Male Analysts are Overconfident while Female Analysts are Not*

Katrien Bosquet† Peter de Goeij‡ Kristien Smedts§

January 14, 2010

work in progress - please do not quote without permission

Abstract

To study gender differences in overconfidence, we use a two-stage decision model. This model allows for the coexistence of an overconfidence bias and a strategically-driven bias in the forecasting process. Our results confirm the gender stereotype that male analysts are overconfident, while female analysts are not. Male financial analysts overweigh the precision of their private information by about 3%, believing too much in their own abilities. Female analysts, on the other hand, do not attach a weight to their personal information that is too high compared to the rational Bayesian weight. Moreover, both male and female analysts strategically inflate their earnings forecasts, but this strategic behavior disappears after the implementation of the new analyst regulation 2002.

*This version: December 17, 2009. We thank Bertrand Melenberg and seminar participants at Tilburg University for their comments.

†Catholic University of Leuven, AFI, Naamsestraat 69, 3000 Leuven, Belgium. Email: katrien.bosquet@econ.kuleuven.be, tel: (+) 32 (0)16 326675, fax: (+) 32 (0)16 326732. Part of this research is done while Katrien Bosquet was visiting Tilburg University.

‡Tilburg University, Department of Finance and CentER for Economic Research, P.O. Box 90153, 5000LE Tilburg, The Netherlands, email: p.c.degoeij@uvt.nl, phone: (+) 31 (0)13 4662083, fax: (+) 31 (0)13 4662875.

§Catholic University of Leuven, AFI, Naamsestraat 69, 3000 Leuven, Belgium. Email: kristien.smedts@econ.kuleuven.be, tel: (+) 32 (0)16 326739, fax: (+) 32 (0)16 326732.

1 Introduction

In recent years, financial economists are focused on studying psychological phenomena in market participants' behavior, and a shift from the neoclassical paradigm towards a more behaviorally-based paradigm can be observed. As noted by Shefrin (2008), the basis for the debate about the paradigm shift can be found in the decision making process of economic agents. Traditionally, financial economists model this decision making process on assumptions of rationality. More recently, it is acknowledged that behaviorally based assumptions are often more realistic. This paper focusses on one of these psychological finance theories, namely overconfidence. While many definitions of overconfidence exist, the overconfidence hypothesis relates to the overestimation of one's own abilities.

In a finance context, overconfidence is the behavioral bias that is most documented. Barber and Odean (2001) show overconfident investors trade excessively hurting their performance. Hilary and Menzly (2006) provide evidence that analysts with a short-lived past success become overconfident in their forecasting ability. The current paper also analyzes overconfidence in financial analysts' decision making. However, we do not analyze the mere presence of overconfidence among financial analysts, but we study its gender stereotype suggesting that men are more overconfident than women.

Many existing studies show that men are more overconfident than women. Estes and Hosseini (1998) suggest that the most important factor for explaining investment decision confidence is the decision maker's gender. Deaux and Farris (1977) find that men claim more ability than do women and Prince (1993) reveals that men feel more competent than women with regard to financial matters. While previous studies are often performed on small-scale experimental data, we contribute to the existing evidence by using a large real-life dataset of financial analysts' earnings forecasts. To study gender differences in overconfidence, we use the two-stage model of Bosquet et al. (2009). Following Daniel et al.(1998), this model defines an overconfident analyst as one who overestimates the precision of his private information signal. The two-stage model we use, is an extension of Chen and Jiang (2006) by simultaneously allowing for an overconfidence bias, as well as a strategically-driven bias in the forecasting process.

We confirm the gender stereotype that male analysts are overconfident while female analysts are not. Male financial analysts overweigh the precision of their private information by about 3%, believing too much in their own ability to process information. Female analysts do not behave in the same way, and attach a weight to their personal information that is in line with the rational Bayesian weight. Moreover, this male and female behavior is significantly different from each other. We do observe a slight convergence in behavior after 2002. However, the estimation results show that male analysts are still overconfident in recent years, while female analysts are not. Finally, both male and female analysts strategically inflate their earnings forecasts, but this bias seems to be eliminated by the new analyst regulation implemented in 2002.

The remainder of this paper is organized as follows. Section 2 briefly explains the two stage model for the decision making process of financial analysts. Section 3 presents the data and descriptives. The empirical results are discussed in Section 4, and finally, section 5 concludes.

