Journal of Economic Behavior & Organization xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

The dynamics of overconfidence: Evidence from stock market forecasters

Richard Deaves[a,b], Erik Lüders[c,b], Michael Schröder[b,*]

[a] DeGroote School of Business, McMaster University, Hamilton, Ontario, Canada
[b] Center for European Economic Research (ZEW), Mannheim, Germany
[c] McKinsey & Company, Inc, Frankfurt/Main, Germany[1]

ARTICLE INFO

Article history:
Received 8 November 2007
Received in revised form 16 April 2010
Accepted 11 May 2010
Available online xxx

JEL classification:
D83
D84
G1

Keywords:
Overconfidence
Stock market
Behavioral finance
Learning
Panel data

ABSTRACT

As a group, market forecasters are overconfident in the sense that they are miscalibrated. While overconfidence is persistent, respondents do exhibit some degree of rational learning in that they widen confidence intervals after failure as much as they narrow them after success. Market experience exacerbates overconfidence, primarily through knowledge deterioration.

1. Introduction

There is abundant evidence that most people most of the time are overconfident in the sense that they overestimate the precision of their knowledge, a phenomenon known as miscalibration.[2] In one variant of a calibration test, 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), usually markedly below *x*% of their intervals bracket is the true answer. Suboptimal financial decision-making, ranging from excessive trading (Barber and Odean, 2000; Deaves et al.,

* Corresponding author at: Center for European Economic Research (*ZEW*), P.O. Box 103443, D-68034 Mannheim, Germany. Tel.: +49 621 1235140; fax: +49 621 1235223.
E-mail address: schroeder@zew.de (M. Schröder).

[1] Any views expressed represent those of the authors only and not necessarily those of McKinsey & Company, Inc.

[2] Universal agreement does not exist on what overconfidence is. While in the present paper, we use the term "overconfident" in the sense of miscalibrated, there are other perspectives or metrics. The better-than-average effect (Svenson, 1981) induces most of us to think that we are more skilled at tasks or knowledgeable about information than average. Those subject to illusion of control (Langer, 1975) believe that they can somehow exercise more control over external events than can realistically be possible. Related to the latter is excessive optimism (Armor and Taylor, 2002), which exists when people assign unreasonably high/low probabilities to good/bad events. Importantly, we observe that people's overconfidence "scores" using different metrics often have low correlations (Glaser et al., 2005). See (Lichtenstein et al., 1982) for a review of miscalibration. Gigerenzer (1991) argues that miscalibration is largely frame-induced.

0167-2681/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jebo.2010.05.001

2008) and underdiversification (Goetzmann and Kumar, 2008) on the part of investors; or excessive entry into markets (Camerer and Lovallo, 1999), allowing investment to be dictated by cashflows (Malmendier and Tate, 2005) and over-investment (Ben-David et al., 2007) on the part of entrepreneurs and managers, has been argued to flow from various forms of overconfidence.

It is not just neophytes who are overconfident: more surprisingly, so are experts in their fields.[3] This includes practitioners in business and finance.[4] The insight that overconfidence is less likely if feedback is frequent and unambiguous (Fischhoff et al., 1977) leads to the expectation that this behavioral flaw is less common for frequent forecasters of unambiguous events. One example is individuals who earn their living in part by making stock market forecasts. Given the profit motive that is front and center in financial markets, weak performance should make survival more difficult, encouraging a sense of self-awareness by survivors.[5] 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 harm.

If people recall their successes and failures equally clearly, over time they should obtain an accurate view. In this sense, experience should engender wisdom. On the other hand, the prevalence and persistence of overconfidence suggest that the forces able to eliminate it are weak. Most problematically, self-attribution bias leads us to remember our successes with great clarity, if not embellishment (Miller and Ross, 1975).[6] Along these lines, in the dynamic model of Gervais and Odean (2001), past successes exacerbate overconfidence, while past failures tend to be downplayed.[7] The result is that those who have had the good fortune of being successful in their fields might for a time be more overconfident than new entrants. Eventually, however, experience should reveal to people their true knowledge level.

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* which is a monthly survey of financial market practitioners in Germany. Respondents are asked for 90% confidence intervals for the level of the *DAX* 6 months ahead. While an analysis of forecasting accuracy is far from unusual, the 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 begin by investigating whether, consistent with previous evidence, the respondent group as a whole is overconfident. Next we explore whether survey respondents adjust their confidence intervals in response to past successes and failures. The relationship between experience and miscalibration is then considered. To preview, market forecasters are overconfident. While overconfidence persists, some learning does seem to occur as confidence intervals widen with failure and narrow with success in equal measure. Greater market experience is associated with higher levels of overconfidence. 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 6 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 Indicator of Economic Sentiment*, an indicator which, together with the *ifo Business Climate* index, is one of the most important and most closely followed economic indicators in Germany.[8]

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 a value-weighted index composed of the 30 largest and most important German companies traded on the *German Stock Exchange* in Frankfurt.[9] 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* 6 months in

[3] Two examples are clinical psychologists (Oskamp, 1965) and engineers (Kidd, 1970).

[4] Two examples of practitioners being overconfident are managers (Russo and Schoemaker, 1992) and investment bankers (Staël von Holstein, 1972).

[5] 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).

[6] Related to this is cognitive dissonance, which sometimes induces us to forget what is unpleasant or did not go our way (Festinger and Carlsmith, 1959), and confirmation bias, the tendency to search out evidence consistent with one's prior beliefs and to ignore conflicting data, may also contribute (Forsythe et al., 1992).

