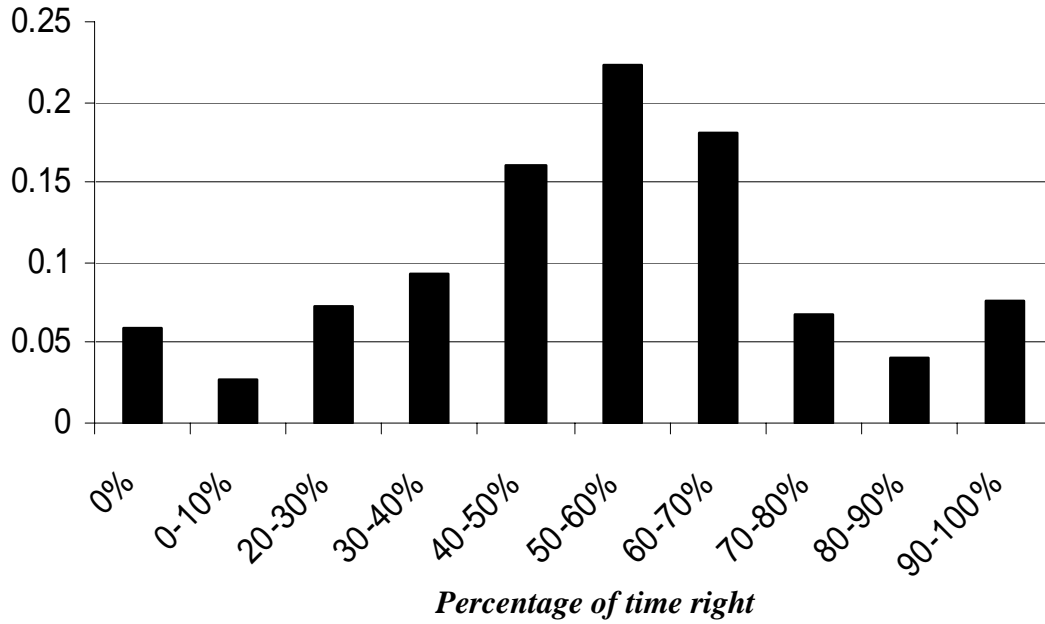

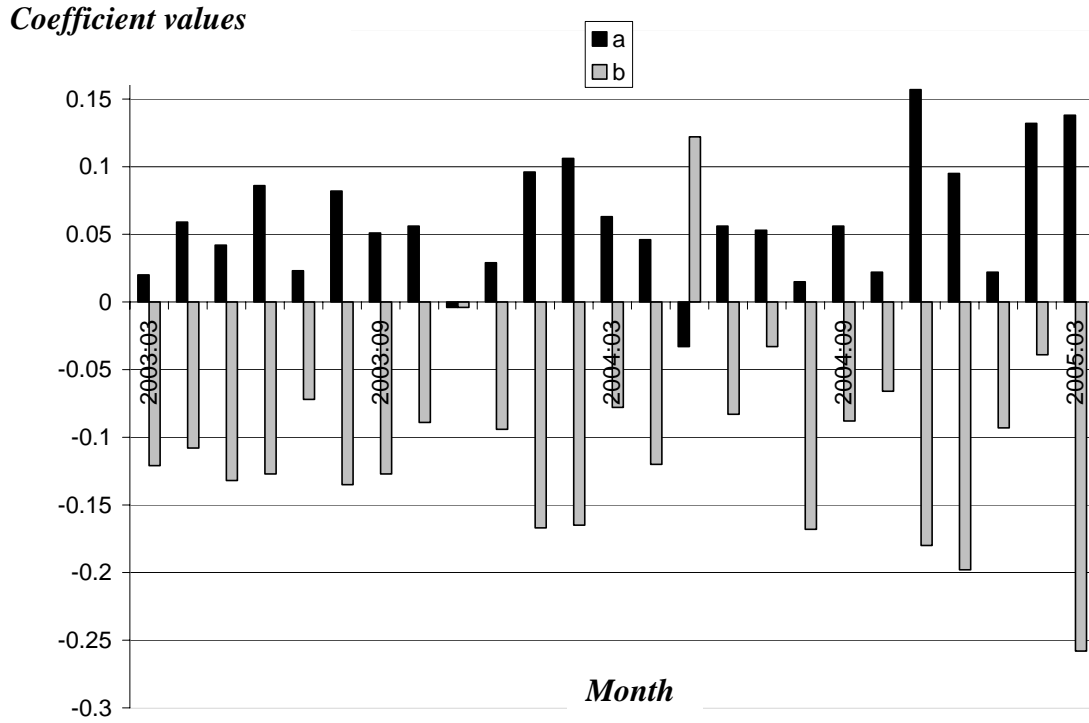*Percentage of time right*

Figure 5: Month-by-month intercept and slope coefficients for regression of changes in relative confidence interval width on previous forecasting success

Coefficient values

Month

REFERENCES

Barber, B., and T. Odean, 1999. "The courage of misguided convictions." *Financial Analysts Journal* Special Issue on Behavioral Finance: 41-55.

Barber, B., and T. Odean, 2000. "Trading Is hazardous to your wealth: The 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: 261-92.

Beyer, S., 1990. "Gender differences in the accuracy of self-evaluations and performance," *Journal of Personality and Social Psychology* 59: 960-70.

Bhandari, G., and R. Deaves, 2005. "The demographics of overconfidence," Forthcoming in the *Journal of Behavioral Finance.*

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.

Daniel, K. D., D. Hirshleifer and A. Subrahmanyam, 1998. "Investor psychology and investor under- and overreactions," *Journal of Finance* 53: 1839-86.

Daniel, K. D., D. Hirshleifer and A. Subrahmanyam, 2001. "Mispricing, covariance risk, and the cross section of security returns," *Journal of Finance* 5: 921-65.

Deaves, R., E. Lüders and G. Y. Luo, 2005. "An experimental test of the impact of overconfidence and gender on trading activity," Working paper.

Fischhoff, B., 1982. "For those condemned to study the past: Heuristics and biases in hindsight," in Kahneman, D., A. Slovic and A. Tversky (editors), *Judgment under uncertainty: Heuristics and biases*, Cambridge University Press, Cambridge.