2 The Analyst's Forecasting Model

To study gender differences in overconfidence, we use a two-stage model of Bosquet et al. (2009) that defines financial analysts' earnings forecast formation. Following Daniel et al. (1998), this model defines an overconfident analyst as one who overestimates the precision of his private information signal. The two-stage model we use, is an extension of Chen and Jiang (2006) by simultaneously allowing for an overconfidence bias (an unconscious bias), as well as a strategically-driven bias (a conscious bias) in the forecasting process. In particular, in the first stage a financial analyst collects private and public information and processes the data to obtain his earnings forecast. In arriving at his judgement, the analyst is likely unconsciously influenced by a behavioral bias, such as overconfidence. Once this fundamental analysis is complete, a financial analyst can, in the second stage, consciously deflate or inflate his forecast for strategic reasons (conflicts of interest). We provide an overview of the model below.[1]

The analyst's best conditional forecast of actual earnings (a) given his private information (x) and his public information (c), using Bayes' rule, is as follows:

$$E\left[a \mid x, c\right] = hx + (1-h)\,c \tag{1}$$

where $h \in [0,1]$ is the rational Bayesian weight for the analyst's private information. When an analyst processes publicly available information and weights his private information, behavioral biases could lead him to use a personal weighting scheme:

$$F = kx + (1-k)\,c \tag{2}$$

with $k \in [0,1]$ and F the first stage earnings forecast. This first stage earnings forecast is not revealed to the public. The degree of overconfidence is captured by the ratio of the weights $\frac{k}{h}$. When the personal weighting scheme corresponds to the rational scheme $\frac{k}{h} = 1$ and there is no behavioral bias. However, when the analyst overestimates his private information signal $\frac{k}{h} > 1$ reflecting an overconfident financial analyst. Alternatively,

[1] For a thorough understanding of the model we refer to the appendix or the paper of Bosquet et al. (2009).

when the private information signal is underestimated $\frac{k}{h} < 1$ and the analyst is herding.

Apart from the unconscious bias in the analyst's forecast, strategic incentives might induce the analyst to update his initial forecast F by a multiplication factor s:

$$f = sF = s\left[kx + (1-k)c\right] = skx + (s - sk)c \tag{3}$$

with f the earnings forecast issued by the financial analyst and observed by the public. The direction of change is determined by the size of s. $s > 1$ corresponds to an inflation of the initial forecast, while $s < 1$ reflects a deflation of the initial forecast.

By comparing the forecasts that are issued by male analysts with those that are issued by female analysts, this model sheds light on gender differences in overconfidence and strategic incentives. Based on the gender literature, we hypothesize that male analysts are more overconfident compared to female analysts, while we have no gender-driven priors with respect to strategic incentives:

$$H_1 \quad : \quad \left(\frac{k}{h}\right)_{male} > \left(\frac{k}{h}\right)_{female} \tag{4}$$

$$H_2 \quad : \quad s_{male} = s_{female} \tag{5}$$

To test our hypotheses, we use following reduced form equation:

$$\begin{aligned} E\left[f - a \mid x, c\right] &= \left(\frac{k-h}{k}\right)(f-c) + \left(\frac{hs-h}{sk}\right) f \\ &= \beta(f-c) + \gamma f \end{aligned} \tag{6}$$

This reduced form allows for a clear identification of the degree of overconfidence with $\frac{k}{h} = 1 + \beta$ and the size of strategic incentives as $s = 1 + \gamma$.[2] To obtain estimates for the model parameters $\frac{k}{h}$ and s, following regression equation is considered:

$$FE_{ikt} = \alpha + \beta Dev_{ikt} + \gamma f_{ikt} + \delta_i X_{ikt} + \varepsilon_{ikt} \tag{7}$$

with i an analyst identifier (who issues the forecast), k a company identifier (on which company a forecast is issued) and t a time identifier (to which quarter the forecast pertains). $FE_{ikt} = f_{ikt} - a_{kt}$ is the forecast error made by the individual analyst i for quarter t for company k, calculated as the difference between the forecasted f_{ikt} and actual a_{kt} earnings per share. $Dev_{ikt} = f_{ikt} - c_{k\tau}$ is the deviation from the consensus and is determined as the difference between the analyst's forecast f_{ikt} and the consensus forecast $c_{k\tau}$ concerning the earnings per share. The consensus forecast proxies for

[2] This is obtained by applying a first-order Taylor approximation to both coefficients in equation (6) around $\frac{k}{h} = 1$ and $s = 1$

available public information up to the point in time an analyst issues his forecast. This consensus forecast at time τ (within the quarter) is calculated as the mean of all the analysts' estimates up to τ, excluding the estimate of the analyst i, who makes his estimate for quarter t (analogous to Chen and Jiang, 2006 and Zitzewitz, 2001). The forecast error, the deviation from consensus as well as the forecast the analyst makes are deflated by the share price. Following Clement (1999), Size of the covered firm ($Size_{kt}$), general ($TotExp_{it}$) and firm specific experience ($FirmExp_{ikt}$), two measures of task complexity ($FirmCompl_{it}$ and $IndCompl_{it}$) and forecast age (Age_{ikt}) are added as control variables (X_{ikt}). A full description of all variables and their summary statistics can be found in Table 3, in the Appendix. More information on the estimation technique for equation (7) follows in Section 4 below.