[7] Other theoretical models of overconfidence include Kyle and Wang (1997), Odean (1998) and Daniel et al. (1998, 2001).

[8] The *ifo Business Climate* 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 provide assessments of their current business situation and expectations for the next 6 months. These enterprises can characterize their situation as "good," "satisfactory" or "poor," and business expectations for the next 6 months as "more favorable," "unchanged" or "more unfavorable."

[9] More specifically, the *DAX* is a net-dividend-reinvested index, which means it includes after-tax dividends which are reinvested in the index.

Table 1
Demographic characteristics of survey respondents.

	Mean	Median	Min	Max	SD	*N*
Age	42	42	24	65	8.7	236
Gender (male = 1)	0.95	1	0	1	0.21	244
Years in financial markets	15	13	1	45	9.6	241
Own research? (yes = 1)	0.82	1	0	1	0.39	240
Education (university equivalent = 1)	0.58	1	0	1	0.49	244
Field (economics/business = 1)	0.95	1	0	1	0.22	199

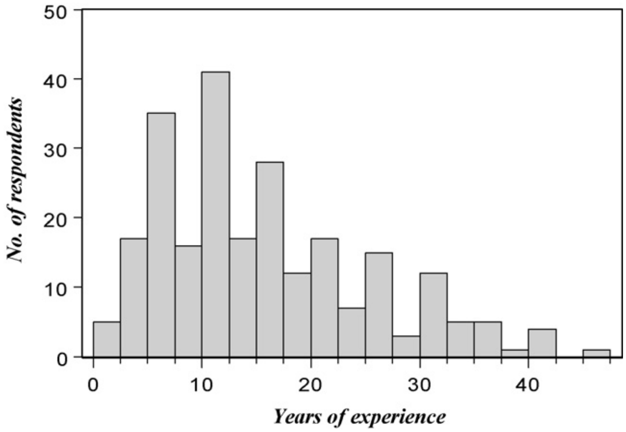

Fig. 1. Frequency distribution for years of experience in financial markets.

the future, as well as lower and upper bounds forming 90% confidence intervals, began to be solicited. The analysis here is restricted 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, the gender, age, educational level, educational field and professional experience of respondents are known—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.[10]

Table 1 provides a demographic snapshot. It is apparent that this group is predominately male and well-educated, tends to have had an educational specialization in economics or business, and does its own research.[11] On average, the group has good experience in financial markets. Since market experience (or years in markets) will be a key variable in the analysis, its distribution is shown in Fig. 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 properly calibrated in the sense that eventual *DAX* realizations fall within their *x*% confidence ranges *x*% of the time.

As mentioned earlier, there is evidence that experts in their fields (including finance) exhibit overconfidence, so we are not striking new ground here. Nevertheless, one might expect only moderate levels of overconfidence in the present context, 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 the sample do have extensive experience in financial markets. Hypotheses 2 and 3 are dynamic in nature, and focus on the cross-section. The former focuses on forecasting success:

[10] Turnover (i.e., a new person in the company responding to the survey) could color results. It seems to have been quite minimal, however: for example, during 2004 it was on average 1.17 per month.
[11] There are 30 *Ph.D.*s.

Hypothesis 2. Rational learning takes place, in the sense that after success, measured by *DAX* realizations falling within confidence ranges, confidence ranges contract, while after failure, measured by *DAX* realizations falling outside of confidence ranges, confidence ranges expand. Moreover, contrary to self-attribution bias, these adjustments are of the same magnitude.

The alternative hypothesis is that past success triggers self-attribution bias that exacerbates overconfidence, while failures are downplayed (Gervais and Odean, 2001). Because of this lack of symmetry, rational learning does not occur. Hypothesis 3 focuses on experience.

Hypothesis 3. Experience in financial markets reduces overconfidence.

Since survival in financial markets is a signal of persistent success, one 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. The model of Gervais and Odean (2001) suggests the latter is the weaker force, especially at low to moderate levels of experience.

4. Empirical results and discussion

4.1. *Hypothesis 1*

Survey respondents must produce 90% confidence intervals for the *DAX* 6 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 6 months ahead. If overlapping observations are used, 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 6 months to see if individuals are right, while eschewing overlapping observations. Unfortunately, for a barely minimal number of independent observations, one would have to possess 10 years of data (as opposed to just over 2 years).

An imperfect work-around begins by inferring 1-month forecast confidence intervals from 6-month intervals based on the assumption that forecasters believe that the growth rate in the *DAX* will be constant over the next 6 months.[12] 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 1-month forecast point estimate (F1) is calculated assuming constant growth in the *DAX*:

$$F1 = \left[\frac{F6}{Current\ Dax}\right]^{1/6} \times Current\ DAX \tag{1}$$

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 one has to do is multiply U6–F6 and F6–L6 by $1/\sqrt{6}$ to arrive at U1–F1 and F1–L1. And since F1 is known, it is straightforward to calculate L1 and U1. Finally, using L1 and U1, one merely has to note whether or not the eventual realization of the *DAX* falls between the bounds. If this technique is sound and respondents are well-calibrated, this should occur 90% of the time.

To consider the viability of this approach, suppose that the current level of the index is below a forecaster's estimate of intrinsic value. Assuming the expectation is for gradual convergence of value and price no sooner than 6 months in the future, then this technique will provide an unbiased 1-month ahead forecast. If convergence is expected to occur earlier than 6 months, then this technique will induce a downward bias. On the other hand, if the current level of the index is *above* a forecaster's estimate of intrinsic value, then some of the time an *upward* bias will be induced. We believe that if these two situations are equally likely, imputed 1-month ahead forecasts will be unconditionally unbiased estimates of true 1-month ahead forecasts. While true 1-month ahead forecasts would be preferable, they are not available and the resultant noise cannot be avoided.