Fischhoff, B., P. Slovic and S. Lichtenstein, 1977. "Knowing with certainty: The appropriateness of extreme confidence," *Journal of Experimental Psychology* 3: 552-64.

Forsythe, R., F. Nelson, G. Neumann and J. Wright, 1992. "Anatomy of an experimental political stock market," *American Economic Review* 82:1142-61.

Gervais, S., and T. Odean, 2001. "Learning to be overconfident," *Review of Financial Studies* 14: 1-27.

Glaser, M., and M. Weber, 2003. "Overconfidence and trading volume," Working paper.

Glaser, M., and M. Weber, 2004, "Which past returns affect trading volume?" Working paper.

Glaser, M., T. Langer and M. Weber, 2005, "Overconfidence of professionals and laymen: Individual differences within and between tasks?" Working paper.

Graham, J.R., and C. R. Harvey, 2003. "Expectations of equity risk premia, volatility and asymmetry," Working paper.

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.

Keren, G., 1991. "Calibration and probability judgments: Conceptual and methodological issues," *Acta Psychologica* 77: 217-73.

Kyle, A. S., and F. A. Wang, 1997. "Speculation duopoly with agreement to disagree: Can overconfidence survive the market test?" *Journal of Finance* 52: 2073-90.

Langer, E. J., 1975. "The illusion of control," *Journal of Personality and Social Psychology* 32: 311-28.

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: 951-55.

Lichtenstein, S., B. Fischhoff and L.D. Phillips, 1982. "Calibration of probabilities: The state of the art to 1980," in Kahneman, D., A. Slovic and A. Tversky (editors), *Judgment under uncertainty: Heuristics and biases*, Cambridge University Press, Cambridge.