3 Data and Descriptive Statistics

Analysts' quarterly earnings forecasts and stock price data are obtained from the Institutional Broker Estimate System (I/B/E/S) database, part of Thomson Financial. The earnings forecasts pertain to the period 1996 until 2006. The database is restricted to highly covered United States companies with fiscal year end December. High coverage is ensured by demanding a minimum average coverage of three analysts, and deleting firms which have an average market capitalization below \$100 million, an average market price below \$5, or a market price below \$1. Furthermore, the dataset is stripped from errors and potential companies in difficulties. As quarterly filings must be with the SEC within 45 days subsequent to the end of a quarter, observations of companies reporting later than 45 days after the last day of the end of the quarter, are eliminated. When revisions are made, only the first forecast is kept (Chen and Jiang, 2006; Hilary and Menzly, 2006). The first forecast is timely and thus most useful to investors. Moreover, we drop negative earnings forecasts, ensuring a sample of forecasts with maximum effort and processing of information. Hayes (1998) argues that the incentive to gather information are most intense for stocks that are anticipated to give strong performance and McNichols and O'Brian (1997) indicate that analysts drop stocks with unfavorable future prospects. Also, the dynamics and interpretation of the behavioral process as well as the strategic incentives can differ for negative and positive forecasts. The final dataset contains 322,123 quarterly earnings forecasts.

To examine differences between male and female analysts, the gender of each analyst is determined. In the I/B/E/S Brokers Translation File every analyst is listed by last name and first initial. This dataset is merged with data from the corresponding annual edition of Nelson's Directory of Investment Research. This way the analyst's full name can be obtained on which basis gender is determined. We rely on a program that uses

Google's database to analyze common patterns involving a first name[3]. The program determines whether the first name is more commonly used for a man or for a woman. If there is uncertainty about the gender of an analyst, the history of that analyst is examined using the Internet to try to find out whether the analyst is male or female. For approximately 95% of the observations in the full dataset it is possible to determine the gender of the analyst.[4] The other 5% of observations is removed from the dataset. The gender dataset contains 302,475 observations of 5,327 analysts on 2,770 firms. Male analysts issued 85% of the earnings forecasts, while the remaining 15% is issued by female analysts.

To obtain a first insight into potential gender differences in overconfidence, Figure 1 plots the forecast accuracy of male versus female analysts, which is measured as the ratio of absolute forecast error of female analysts and the absolute forecast error of male analysts. A ratio smaller than one, indicates that female analysts are more accurate compared to male analysts. Hilary and Menzly (2006) show that forecast accuracy and overconfidence are negatively related, implying that overconfidence hurts forecasting performance. Figure 1 shows that before 2002, female analysts are more accurate than male analysts. A superior forecasting ability of female analysts is also found by Kumar (2009). This figure suggests that male analysts are likely more prone to overconfidence compared to female analysts for that period, confirming the gender stereotype. After 2002, the results indicate the opposite with male analysts being more accurate than female analysts. This possibly reflects a change in analysts' forecasting behavior, driven by the new financial analyst regulation (NASD rule 2711) that is enforced in May 2002 and is targeted at reducing conflicts of interest. Before 2002, financial analysts resorted to strategical behavior to please management of the stock covered (Ljunqvist et al., 2007). The next section provides a more in-depth analysis of the forecasting behavior and its evolution over time using the two-stage model of Section 2.

4 Empirical Results

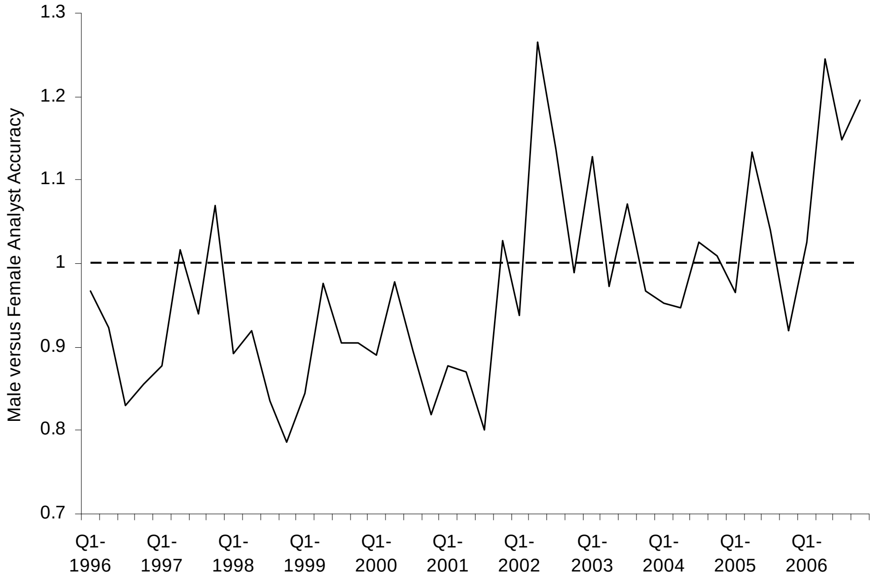
The gender dataset contains financial analysts' earnings forecasts for a particular stock at a certain point in time. This three way panel possibly contains unobserved effects such as analyst, time, firm and industry effects. To control for these unobserved effects, Petersen (2008) argues that OLS with clustered standard errors (if necessary multi-way) is the best estimation method. Comparing clustered standard errors, in each dimension or multiple dimensions, with White standard errors, we conclude that standard errors