One further complication has to be addressed. Respondents typically have about 2 weeks to make their 6-month ahead forecasts. Since they are not told to do otherwise, logically respondents would be making their forecasts for *exactly* 6 months in the future. Using the 1-month interpolation procedure described earlier, the derived values are then appropriate for exactly 1 month in the future. But 1 month in the future occurs on different dates for different people. So what one in fact has to do is use respondent-specific "current" *DAX* values and compare intervals to respondent-specific "eventual" *DAX* values.

While respondents should be right 90% of the time, Fig. 2 shows that in actual fact this occurs far less often than that.[13] Note that at best about 80% of respondents are "right," and quite often less than 50% are right. The inference is that the survey

[12] We also examined the time series of the history of the *DAX* in order to ascertain how wide a 90% confidence interval *should* be at each point in time, and then compared the width of respondent confidence intervals to this ideal width. On this basis, forecasters were also overconfident. One possible critique of this sort of test is that forecasters can use information other than the history of the *DAX* in constructing their intervals.

[13] 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.

Please cite this article in press as: Deaves, R., et al., The dynamics of overconfidence: Evidence from stock market forecasters. J. Econ. Behav. Organ. (2010), doi:10.1016/j.jebo.2010.05.001

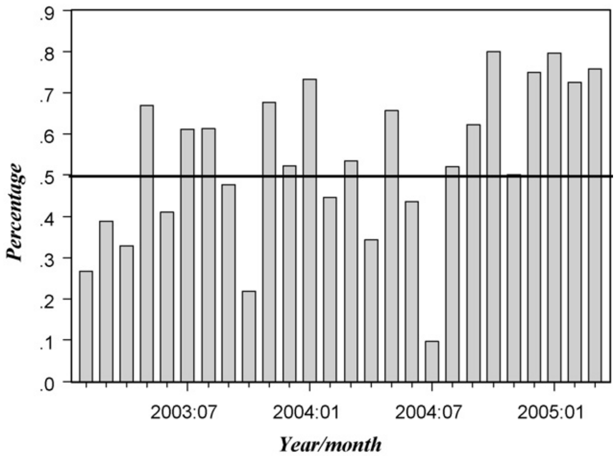

Fig. 2. Percentage of time that respondents' intervals contain eventual value of DAX (time series).

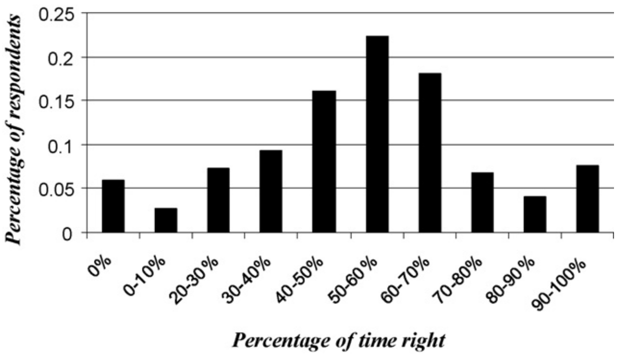

Fig. 3. Percentage of time that respondents' intervals contain eventual value of DAX (cross-sectional).

respondents are overconfident. Interestingly, it seems that respondents become more accurate during the last 6 months of the sample. Whether this is because conditions make their task easier or because learning is taking place is an open question.

It is also instructive to view overconfidence cross-sectionally. In Fig. 3, a frequency distribution of the percentage of times individuals are right over the full sample is shown.[14] The mode suggests that many are right between 40% and 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 conclude, consistent with previous evidence on the overconfidence of experts, the *ZEW* survey respondents are no different in this regard.[15] Substantially more than 10% of their imputed 90% confidence intervals fail to contain the eventual level of the *DAX*. There are, however, suggestive signs of some learning occurring by the end of the sample.

4.2. Hypothesis 2

We next consider 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 without a commensurate widening following failure. If so, respondents are learning to be overconfident. The maintained hypothesis is that rational learning occurs, suggesting that after failure intervals expand as much as they contract after success.

Using the technique described above, 1-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 1-month forecasts performed and adjust behavior in response. For this reason the associated noise might make it difficult to detect any innate dynamics.

[14] We also generated an analogous figure using only 6-month ahead non-overlapping forecasts. The results, available from the authors on request, are broadly similar.

[15] Not only are forecasters miscalibrated, they are also overly pessimistic. The average forecast error (forecast minus realization) is −40.5, which is highly statistically significant (p-val = 0.000).

Table 2
Pooled panel regressions of changes in relative confidence interval width on previous forecasting success and control variables $\Delta CIRW_{i,t} = a + bRIGHT_{i,t-1} + c\,\Delta\log(DAX_{i,t}) + d\,\Delta\log(VDAX_{i,t})$.

