Overconfidence, Competence and Trading Activity

Liu-Yi Yeoh and Andrew Wood

March 2011

ABSTRACT

Despite the theoretical and intuitive appeal of a link between overconfidence and trading volume, supporting evidence is provided by proxies for overconfidence such as gender and past returns while interval based measures of confidence are uncorrelated with trading activity. We design an experiment that differs from previous studies by giving participants freedom to trade using real stock prices over an extended period of eight weeks. Overconfident participants undertook smaller but more frequent trades. Participants who believed they would finish high in the game undertook larger but less frequent trades. Men log in more often but trade less than women, both in absolute terms and relative to the number of log-ins, but when they do trade the size of those trades are significantly larger, both in statistical and economic terms.

Corresponding author: Andrew Wood, Essex Business School, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK. Tel.: +44 1206 872402. Fax.: +44 1206 873429. Email: wooda@essex.ac.uk.

1. Introduction

Overconfidence has been described as one of the most robust findings from the Judgment and Decision Making literature and has been applied to a number of puzzles within finance, most prominently to explaining high levels of trading activity. Despite the intuitive appeal linking overconfidence to trading activity, there is limited evidence supporting the relationship and controversies surround the measurement of overconfidence and the precise interpretation of those measures. In this paper we contribute to this debate by using evidence from an experiment designed to distinguish between different types of overconfidence and their effect on trading volume and to throw light on some of the methodological and empirical controversies.

In an influential paper, Odean (1998) has argued that overconfident investors overestimate the accuracy of their own evaluations resulting in an under-estimation of risk and increasing the differences of opinions between traders, thereby resulting in higher trading volume. In a recent contribution to this debate, Graham et al. (2009) have suggested a rival psychological explanation for trading activity: investors who believe they have a high level of relevant knowledge and skill are more inclined to take risks causing them to trade more frequently. They argue that not only is competence a rival explanation for high levels of trading, but existing evidence for overconfidence affecting trading may in fact be due to measures of overconfidence picking up a competence effect.

A difficulty in testing these models is that, when using market data, it is often necessary to use quite crude proxies for overconfidence. For example, Barber and Odean (2001) use gender as a binary indicator of overconfidence, with men presumed to be more overconfident than women. Ekholm (2006) and Ekholm and Pasternack (2007) assume overconfidence decreases with investor size and Statman et al. (2006) assume that overconfidence increases and falls following rises and declines in the market. The alternative approach to testing hypotheses relating to the effect of overconfidence on trading behaviour is to adopt an experimental framework which enables the use of more refined measures of overconfidence at the expense of a loss of realism in the investment environment. Contrasting results from these two approaches provides something of a paradox. Those studies that use proxies for overconfidence such as gender and past returns alongside real trading data are supportive of the excess trading hypothesis. By contrast those studies utilising more direct measures of overconfidence from survey data or within an experimental framework produce at best limited support for the excess trading hypothesis.

We are left with the question of whether the evidence from proxies overstates the case for overconfidence causing excess trading because those proxies are picking up other causal factors, or whether surveys and experiments are failing to either capture an accurate measure of overconfidence or to satisfactorily recreate a trading environment. In the current study we adopt an experimental framework that is designed to overcome some of the shortcomings of previous comparable studies. In particular, whereas previous experimental studies use artificial data within extremely time constrained games set within a laboratory environment, our experiment takes the form of an eight week trading competition using London Stock Exchange share prices. In this way, players are free to trade when and where they like, with the only constraint that they need to be able to go on-line. We believe this freedom is an important and distinctive feature of our framework since it enables different levels of overconfidence to fully come into play in influencing individual trading activity. In response to well known criticisms relating to the number and nature of the questions used to construct calibration-based measures of overconfidence (Lichtenstein et al. 1982), we use confidence intervals based on 40 finance related questions, at least twice the number used in comparable studies.