[3] http://www.gpeters.com/names/baby-names.php

[4] Also Green et al. (2009) and Kumar (2009) are able to match approximately 95% of the observations with gender.

Figure 1: **Absolute Forecast Error of Male and Female Analysts**

The figure plots the forecast accuracy of male versus female forecasts over the time period 1996-2006. Forecast accuracy is computed as the ratio of the absolute forecast. error of female analysts and the absolute forecast error of male analysts.



clustered by business group are sufficient.[5] Moreover a firm fixed effects estimation is used to allow for more efficient parameter estimates. To investigate gender differences in overconfidence ($\frac{k}{h}$) a dummy approach is used. Compared to a split sample approach, such dummy approach has the advantage of allowing for a covariance between the male and female observations.

Given the observed trend break in forecasting accuracy of male versus female analysts, we estimate the forecasting model for the full sample Q1-1996 until Q4-2006, as well as for the subsamples Q1-1996 until Q1-2002 (pre-2002 sample) and Q2-2002 until Q4-2006 (post-2002 sample). This break-up of our sample also captures the change in new analyst regulation that is enforced in may 2002. The full sample estimation results can be found in Table 1.

The table reports the estimates of the weighting factor $\frac{k}{h}$, capturing overconfidence, and the strategic incentives factor s. These estimates are obtained from the reduced form estimation equation (7), whose summary statistics and estimation results can be found in the Appendix in Table 3 and Table 4, respectively. The delta method is used to

[5] Using the proprietary classification scheme of Thomson Financial to categorize companies into homogenous groups according to business lines results into 211 clusters. This group classification allows for a sufficient number of clusters, which is necesary to obtain unbiased standard errors (see Thompson 2006 and Petersen 2008).

Table 1: **Estimation of Overconfidence and Strategic Behavior (1996-2006)**

This table presents estimation results for the weighting factor $\frac{k}{h}$ and the strategic factor s, defined in equation (6). These factors are extracted from the reduced form estimation equation (7). For both factors a two sided hypothesis test of whether they are significantly different from 1 is performed. For the difference with respect to the weighting factor a one sided t-test is performed to determine whether male analysts are more overconfident than female analysts. For the difference with respect to the strategic factor a two sided t-test is performed to determine whether male analysts behave strategically different from female analysts. *,**,*** indicate significance at the 10%, 5% and 1% level, respectively. The sample runs from Q1-1996 until Q4-2006.

	Male			Female			Ind. Diff.
	Coeff.	s.e	t-value	Coeff.	s.e.	t-value	t-value
$\frac{k}{h}$	1.027	0.007	4.132***	1.008	0.012	0.646	1.434*
s	1.026	0.014	1.772*	1.037	0.018	2.063**	-0.501

obtain the appropriate standard errors. The estimates in Table 1 shows that male analysts are overconfident. Compared to the rational Bayesian weight, they overestimate their private information by 2.7%. This result is significant at the 99% confidence level. For female analysts the degree of overweighting is estimated at only 0.8%, and is not statistically significant. This leads us to conclude that the female personal weighting scheme is close to the rational scheme. Moreover, the table also shows that the hypothesis (4) that male analysts are more overconfident than female analysts holds. The test whether the male and female estimates of overconfidence are significantly different from each other is not rejected at the 90% confidence level. The gender stereotype that males are overconfident compared to females is thus confirmed in our analyst sample. While female analysts use rational Bayesian weights in processing public and private information, male analysts attach a weight to their personal information that is too high. They believe, too much, in their own abilities and information processing skills.

The table also reports the factor reflecting strategic incentives of financial analysts. For male as well as female analysts we find a significantly positive strategic factor (significant at the 90% and 95% confidence level respectively). While female analysts inflate their forecast more than male analysts (3.7% versus 2.6%), hypothesis (5) which states that there is no gender difference in strategic behavior cannot be rejected. This is also as expected as the existing gender literature gives us no prior with respect to possible differences in strategic behavior.