	Obs	a	b	c	D	R^2 (Adj. R^2)
Panel A						
Full sample	4966	0.046 (0.000) [0.000]	−0.09 (0.000) [0.000]	–	–	0.004 (0.004)
2003:03–2004:03	2737	0.048 (0.000) [0.000]	−0.103 (0.000) [0.000]	–	–	0.016 (0.016)
2004:04–2005:03	2229	0.043 (0.077) [0.042]	−0.075 (0.045) [0.043]	–	–	0.002 (0.001)
Panel B						
Full sample	4966	0.044 (0.000) [0.000]	−0.090 (0.000) [0.000]	−0.044 (0.875) [0.841]	−0.058 (0.694) [0.548]	0.004 (0.004)
2003:03–2004:03	2737	0.047 (0.000) [0.000]	−0.102 (0.000) [0.000]	0.151 (0.496) [0.389]	0.062 (0.579) [0.500]	0.017 (0.015)
2004:04–2005:03	2229	0.039 (0.130) [0.111]	−0.075 (0.059) [0.072]	−0.499 (0.642) [0.676]	−0.188 (0.568) [0.460]	0.002 (0.001)

Notes: $\Delta CIRW_{i,t}$ = change in width of confidence interval (relative to the cross-sectional average width) at time t for respondent i; $RIGHT_{i,t-1}$ = indicator variable equaling one when the respondent's prior interval contained eventual DAX value and zero otherwise; $\Delta\log(DAX_{i,t})$ = individual-specific change from month $t-1$ to t in log of level of DAX; $\Delta\log(VDAX_{i,t})$ = individual-specific change in expected volatility of log of DAX from month $t-1$ to t; last two variables are measured from 1 month *plus a day prior* to an individual's forecast to *1 day prior* to same individual's forecast; p-values are in parentheses/square brackets behind coefficient estimates; standard errors are estimated using panel-corrected approach of Beck and Katz (1995) which corrects either for heteroscedasticity or serial correlation: first figure (in parentheses) below estimate shows p-value in case of serial correlation correction, while second figure [in square brackets] shows p-value when standard errors are corrected for cross-sectional heteroscedasticity; there are 323 cross-sectional units in full sample.

Regressions of the following form are run:

$$\Delta CIRW_{i,t} = a + bRIGHT_{i,t-1} + c\,\Delta\log(DAX_{i,t}) + d\,\Delta\log(VDAX_{i,t}) \tag{2}$$

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; $RIGHT_{i,t-1}$ is an indicator variable equaling one when the respondent's *prior* interval contained the eventual DAX value and zero otherwise. Note that a equals the change based on wrong answers; $a+b$ equals the change for right answers. Individual-specific changes from t to $t-1$ in the log of the level of the DAX ($\Delta\log(DAX_{i,t})$), or returns, and comparable changes in the expected volatility of the DAX ($\Delta\log(VDAX_{i,t})$) are used as control variables.[16] Note that the latter is the implied volatility of option contracts on stocks included in the DAX.[17] It is important to note that, at the time respondents provide predictions, they know the values of these control variables, since these control variables are measured from 1 month plus a day prior to an individual's forecast to 1 day prior to the same individual's forecast. By including these control variables it is possible to test whether predictions of the participants are influenced by recent changes in stock market levels or volatility.

Panel A of Table 2 provides full-sample regression results without the inclusion of the control variables.[18] Additionally, the full sample is partitioned into roughly equal subsamples. Note that all estimations have been done applying a pooled regression approach. The standard errors have been corrected by using the method of "panel-corrected standard errors" of Beck and Katz (1995). Thus the standard errors are corrected either for heteroscedasticity or serial correlation, with the table providing p-values for both types of corrections. We also tested the validity of the pooled regression approach in contrast to random and fixed effects panel models. A Hausman test showed that the fixed effects model was preferable to the random effects model as the null hypothesis was rejected at the 5% level. But, as the null hypothesis that all fixed effects were jointly redundant could not be rejected at any conventional significance level, the pooled regression model is appropriate.

For the full sample as well as for the first subperiod, there is very strong evidence that forecasters adjust the width of their confidence intervals based on how well their intervals from the previous forecast performed.[19] Typically, when respondents get it wrong, they expand their intervals; when they get it right, they contract their intervals.[20] Even for the second subperiod the evidence is fairly solid, in that the slope coefficient is still significant at 1%. Panel B shows results when the two control variables are included as regressors. Both control variables do not have a significant influence on $\Delta CIRW_{i,t}$, and their inclusion only slightly changes the estimates (and significance) of the parameters a and b.

A few other salient points suggest themselves. First, these results are quite robust. Cross-sectional forecast-by-forecast regressions are run (without the control variables) and the coefficient estimates are shown in Fig. 4. 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 ($a+b<0$) and expand after failure ($a>0$).

[16] Though both the dependent and control variables are subscripted at t (the forecast month), all (or virtually all) of the change in the DAX (and its volatility) would have been available to market forecasters.

[17] For more information, see www.exchange.de. The approach for calculating $VDAX$ is similar to other volatility indices such as the VIX (on the *S&P 500*) or the VXD (on the *Dow*).

[18] We also generated an analogous panel using only 6-month ahead non-overlapping forecasts. The results, available from the authors on request, are broadly similar.

[19] 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.

[20] As for significance in an economic sense, for the full sample a has a value of 0.046 and b has a value of −0.090, implying a change of −4.4% (=4.6–9%) in the average confidence interval in the case of a correct interval forecast, and 4.6% in the case of an incorrect interval forecast.