Broadly speaking our findings provides support for a trading-confidence relationship. Participants who scored highly for our miscalibration measure of overconfidence traded more frequently than those who were underconfident, while those who believed they would perform well in the game undertook larger trades on average. However, these same two measures were associated with smaller transactions and less frequent trades respectively. Our strongest result in terms of both statistical and economic significance involved gender with men trading substantially larger values than women with this finding being robust to each of our measures of confidence. Against this, we find that although men appear to be more engaged with the game in that they logged on more frequently, the did not trade more frequently than women. This pattern of results suggests there is more to the role of gender than a simple overconfidence story.

The paper proceeds as follows. Section 2 provides the related literature review, Section 3 describes our experimental design. Section 4 presents our results and Section 5 concludes.

2. Hypotheses

The primary hypothesis relating overconfidence to trading can be succinctly summarised as:

H1: trading volume increases with confidence, with overconfident traders exhibiting higher volume than less confident traders.

In order to test H1 we need to define trading volume. The conventional approach in the overconfidence literature is that overconfident traders have a higher turnover of shares because investors over-estimate the precision of their own signal and undervalue the signals received by others. However, it may be that overconfident traders undertake larger transactions but once the trade is made the stocks are held for longer. We therefore not only test the effect of overconfidence on the volume of trades, but also for the effect on the number of transactions and the size of individual transactions. This leads to a refinement of H1 as follows:

H1a: the number of transactions increases with confidence

H1b: the average size of transactions increases with confidence.

If trading volume increases with confidence and investment performance increases sufficiently to outweigh any increase in transaction costs, it suggests that confidence is justified. For a positive relationship between confidence and trading volume to be attributed specifically to overconfidence as opposed to a well founded belief in one's own ability, we need to confirm that the higher confidence is not justified. There are two ways of doing this; we can test that a positive relationship between confidence and performance does not exist and, more importantly, that there is not a positive relationship between trading volume and performance.

H2: investment performance does not increase with confidence

H3: investment performance does not increase with trading volume

Having previously shown that individual traders' turnover and profits are negatively related, Barber and Odean (2001) test their conjecture that this is due to overconfidence. They premise their test on the assumption that men are more overconfident than women, citing a Gallup poll conducted for PaineWebber that indicates men expect higher returns than women. Using brokerage data, Barber and Odean (2001) find that men traded more than women such that the portfolio turnover of male participants was on average 45% higher. Furthermore, while both samples of men and women achieved negative net returns, the gross returns earned by men were slightly higher than those achieved by women. But after taking account of the higher transaction costs, men's net returns were lower than women's, albeit with statistically insignificant differences. This much cited paper leads to the following hypotheses:

H4: men are more overconfident than women

H5: Men expect higher returns than women

H6: men trade more than women

A problem with the evidence that does exist of a relationship between measures of confidence and trading activity is that the precise causal mechanism can be questioned. In a recent study, Graham et al. (2009) suggest that investors who believe themselves to be better than average also believe themselves to be competent. Citing Heath and Tversky (1991), they note that when people feel more competent, either because they believe themselves to be particularly skilful or knowledgeable, they are more likely to take gambles. It follows that investors who reveal their self to be confident are also likely to feel competent, thereby giving rise the following hypothesis that is distinct from the central overconfidence-trading volume hypotheses:

H7: high perceived competence leads to more trading

3. The Experiment Design

Participants for a financial trading competition were recruited from students at the University of Essex who responded to poster and e-mail advertisements. A total of 260 participants signed up for the competition, with the final sample reduced to 146 once students who failed to meet the requirement of logging on at least once every two weeks during the eight weeks of the competition.[1] As a prerequisite to take part in the competition, participants were first required to fill out a questionnaire to elicit measures of overconfidence and competence. Questionnaires were self-paced, and required approximately 45 minutes to complete.

i) the trading game

Previous experimental studies that have tested for an overconfidence effect on trading volume have made used of trading data produced within a highly structured laboratory environment. We designed our experiment to enable participants to be as free as possible, both in terms of what they invest in and how frequently they choose

[1] During the eight weeks, participants are required to log in once every two weeks. This does not mean that they have to transact on each log in, it merely allows us to ensure that participants are still 'staying in' the challenge. Participants who administer a buy and hold strategy will have informed us at the beginning of the experiment and are therefore exempt from the requirement to login fortnightly. 46 of the participants chose a buy and hold strategy.

to trade. To achieve this we ran a trading competition lasting a full eight weeks. Each participant was provided with an online trading account with a nominal sum of £10,000 access to approximately 2,750 stocks listed on the London Stock Exchange.