As mentioned above, the forecasting model is also estimated for two subsamples. The estimation results are shown in Table 2. The corresponding reduced form estimates can be found in Table 5 in the Appendix. Our main conclusions based on the full sample estimations also hold in the subsamples, but some interesting dynamics are uncovered. The estimation results indicate that for the pre-2002 period and the post-

Table 2: **Estimation of Overconfidence and Strategic Behavior (Pre-2002 and Post-2002)**

This table presents estimation results for the weighting factor $\frac{k}{h}$ and the strategic factor s, defined in equation (6). These factors are extracted from the reduced form estimation equation (7). For both factors a two sided hypothesis test of whether they are significantly different from 1 is performed. For the difference with respect to the weighting factor a one sided t-test is performed to determine whether male analysts are more overconfident than female analysts. For the difference with respect to the strategic factor a two sided t-test is performed to determine whether male analysts behave strategically different from female analysts. To test for a structural change over time in the behavior of the male or female analysts, two sided t-tests are performed for the overweighting and the strategic incentives' factor. *,**,*** indicate significance at the 10%, 5% and 1% level, respectively. The subsamples are defined as Q1-1996 until Q1-2002 (pre-2002) and Q2-2002 until Q4-2006 (post-2002).

	Male			Female			Ind. Diff.
	Coeff.	s.e	t-value	Coeff.	s.e.	t-value	t-value
				Pre-2002			
$\frac{k}{h}$	1.037	0.012	2.968***	1.006	0.013	0.446	1.686**
s	1.046	0.022	2.098**	1.063	0.020	3.123***	-0.599
				Post-2002			
$\frac{k}{h}$	1.024	0.009	2.719***	1.009	0.018	0.499	0.722
s	1.004	0.020	0.218	1.014	0.025	0.555	-0.301
	Ind. Diff . over time			Ind. Diff . over time			
	t-value			t-value			
$\frac{k}{h}$	-0.830			0.143			
s	-1.394			-1.514			

2002 period, male financial analysts significantly overweigh private information. In both cases the estimates are significant at the 1% significance level. Across time, a decrease in overweighting of private information from 3.7% to 2.4% for male analysts is observed. In the more recent subsample, male analysts thus become less overconfident. However, the change in behavior over time is not statistically significant. For female analysts, we find that they do not overweigh private information, neither in the pre-2002 period nor in the post-2002. For both subsample, female analysts thus process information in line with the rational Bayesian weights. Even though the overweighting factor slightly increases from 0.6% to 0.9%, this increase is not statistically significant. Moreover, while the full sample results indicate that the male and female overweighting is estimated significantly different from each other, the subsample estimates show that this is completely driven by a different weighting behavior in the period before 2002. For the first subsample, we confirm the hypothesis (4) at the 5% significance level, but this is no longer true fro the second subsample. This indicates that male and female behavior has converged, although at a small (insignificant) pace.

The table also gives the results of the strategic incentive parameters of male and female analysts over time. Again we observe that the full sample results are mainly driven by the pre-2002 period. Before 2002, both male and female analysts inflate their forecasts for strategic reasons. Male analysts inflate their fundamental forecast by 4.6%, while female analysts are even more responsive to strategic incentives resulting in an inflation factor of 6.3%. Similar to our complete sample findings, we cannot reject the hypothesis of no gender differences in strategic behavior. However, after 2002 this strategic inflation behavior disappears by analysts disappears. The new regulation that is enforced in May 2002 is targeted at reducing conflicts of interest and is clearly very successful at it. Apparently, regulation is effective as it makes analysts more immune to inflate their fundamental forecast to e.g. please management or generate trading commission.

To sum up, the empirical results show that male analysts are overconfident while female analysts are not. According to Hilary and Menzly (2006), overconfidence hurts an analyst's performance. If overconfidence is the main driver of forecast accuracy, female analysts should be more accurate than male analysts throughout the entire forecasting period. However, looking back at Figure 1, we observe a switch in relative forecast accuracy. While male analysts are indeed less accurate before 2002, they are more accurate than female analysts afterwards. Even though overconfidence most likely still hurts male analysts' performance, other factors are clearly more important in determining forecast accuracy. Two factors are put forward by the literature. First, strategic behavior can impact forecast accuracy because of management guidance (Hutton, 2005). Second, the mere environment of the analysts' profession can influence forecast accuracy because of its competitiveness (Gneezy et al., 2003). Using these two factors; two alternative theories are set forth: (i) a disproportionate effect of regulation on female analysts, (ii) females' worse performance in a competitive environment.

The basis of the first theory is the larger effect of strategic behavior on forecast accuracy compared to overconfidence. Our empirical results indicate that strategic behavior in the pre-2002 period is largest for female analysts. Therefore it is likely that the 2002 regulation had a disproportionately larger effect on the accuracy of female analysts. The loss of managerial guidance results in a larger decline in performance of female analysts compared to male analysts. Thus, the more severe impact of the regulation on female analysts results in their worse performance even though the male analysts are still overconfident.