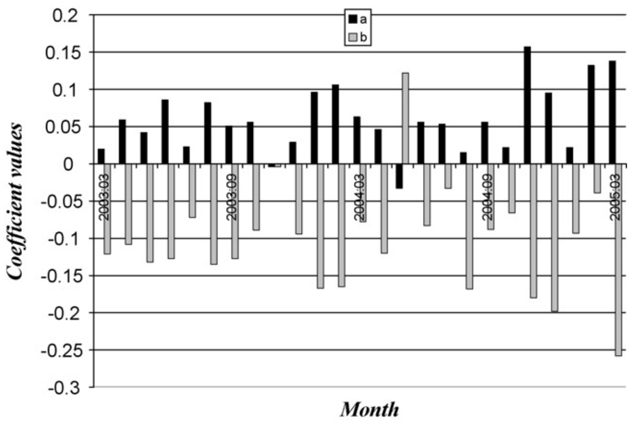

Fig. 4. Month-by-month intercept and slope coefficients for regression of changes in relative confidence interval width on previous forecasting success.

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).[21] Thus we cannot reject the maintained hypothesis that rational learning is taking place. Respondents, while on average remaining overconfident, are not learning to be more overconfident. 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 it was necessary to impute 1-month intervals from 6-month intervals obviously exacerbates the problem. Nevertheless, quite reliably, success, or its absence, is an important force.

Another perspective can be obtained at the level of the markets. While overconfidence can afflict *individuals*, it can also be endemic in *markets*. The most likely trigger is high past market returns. Statman et al. (2006) show that lagged market returns are correlated with changes in trading activity.[22] The argument is that higher market returns, though benefiting all, lead some investors to mistakenly attribute rising portfolio values to a higher level of information precision. This constitutes *indirect* evidence of an increase in overconfidence through self-attribution bias, since some behavioral models associate overconfidence with excessive trading.[23] 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. While it would clearly be best to examine the overconfidence of investors, given the data that we have at hand, we can use the overconfidence of stock market forecasters as a proxy. This of course relies on the notion that stock market forecasters are also investors and greater overconfidence in one realm is likely to spill over to greater overconfidence in other realms.

Along these lines, the dynamics of market-wide overconfidence is examined by considering whether high recent market returns induce survey respondents to become more overconfident. Returning to 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 (along with *DAX* volatility), we again note that the p-values for the latter are never close to indicating significance.

A decomposition of the variance of the full panel of imputed 1-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 one is more likely to find an impact on aggregate overconfidence if the dependent variable is the change in the *average* confidence interval width. For this reason, the following regression is run:

$$\Delta ACIW_t = a + bPRIGHT_{t-1} + c\,\Delta\log(DAX_t) + d\,\Delta\log(DAX_{t-1}) + e\,\Delta\log(VDAX_t) \tag{3}$$

where $\Delta ACIW_t$ is the change in the width of the average confidence interval at time t; $PRIGHT_{t-1}$ is the percentage of respondents whose *prior* intervals contained the eventual *DAX* value. The other previously defined variables have been averaged over all individuals at each point in time.

Table 3 confirms our suspicions.[24] The first regression investigates whether average overconfidence increases with the percentage of respondents who were previously right. This is analogous to Panel A of Table 2, but with averaging done over

[21] It is not possible to reject the hypothesis that $2a + b = 0$.
[22] Glaser and Weber (2009) 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.
[23] See Odean (1998).
[24] Note that the p-values shown in Table 5 are corrected for heteroscedasticity and serial correlation using the approach of Newey and West (1987).

Please cite this article in press as: Deaves, R., et al., The dynamics of overconfidence: Evidence from stock market forecasters. J. Econ. Behav. Organ. (2010), doi:10.1016/j.jebo.2010.05.001

Table 3
Time-series regressions of changes in average confidence interval width on previous average forecasting success and control variables $\Delta ACIW_t = a + bPRIGHT_{t-1} + c\,\Delta\log(DAX_t) + d\,\Delta\log(DAX_{t-1}) + e\,\Delta\log(VDAX_t)$.

Sample	a	b	c	D	e	R^2 (Adj. R^2)
2003:03-2005:03	−0.009 (0.146)	0.010 (0.297)	–	–	–	0.075 (0.035)
2003:03-2005:03	−0.009 (0.027)	0.011 (0.107)	−0.054 (0.064)		–	0.211 (0.139)
2003:03-2005:03	−0.007 (0.103)	0.007 (0.341)	−0.065 (0.037)	0.030 (0.270)	–	0.256 (0.150)
2003:03-2005:03	−0.006 (0.125)	0.005 (0.485)	−0.108 (0.026)	0.040 (0.153)	−0.024 (0.229)	0.309 (0.171)

Notes: $\Delta ACIW_t$ = change in width of average confidence interval at time t; $PRIGHT_{t-1}$ = percentage of respondents whose prior intervals contained eventual DAX value; $\Delta\log(DAX_t)$ = average of individual-specific changes from t to $t-1$ in log of level of *DAX*; $\Delta\log(VDAX_t)$ = average of individual-specific changes in expected volatility of log of *DAX*; p-values are in brackets behind coefficient estimates; standard errors have been adjusted for serial correlation and heteroscedasticity (using procedure of Newey and West, 1987).

all individuals. There is no discernible impact. Turning to the other regressions, as in Panel B of Table 2, changes in the level of the *DAX* along with changes in its volatility have been used as control variables. More precisely, individual-specific values are used, logs are taken and variables are differenced from the day before an individual's forecast to 1 month before. Then these cross-sectional variables are averaged over all individuals at each point in time to create time-series variables.

The second regression includes as an additional regressor past market returns, while the third includes both the latter and a once-lagged counterpart. Taken together, there is evidence that, despite the sparse number of time-series observations, high market returns in the recent past push the entire market towards greater overconfidence.[25] The fourth regression indicates that this finding is robust to the inclusion of *DAX* volatility as an additional control variable.[26]

In sum, while we are not able to reject rational learning in that confidence intervals expand after failure as much as they contract after success, evidence at the level of the market is less amenable to rationality. People seem to confound market success and personal success, leading to an increase in overconfidence.