Lundeberg, M. A., P. W. Fox and J. Punccohar, 1994. "Highly confident but wrong: Gender differences and similarities in confidence judgments," *Journal of Educational Psychology* 86: 114-21.

Malmendier, U., and G. Tate, 2005. "CEO overconfidence and corporate investment," Forthcoming in *Journal of Finance.*

McClelland, A., and F. Bolger, 1994. "The calibration of subjective probabilities: Theories and models 1980-93," in Wright, G., and P. Ayton (editors), *Subjective Probability*, Wiley, Chichester.

Miller, D. T., and M. Ross, 1975. "Self-serving bias in attribution of causality: Fact or fiction?" *Psychological Bulletin* 82: 213-25.

Odean, T., 1998. "Volume, volatility, price and profit when all traders are above average," *Journal of Finance* 53: 1887-1934.

Statman, M, S. Thorley and K. Vorkink, 2004. "Investor overconfidence and trading volume," Working paper.

Svenson, O., 1981. "Are we all less risky and more skillful than our fellow drivers?" *Acta Psychologica* 47: 143-48.

Taylor, S. E., and J. D. Brown, 1988. "Illusion and well-being: A social psychological perspective on mental health," *Psychological Bulletin* 103: 193-210.

ENDNOTES

[1] See Barber and Odean (1999) for numerous references.

[2] Also working in an experimental setting, Biais, Hilton, Mazurier and Pouget (2005) are unable to conclude that overconfidence leads to an increase in trading intensity, though they do find that it serves to significantly reduce profits. Glaser and Weber (2003), combining a survey and naturally-occurring market data, document the correlation between the better-than-average effect variant of overconfidence and trading activity.

[3] The *IFO* index is based on a survey of about 7,000 companies in Germany on their business expectations. Every month the *IFO Institute* asks about 7,000 enterprises in manufacturing, construction, wholesaling and retailing to give their assessments of the current business situation and their expectations for the next six months. The enterprises can characterize their situation as "good," "satisfactory" or "poor," and their business expectations for the next six months as "more favorable," "unchanged" or "more unfavorable."

[4] While the questionnaire is sent to the same addresses every month, the person in the department filling it out and submitting the forecasts might change without *ZEW* being notified. This potentially adds noise to our results.

[5] There are 30 *Ph.D.s*.

[6] Deaves, Lüders and Luo (2004) do not corroborate this finding, but this likely due to the fact that their female subjects are business students rather than belonging to a diverse group.

[7] This is a conservative estimate. As return distributions such as the *DAX* exhibit leptokurtosis the correct lower and upper limits based on a bootstrap would actually be wider.

[8] It should not be too surprising to hear that the gap between the two curves is always highly significant (p-value=0.000).

[9] For each date it can easily be rejected (with a p-value of 0.000) that the percentage of correct respondents is 90% or more.

[10] This is not the same as having accurate point estimates. In fact when we regress confidence interval changes on accuracy in this sense the coefficient is insignificant.

[11] It is not possible to reject the hypothesis that *2a+b=0.*

[12] Current month success rates are definitionally 0% and 100% for the two groups.

[13] The last several months are an exception, as three out of six times the "right" group gets it right at a rate of 89% or better in the subsequent month.

[14] In Bhandari and Deaves (2005), due to how overconfidence was defined, it was possible to make the metric for overconfidence exactly equal to the difference between the metrics for certainty and knowledge. Unfortunately this is not possible here.

[15] For the two knowledge regressions, heteroscedasticity could not be rejected (based on White tests) at conventional levels, so reported standard errors are consistent in the presence of heteroscedasticity.

[16] It may be a little surprising that gender does not impact overconfidence. This is partly due to the small percentage of females in the sample (see *Table 1)* and the fact that females in the forecasting field are likely a more overconfident group than the general female population (see footnote 6).