The literature shows that male analysts outperform female analysts in a competitive environment. The second theory assumes a larger impact of competitiveness on accuracy compared to overconfidence. In absence of conflicts of interest, this results in a higher relative accuracy of male analysts, as observed in the post-2002 period. However, in the pre-2002 period, the larger strategic behavior of female analysts drives their higher relative performance.

5 Conclusion

Gender stereotyping suggests that men are more overconfident than women. This paper analyzes whether this is the case for earnings forecasts issued by male and female financial analysts. While multiple definitions of overconfidence exist, this paper defines an overconfident analyst as one who overestimates the precision of his private information signal. To identify overconfidence in earnings forecasts, we use a two-stage model that allows for the coexistence of an overconfidence bias (an unconscious bias), as well as a strategically-driven bias (a conscious bias) in the forecasting process.

Our analysis confirms the gender stereotype that males are more overconfident than females. We show that male analysts are overconfident, while female analysts are not. Male analysts believe too much in their own skills and information processing abilities, and therefore attach a weight to their personal information that is too high compared to the rational Bayesian weight. Female analysts are not prone to overweigh the precision of their private information, and rely on the rational Bayesian weights to process public and private information. This behavior is confirmed in the subsample analyses. During the entire sample period, male analysts are overconfident, while female analysts are not, but only in the pre-2002 period the gender difference in overconfidence is statistically significant.

References

[1] Barber B. M. and Odean T. 2001. Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment. *Quarterly Journal of Economics*, Vol.116(1). pp 261-292.

[2] Bosquet K., de Goeij P. and Smedts K. 2009. Coexistence and Dynamics of Overconfidence and Strategic Incentives. *Working Paper*.

[3] Chen Q. and Jiang W. 2006. Analysts' Weighting of Private and Public Information. *Review of Financial Studies*, Vol.19(1). pp 319-355.

[4] Chen S. and Matsumoto D. A. 2006. Favorable versus Unfavorable Recommendations: The Impact on Analyst Access to Management-Provided Information. *Journal of Accounting Research*, Vol.44(4). pp 657-689.

[5] Clement M. B. 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics*, Vol.27. pp 285-303.

[6] Daniel K., Hirshleifer D. and Subrahmanyam A. 1998. Investor Psychology and Security Market Under- and Overreactions. *Journal of Finance*, Vol.53(6). pp 1839-1885.

[7] Deaux K. and Farris E. 1977. Attributing Causes for One's Own Performance: The Effects of Sex, Norms, and Outcome. *Journal of Research in Personality*, Vol.11(1). pp 59-72.

[8] Estes R. and Hosseini J. 1988. The gender gap on wall street: An empirical analysis of confidence in investment decision making. *Journal of Psychology*, Vol.122(6). pp 577-590.

[9] Gervais S. and Odean T. 2001. Learning to be overconfident. *Review of Financial Studies*, Vol.14. pp 1–27.

[10] Gneezy U., Niederle M., and Rustichini A. 2003. Performance in Competitive Environments: Gender Differences. *Quarterly Journal of Economics*, Vol. 118(3). pp 1049-74.

[11] Green T. C., Jegadeesh N. and Tang Y. 2009. Gender and Job Performance: Evidence from Wall Street. *Financial Analyst Journal*, Vol.65(6).

[12] Hayes R. 1998. The impact of trading commission incentives on analysts' stock coverage decisions and earnings forecasts. *Journal of Accounting Research*, Vol.36. pp 299-320.

[13] Hilary G. and Menzly L. 2006. Does Past Success Lead Analysts to Become Overconfident. *Management Science*, Vol.52(4). pp 489-500.

[14] Hutton A. P. 2005. Determinants of Managerial Earnings Guidance Prior to Regulation Fair Disclosure and Bias in Analysts' Earnings Forecasts. *Contemporary Accounting Research*, Vol.22(4). pp 867-914.

[15] Kumar A. 2009. Self-Selection and the Forecasting Abilities of Female Equity Analysts. Forthcoming, *Journal of Accounting Research.*

[16] Ljungqvist A.; Marston F.; Starks L. T.; Wei K. D. and Yan H. 2007. Conflicts of interest in sell-side research and the moderating role of institutional investors. *Journal of Financial Economics*, Vol.85(2). pp 420-456.

[17] McNichols M. and O'Brien P. 1997. Self selection and analyst coverage. *Journal of Accounting Research*, Vol.35. pp 167- 199.

[18] Petersen M. A. 2008. Estimating standard errors in Finance Panel Data Sets: Comparing Approaches. *Forthcoming, Review of Financial Studies.*

[19] Prince M. 1993. Women, men, and money styles. *Journal of Economic Psychology*, Vol. 14(1). pp 175-182.