4.3. Hypothesis 3

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, two of the questions were years of experience working in financial markets and age. As for the former, recall that the model of Gervais and Odean (2001) suggests that experience, at least up to a point, will exacerbate overconfidence.

In exploring this issue, it is possible to broaden the focus. Overconfidence can be partitioned into certainty (or, synonymously, confidence as opposed to overconfidence) and knowledge, where it is to be noted that overconfidence is the difference between certainty and knowledge.[27] The advantage of this partitioning is that it allows for a separate assessment of the impact of experience on the constituent components.

To more specifically define these components, certainty is expressed in terms of an individual's confidence interval width relative to the cross-sectional average at each point in time. A time-series average is then taken to obtain an individual-specific average relative confidence interval width. The greater is this value, the *less* confident (or certain) is the individual.[28] Thus the following metric for confidence is appropriate:[29]

$$\textit{Certainty}_i = \frac{Max(Avg\ rel\ conf\ range\ width) - Avg\ rel\ conf\ range\ width_i}{Max(Avg\ rel\ conf\ range\ width) - Min(Avg\ rel\ conf\ range\ width)} \tag{4}$$

25 While researchers have investigated the relationship between the conditional mean and the conditional variance of stock returns, the evidence is mixed (Glosten et al., 1993).

26 The value of c seems to be significant from an economic point of view. Compared to the average confidence interval (0.082), the *DAX* coefficient (between −0.054 and −0.108) is fairly high. The mean of $\Delta(\log DAX)$ is 0.01356 and its standard deviation is 0.0553, so a *DAX* change of one standard deviation unit above the mean leads to a change in *ACIW* of between −0.0037 and −0.0074 (which is about 4.5–9.0% of the mean of *ACIW*).

27 In Bhandari and Deaves (2006), 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.

28 Note that we do not use the average confidence interval for each respondent. One thing to keep in mind is that not all respondents always participate. Given that market volatility will be time-varying, and that intervals are likely to be wider when volatility is high, if a particular individual has mostly made forecasts during high-volatility months, then his average confidence interval will tend to be wide *for this reason*, while, if he has tended to make forecasts when volatility has been low, then his average confidence interval will tend to be narrow *for this reason*. By using at each point in time one's confidence interval width *relative* to the average at that point in time we have sought to control for time-varying volatility.

29 As an example, suppose someone, for a given forecast month, has a confidence range of 900, while the average over all forecasters for this forecast month is 1,000. Thus his relative confidence range width is 0.9. Further suppose there are only two other forecast months and this particular forecaster for these months has relative confidence range widths of 0.7 and 0.5. His overall average relative confidence range width would then be 0.7. To calculate his certainty variable, we order all individuals by average relative confidence range width. The person with the highest automatically has a certainty level of one, and the person with the lowest automatically has a certainty level of zero. If the highest is 1.7 and the lowest is 0.3, then the individual in question would have a certainty level of (1.7 − 0.7)/(1.7 − 0.3) = 0.714.

Please cite this article in press as: Deaves, R., et al., The dynamics of overconfidence: Evidence from stock market forecasters. J. Econ. Behav. Organ. (2010), doi:10.1016/j.jebo.2010.05.001

Table 4
Summary statistics and correlations for knowledge, confidence, and overconfidence.

	Correlation matrix		
	Knowledge	Certainty	Overconfidence
Knowledge	1	0.054	−0.507
Certainty	–	1	0.459
Overconfidence	–	–	1
	Other summary statistics		
	Knowledge	Certainty	Overconfidence
Mean	0.724	0.761	0.479
Median	0.740	0.788	0.500
Standard Dev.	0.119	0.129	0.236

Table 5
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)	Obs.
Certainty	0.750 (0.000)	0.0009 (0.232)	–	0.005 (0.001)	239
(Age ortho.)	0.750 (0.000)	0.0008 (0.291)	0.0015 (0.471)	0.007 (−0.002)	231
(Years in markets ortho.)	0.720 (0.000)	−0.0004 (0.843)	0.0015 (0.195)	0.007 (−0.002)	231
Knowledge	0.750 (0.000)	−0.0019 (0.034)	–	0.050 (0.046)	240
(Age ortho.)	0.760 (0.000)	−0.0019 (0.041)	−0.0002 (0.809)	0.052 (0.044)	232
(Years in markets ortho.)	0.800 (0.000)	−0.0017 (0.095)	−0.0002 (0.060)	0.052 (0.044)	232
Overconfidence	0.431 (0.000)	0.0041 (0.002)	–	0.040 (0.036)	239
(Age ortho.)	0.429 (0.000)	0.0041 (0.003)	−0.0022 (0.518)	0.042 (0.034)	231
(Years in markets ortho.)	0.350 (0.000)	0.0058 (0.054)	0.0035 (0.030)	0.042 (0.034)	231

Notes: p-Values are in brackets behind coefficient estimates; standard errors have been estimated using White correction for heteroscedasticity.

Knowledge is based on average absolute forecast errors scaled by *DAX* realizations.[30] It is defined as follows:

$$Knowledge_i = \frac{Max(Avg\,abs\,scaled\,fore\,error) - Avg\,abs\,scaled\,fore\,error_i}{Max(Avg\,abs\,scaled\,fore\,error) - Min(Avg\,abs\,scaled\,fore\,error)} \quad (5)$$

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 0.1 would reflect neither underconfidence nor overconfidence, but values above this threshold reflect overconfidence.