Our experiment benefits from the online trading platform, *Finesse*, which resembles current online trading facilities provided by stockbrokerage firms (e.g. etrade, iwebsharedealing, etc).[2] The interface and functions mimic an online trading account such that participants are able to closely monitor the performance of their portfolio, search for price, P/E and EPS information on all stocks, as well as enabling the buying and selling of stocks with just a few clicks of the mouse. We believe the realism of the trading platform, the eight week span of the game and the absence of intervention from the organisers combine to create an environment which enables different levels of overconfidence to fully come into play in influencing trading decisions.

The effectiveness of financial incentives used in experimental markets has been inconclusive with a complex relationship between the level of remuneration and performance which depends on the nature and duration of the task (Camerer and Hogarth 1999). In their survey of 74 experiments, Camerer and Hogarth find that for many economic type experiments, such as trading in markets, bargaining in games and choosing among risky gambles, the overwhelming finding is that increased incentives do not change average behaviour substantively, although the variance of responses often decreases. In comparable experimental studies examining the link between psychometric variables and trading activity, Deaves et al (2009) did not use incentives while Biais et al (2005) rewarded students in terms of grades that would contribute to the final grade of the course taught on stock markets.

We felt that given the voluntary nature of our competition along with the long period over which it takes place it was important to incentivise participants. In addition to financial prizes for the best three performances according to final portfolio value, there were a further five prizes which would be determined by a lottery draw. The purpose of the lottery draw was to minimise the winner-takes-all effect.[3] To this end, the number of lottery tickets allocated to each participant was determined by the

[2] Finesse is an on-line trading platform created by the universities of Dundee, Glasgow Caledonian, St Andrews and Strathclyde to support the teaching and learning of finance. Using data that is updated from the London Stock Exchange with a fifteen minute lag, Finesse enables students to construct and manage their own investment portfolios.

[3] The total value of prizes was £700

rank of portfolio value at the end of the eight weeks, such that the higher the portfolio rank, the greater the reward in terms of lottery tickets. In this way, the additional lottery prize draw encouraged participants to remain involved in the challenge even when the portfolio had performed poorly, and to discourage participants from taking unnecessarily risky bets to achieve the top three portfolio values. The incentive to stay in the game come what may was further supported by participants not knowing where they stood vis-à-vis rival players.

ii) The data

When used in theoretical models of financial market behaviour, overconfidence is modelled as the overestimation of the precision of private information (Glaser and Weber 2007). Conceptual and measurement considerations outlined in a large literature from the field of cognitive psychology suggest a greater complexity. In a recent review of this literature, Moore and Healy (2008) identify three distinct types of overconfidence. The largest literature relates to what they describe as overestimation; the phenomenon where individuals overestimate their performance or ability. The second type is where individuals have excessive certainty regarding the accuracy of their beliefs as signified by unduly narrow confidence intervals around their answers to questions and is labelled overprecision. The third definition is where people believe themselves to be better than others. Moore and Healy describe this as overplacement, although it is more commonly referred to as the better-than-average effect.

With its easy application to risk, Moore and Healy's notion of overprecision best fits with the financial economist's view of overconfidence. Despite this, many empirical studies are unable to provide measures of overprecision and must resort to proxies. Even those studies that utilise information from questionnaires that provide a measure of overprecision, also invariably include other measures such as better-than-average which, despite its inclusion in Moore and Healy's taxonomy, is potentially problematic since it does not distinguish between those who consider themselves better than average because they are overconfident from those who are well calibrated and are genuinely above average.[4]

[4] This relates to a further complication that measures of confidence may proxy for a person's competence which, as Heath and Tversky (1991) demonstrate, leads to greater risk taking and therefore more trading (Graham et al. 2009).