[20] Shefrin, H. 2008. *A Behavioral Approach to Asset Pricing*. 2nd edition, Academic Press/Elsevier, Burlington, MA, 604p.

[21] Thompson S. 2006. Simple Formulas for Standard Errors that Cluster by Both Firm and Time. Harvard *Working Paper.*

[22] White H. 1984. Asymptotic Theory for Econometricians. Academic Press.

[23] Zitzewitz E. 2001. Opinion-Producing Agents: Career Concerns and Exaggeration. *Working Paper.*

Appendix

A Two stage model

In the first stage, a financial analyst collects private and public information and processes the data to obtain an earnings forecast. In line with Gervais and Odean (2001) and Chen and Jiang (2006), a is defined as the actual earnings of a firm, which follows a diffuse zero-mean normal distribution; c is defined as a statistic for all *public* information about a:

$$\begin{aligned} c &= a + \varepsilon_c \\ & \quad \varepsilon_c \sim N\left(0, \frac{1}{p_c}\right) \end{aligned}$$

with p_c the precision of the public signal. Next, let x be the analyst's *private* information about a :

$$\begin{aligned} x &= a + \varepsilon_x \\ & \quad \varepsilon_x \sim N\left(0, \frac{1}{p_x}\right). \end{aligned}$$

with p_x the precision of the analyst's private signal and ε_x independent from ε_c. The analyst's best conditional forecast of the actual earnings (a) given his private information (x) and his public information (c), using Bayes' rule, is as follows:

$$E\left[a \mid x, c\right] = hx + (1-h)\,c$$

where $h \approxeq \frac{p_x}{p_x+p_c} \in [0,1]$ is the precision of the analyst's private signal relative to public information. In arriving at his judgement, the analyst is likely unconsciously influenced by a behavioral bias, such as overconfidence. This bias could lead him to use a personal weighting scheme, that deviates from the correct rational scheme:

$$F = kx + (1-k)\,c$$

with $k \in [0,1]$ the actual weight the analyst places on his private signal and F the first stage earnings forecast. This first stage earnings forecast is, however, not necessarily the one that is publicly issued by the analyst. Apart from the unconscious bias in the analyst's forecast, strategic incentives might induce the analyst to update his initial forecast F:

$$f = sF = s\left[kx + (1-k)c\right] = skx + (s - sk)c$$

with f the earnings forecast as observed by the public. The strategic incentives are modeled by introducing a multiplicative factor $s \geq 0$. As analysts are more likely to provide forecasts for stocks for which their true expectations are favorable and are reluctant to issue unfavorable investment information (McNichols and O'Brian, 1997), it is reasonable to assume that $s \geq 0$.

To obtain an expression for $\frac{k}{h}$ and s, we model the conditional expected forecast error:

$$\begin{aligned} E\left[f-a \mid x, c\right] &= skx + (s-sk)c - hx - (1-h)c \\ &= \frac{h}{k}c - \frac{h}{sk}f + (f-c) \\ &= \left(\frac{k-h}{k}\right)(f-c) + \left(\frac{hs-h}{sk}\right)f \\ &= \left(1-\frac{h}{k}\right)(f-c) + \frac{h}{k}\left(\frac{s-1}{s}\right)f \end{aligned}$$

Applying a first-order Taylor approximation to both coefficients in the above equation around $\frac{k}{h}=1$ and $s=1$, the reduced form estimation equation is transformed into:

$$\begin{aligned} E\left[FE \mid x, c\right] &= \left(\frac{k}{h}-1\right)(f-c) + (s-1)f \\ &= \beta(f-c) + \gamma f \end{aligned}$$

B Tables

Table 3 presents summary statistics on all variables used in the empirical equation (7), separately for male and female analysts. At first glance, there are no large differences between male and female analysts with respect to the forecast error, the deviation from consensus and the observed forecast. Concerning the control variables, male analysts release their forecast somewhat earlier in the forecasting period, have about nine months more of experience and follow slightly more stocks than female analysts.

The dummy approach to estimate the male and female forecasting process is summarized as:

$$\begin{aligned} FE_{ikt} &= \alpha_0 + \alpha_1 D_m + \alpha_2 D_f + \beta_m \left(D_m \times Dev_{ikt}\right) + \beta_f \left(D_f \times Dev_{ikt}\right) \\ &\quad + \gamma_m \left(D_m \times f_{ikt}\right) + \gamma_f \left(D_f \times f_{ikt}\right) + \delta_{m,i}\left(D_m \times X_{ikt}\right) \\ &\quad + \delta_{f,i}\left(D_f \times X_{ikt}\right) + \varepsilon_{ikt} \end{aligned} \tag{8}$$

Note that we use both a male dummy $D_m = 1$ in case of a male analyst and $D_m = 0$ otherwise, as well as a female dummy $D_f = 1$ in case of a female analyst and $D_f = 0$ otherwise. This facilitates the extraction of the overconfidence and strategic parameters we are interested, as well as the computation of the associated standard errors.