In Table 4, summary statistics for these variables are provided. Notably, the mean and median respondents are overconfident. Given that overconfidence is the difference between certainty and knowledge, it is not surprising to see that overconfidence is positively correlated with certainty and negatively correlated with knowledge. In contrast, certainty and knowledge are virtually uncorrelated.

Table 5 reports the results of the following cross-sectional regressions:

$$DV_i = a + b \times Years\,in\,markets_i + c \times Age_i \quad (6)$$

where DV_i, the dependent variable, is certainty, knowledge or overconfidence; the cross-sectional independent variables are self-explanatory. Focusing first on the top row of each panel where experience is the sole independent variable, it is apparent that experience is detrimental in that it increases overconfidence. To interpret, 3 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 the results are otherwise. The only significant coefficient is on knowledge, with the interpretation here being that 5 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, there is some evidence in the psychology literature indicating that achievement motivation declines with maturation (Mehrabian and Blum (1996)), and the importance attached to high job demands decreases over time (Warr, 1997).[31] We conjecture that the second effect outweighs the first, causing more experienced people to display lower levels

[30] The reason for scaling absolute forecast errors is that, when the *DAX* is high, absolute forecast errors are likely to be high, and when the *DAX* is low, absolute forecast errors are likely to be low. By scaling by the level of the *DAX* we have in essence converted things from nominal to real magnitudes.
[31] There is also evidence running counter to this. See Kooij et al. (2008) for a discussion and citations.

of knowledge. Finally, since conflicting effects may induce non-linearity, we tested for the latter, but were not able to reject linearity.[32]

Since job experience and age are highly correlated (0.85), it is natural to wonder if it is age rather than experience that is driving the results. To resolve this issue, in the bottom two rows of each panel of Table 5, two regressions where both age and years of experience are independent variables are displayed. In the first, age is orthogonalized, while, in the second, experience is orthogonalized. The advantage of this approach is it allows one to see which variable is the key driver. It turns out to be experience rather than age, since experience (unlike age) is at least marginally significant in the knowledge and overconfidence regressions whether it or age is orthogonalized.

Since demographic data beyond years of experience and age were available, an investigation of the potential impact of other factors on overconfidence (as well as on its components, certainty and knowledge) is possible. There is evidence that certain personal characteristics are correlated with overconfidence. For example, males tend to be more overconfident than females (Lundeberg et al., 1994). Barber and Odean (2001) attribute the greater trading activity of men to this factor. And Bhandari and Deaves (2006) find that highly educated well-paid males are especially susceptible. The other demographic factors of interest available to us were gender, educational level, area of study and whether or not one generated the forecast oneself. Multivariate regressions of certainty, knowledge and overconfidence on all these demographic variables (as well as on age and experience) were performed, and, in all cases, the additional regressors were safely insignificant. At the same time, the impact of experience was easily retained.[33]

To conclude this section, contrary to the maintained hypothesis, experience in financial markets leads to increased overconfidence. It turns out that, inconsistent with self-attribution bias, this operates through reduced knowledge rather than through increased certainty.

5. Conclusion

The present research has explored the statics and dynamics of miscalibration-based overconfidence. Not surprisingly given previous work on experts, the *Finanzmarkttest* respondent group was shown to be overconfident. Nevertheless, some rational learning does seem to have taken place as respondents increase the widths of their confidence intervals with failure as much as they decrease them with success. Less amenable to rationality though is evidence at the level of the market, with high past returns inducing increased aggregate overconfidence. Also pointing away from rational learning is the fact that market experience does not lead to better calibration, though self-attribution bias does not seem to be the driving force behind this finding.

In short, the evidence on the dynamics of overconfidence is mixed and subject to interpretation. Nevertheless, given that there is abundant evidence that overconfidence is a pervasive phenomenon and certainly difficult to “debias” (Fischhoff, 1982), and that so many financial behaviors have been linked theoretically and empirically to the phenomenon, continued research is without doubt called for.

Acknowledgements

The authors would like to thank Patric Anderson, Markus Glaser, Elias Rantapuska and Martin Weber, along with seminar participants at the Financial Management Association meeting in Siena, Italy, the European Financial management Association meeting in Durham, U.K., the Center for European Economic Research (*ZEW*) in Mannheim, Germany, the Deutsche Bundesbank in Frankfurt, Germany, Simon Fraser University in Vancouver, Canada and Thammasat University in Bangkok, Thailand for helpful comments. All remaining errors are the responsibility of the authors.

References

Armor, D., Taylor, S., 2002. When predictions fail: the dilemma of unrealistic optimism. In: Gilovich, T., Griffin, D., Kahneman, D. (Eds.), Heuristics and Biases: The Psychology of Intuitive Judgment. Cambridge University Press, Cambridge, pp. 334–347.
Barber, B., Odean, T., 2000. Trading is hazardous to your wealth: the common stock investment performance of individual investors. Journal of Finance 55, 773–806.
Barber, B., Odean, T., 2001. Boys will be boys: gender, overconfidence, and common stock investment. Quarterly Journal of Economics 116, 261–292.
Beck, N., Katz, J., 1995. What to do (and not to do) with time-series cross-section data. American Political Science Review 89, 634–647.
Ben-David, I., Graham, J., Harvey, C., 2007. Managerial overconfidence and corporate policies. NBER Working Paper 13711.
Bhandari, G., Deaves, R., 2006. The demographics of overconfidence. Journal of Behavioral Finance 7, 5–11.
Camerer, C., Lovallo, D., 1999. Overconfidence and excess entry: an experimental approach. American Economic Review 89, 306–318.
Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and investor under- and overreactions. Journal of Finance 53, 1839–1886.
Daniel, K., Hirshleifer, D., Subrahmanyam, A., 2001. Mispricing, covariance risk, and the cross section of security returns. Journal of Finance 5, 921–965.
Deaves, R., Lüders, E., Luo, G., 2008. An experimental test of the impact of overconfidence and gender on trading activity. Review of Finance 13, 555–575.

[32] We tested whether a linear specification of the equations was appropriate by applying a RESET test. As the included quadratic terms of the three dependent variables were insignificantly different from zero, linearity could not be rejected. To take another perspective, the correlation between years of experience and the latter squared was 0.96. This implies that introducing a non-linear term would lead to severe collinearity.

[33] It may be a little surprising that gender did 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 financial industry are likely a more overconfident group than the general female population (Deaves et al., 2008).

Festinger, L., Carlsmith, J., 1959. Cognitive consequences of forced compliance. Journal of Abnormal and Social Psychology 58, 203–210.
Fischhoff, B., 1982. Debiasing. In: Kahneman, D., Slovic, P., Tversky, A. (Eds.), Judgment Under Uncertainty: Heuristics and Biases. Cambridge University Press, Cambridge, pp. 422–444.
Fischhoff, B., Slovic, P., Lichtenstein, S., 1977. Knowing with certainty: the appropriateness of extreme confidence. Journal of Experimental Psychology 3, 552–564.
Forsythe, R., Nelson, F., Neumann, G., Wright, J., 1992. Anatomy of an experimental political stock market. American Economic Review 82, 1142–1161.
Gervais, S., Odean, T., 2001. Learning to be overconfident. Review of Financial Studies 14, 1–27.
Gigerenzer, G., 1991. How to make cognitive illusions disappear: beyond heuristics and biases. European Review of Social Psychology 2, 83–115.
Glaser, M., Weber, M., 2009. Which past returns affect trading volume? Journal of Financial Markets 12, 1–31.
Glaser, M., Langer, T., Weber, M., 2005. Overconfidence of professionals and laymen: individual differences within and between tasks? Working paper. Available at http://www.papers.ssrn.com/sol3/papers.cfm?abstract_id=712583.
Glosten, L., Jagannathan, R., Runkle, D., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. Journal of Finance 48, 1779–1801.
Goetzmann, W., Kumar, A., 2008. Equity portfolio diversification. Review of Finance 12, 433–463.
Graham, J., Harvey, C., 2003. Expectations of equity risk premia, volatility and asymmetry. Working paper. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=292623.
Hirshleifer, D., Luo, G., 2001. On the survival of overconfident traders in a competitive securities market. Journal of Financial Markets 4, 73–84.
Kidd, J., 1970. The utilization of subjective probabilities in production planning. Acta Psychologica 34, 338–347.
Kooij, D., de Lange, A., Jansen, P., Dikkers, J., 2008. Older workers' motivation to continue to work: five meanings of age. Journal of Managerial Psychology 24, 364–394.
Kyle, A., Wang, F., 1997. Speculation duopoly with agreement to disagree: can overconfidence survive the market test? Journal of Finance 52, 2073–2090.
Langer, E., 1975. The illusion of control. Journal of Personality and Social Psychology 32, 311–328.
Lichtenstein, S., Fischhoff, B., Phillips, L., 1982. Calibration of probabilities: the state of the art to 1980. In: Kahneman, D., Slovic, P., Tversky, A. (Eds.), Judgment Under Uncertainty: Heuristics and Biases. Cambridge University Press, Cambridge, pp. 306–334.
Lundeberg, M., Fox, P., Punccohar, J., 1994. Highly confident but wrong: gender differences and similarities in confidence judgments. Journal of Educational Psychology 86, 114–121.
Malmendier, U., Tate, G., 2005. CEO overconfidence and corporate investment. Journal of Finance 60, 2661–2700.
Mehrabian, A., Blum, J., 1996. Temperament and personality as functions of age. International Journal of Aging and Human Development 42, 251–269.
Miller, D., Ross, M., 1975. Self-serving bias in attribution of causality: fact or fiction? Psychological Bulletin 82, 213–225.
Newey, W., West, K., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.
Odean, T., 1998. Volume, volatility, price and profit when all traders are above average. Journal of Finance 53, 1887–1934.
Oskamp, S., 1965. Overconfidence in case-study judgments. Journal of Consulting Psychology 29, 287–293.
Russo, J., Schoemaker, P., 1992. Managing overconfidence. Sloane Management Review 33, 7–17.
Staël von Holstein, C.-A., 1972. Probabilistic forecasting: an experiment related to the stock market. Organizational Behavior and Human Performance 8, 139–158.
Statman, M., Thorley, S., Vorkink, K., 2006. Investor overconfidence and trading volume. Review of Financial Studies 19, 1531–1565.
Svenson, O., 1981. Are we all less risky and more skillful than our fellow drivers? Acta Psychologica 47, 143–148.
Warr, P., 1997. Age, work and mental health. In: Schaie, K., Schooler, W. (Eds.), The Impact of Work on Older Adults. Springer, New York, pp. 252–303.

Please cite this article in press as: Deaves, R., et al., The dynamics of overconfidence: Evidence from stock market forecasters. J. Econ. Behav. Organ. (2010), doi:10.1016/j.jebo.2010.05.001