Table 3: **Summary Statistics**

This table presents the summary statistics of the forecast error (FE), the deviation from consensus (*Dev*), the (deflated) earnings forecast (f) and the control variables used in equation (7). *FE* is the difference between the earnings forecast and the actual, deflated by the share price. *Dev* is the difference between the earnings forecast and the consensus forecast, deflated by the share price. f is the analyst's earnings forecast, deflated by the share price. *Age* is the number of days between the issue of the analyst's earnings forecast and the reporting date of the actual earnings. *Size* is the logarithm of the market capitalization. *FirmExp* is the number of quarters an analyst has followed a certain stock. *TotExp* is the number of quarters the analyst is present in the data set. For both ability variables data starting from 1992 is used to prevent all analysts from starting with the same experience in 1996. *FirmCompl* is the number of companies an analyst follows during a quarter. *IndCompl* is the number of sectors an analyst follows during a quarter. I/B/E/S identifies 11 sectors using a proprietary classification scheme. Summary statistics are presented separately for male and female analysts.

	Male analysts				Female analysts			
	Mean	Stdev.	Min.	Max.	Mean	Stdev.	Min.	Max.
FE	1.1E-6	0.006	-0.254	0.682	7.2E-5	0.006	-0.113	0.266
Dev	0.005	0.009	-0.112	0.532	0.005	0.001	-0.076	0.267
f	0.014	0.009	0.000	0.532	0.015	0.001	0.000	0.314
Age	74.352	25.765	1.000	143.000	71.798	27.478	1.000	139.000
$Size$	7.909	1.324	-5.319	12.344	8.061	1.333	3.391	12.146
$FirmExp$	11.815	10.498	1.000	59.000	11.280	9.998	1.000	59.000
$TotExp$	21.614	13.344	1.000	59.000	19.293	12.486	1.000	59.000
$FirmCompl$	9.066	4.959	1.000	29.000	8.443	5.027	1.000	29.000
$IndCompl$	1.741	0.944	1.000	9.000	1.561	0.851	1.000	6.000

Table 4: **Gender Differences: Male and Female Dummy**

This table presents estimation results for the reduced form estimation equation (8). Firm fixed effects with clustered standard errors by industry is used to estimate the model. *,**,*** indicate significance at the 10%, 5% and 1% level, respectively. Results for control variables are not reported to facilitate readability. The sample runs from Q1-1996 until Q4-2006.

	Coeff. $(1 \times e^{-3})$	t-value
D_m	0.153	4.59***
D_f	-0.126	-1.05
$D_m \times Dev_{ikt}$ $(\widehat{\beta}_m)$	26.359	4.24***
$D_f \times Dev_{ikt}$ $(\widehat{\beta}_f)$	7.560	0.65
$D_m \times f_{ikt}$ $(\widehat{\gamma}_m)$	24.393	1.82*
$D_f \times f_{ikt}$ $(\widehat{\gamma}_f)$	35.705	2.12**
$\delta_{m,i}$ $(D_m \times X_{ikt})$	*not reported*	
$\delta_{f,i}$ $(D_f \times X_{ikt})$	*not reported*	
No. Obs.	302,475	

Table 5: **Gender Differences: Male and Female Dummy**

This table presents estimation results for the reduced form estimation equation (8). Firm fixed effects with clustered standard errors by industry is used to estimate the model. *,**,*** indicate significance at the 10%, 5% and 1% level, respectively. Coefficient estimates are multiplied by 1000 and results for control variables are left out to facilitate readability. The subsamples run from Q1-1996 until Q1-2002 (Pre-2002) and Q2-2002 until Q4-2006 (Post-2002).

	Pre-2002		Post-2002	
	Coeff.	t-value	Coeff.	t-value
D_m	0.183	2.19**	-0.145	-1.47
D_f	-0.031	-0.15	-0.345	-1.15
$D_m \times Dev_{ikt}$ ($\widehat{\beta}_m$)	33.999	3.05***	24.635	2.92***
$D_f \times Dev_{ikt}$ ($\widehat{\beta}_f$)	11.675	0.13	13.123	0.76
$D_m \times f_{ikt}$ ($\widehat{\gamma}_m$)	42.998	2.30**	3.280	0.17
$D_f \times f_{ikt}$ ($\widehat{\gamma}_f$)	59.608	3.40***	12.277	0.52
$\delta_{m,i}$ ($D_m \times X_{ikt}$)	*not reported*		*not reported*	
$\delta_{f,i}$ ($D_f \times X_{ikt}$)	*not reported*		*not reported*	
No. Obs.	153,239		156,734	