Notwithstanding these concerns, we use information obtained from the questionnaire completed by all participants to construct three psychometric measurements relating to confidence and competence; a miscalibration based measure of confidence (CONF), the better than average effect (BTA) and a measure of competence (COMP).

(i) Confidence Measure (CONF)

Participants were presented with forty two-choice finance related questions and are required to state their level of confidence to which they believe they have provided the correct answer in the form of a probability that must lie between 0.5 and 1.[5] From the responses we compute a measure of confidence, *CONF*, following the method used by Klayman *et al.* (1999). For each participant we compute the proportion of correct answers, P, and the average confidence level reported, Q. Confidence is measured as:

$$CONF = (P - Q)$$

If a participant had correctly answered a large proportion of questions correctly, for example P = 80%, but indicated an average confidence level that is lower, say Q = 70%, the resulting CONF score is negative and the participant is described as underconfident. But if she had indicated a confidence level that is higher, say Q = 90%, she would have a positive CONF measure and said to be overconfident. The average value of CONF across all individuals is 0.0497, with 94 respondents having a positive value and 52 a negative value indicating over and under-confidence respectively. This profile is similar to that of the study by Klayman *et al* (1999) who reported an overall overconfidence score of 0.0458.

(ii) Better than Average Effect (BTA)

When asked about their skills or personal attributes people generally think that they are above average. One of the most cited examples states that when asked about driving safety, 82% of a group of students rank themselves among the top 30% (Svenson 1981). The tendency for individuals to locate themselves within the top half of a performance ranking for a wide variety of tasks has been labelled the better-than-

[5] Forty compares well with comparable studies that construct miscalibration measures of overconfidence using between five and twenty questions (Glaser and Weber 1007; Biais et al. 2005; Deaves et al. 2009), but falls short of the “hundreds” of questions recommended by Lichtenstein et al. (1982). In common with these other studies, we were constrained by the problem of questionnaire fatigue.

average (BTA) effect. Although it is logically impossible for the majority of a population to correctly place themselves amongst the higher deciles, it does not necessarily follow that BTA provides a measure of an individual's overconfidence; if 20% of a population place themselves in the top 10% by ability, some of those 20% may be correct in their personal assessment.

Nevertheless, BTA is related to confidence and is widely used in comparable studies. We therefore construct a measure of BTA by asking participants to predict which decile they are likely to be situated within a ranking of performance once the competition is finished, decile 1 being the highest top 10%, decile 10 being the lowest/bottom 10%. The data shows 72% of participants placed themselves in the top 5 deciles, implying that they assess their skills and abilities as above average.

(iii)Competence (COMP)

The concepts of confidence and competence are closely related. Indeed, our measure of CONF is constructed from two measures of competence, the proportion of correct answers, P, being an objective measure of competence, while the average reported confidence level underlying individual answers provides a subjective measure of competence. By disaggregating the CONF measurement, we define COMP as the proportion, P, of correct answers.

(iv)Review of direct measures and additional proxies

We believe that CONF is our purest measure of confidence that distinguishes between over and underconfidence and is not contaminated by alternative interpretations such as competence. Our other measures are more problematic. Although BTA is commonly used in empirical studies of overconfidence, it does not distinguish between justified and non-justified confidence in ability. However, in the context of investing in financial stocks it could be argued that a belief that you are better than the average investor is akin to believing you will beat the market which is tantamount to revealing yourself to be overconfident. Finally, COMP is a direct measure of subjective competence and is distinct from confidence.

We also use three binary indicators of confidence and/or competence. Following Barber and Odean (2001), we distinguish between male and female participants, on the basis of evidence that men tend to be more overconfident than women. We further hypothesise that the 75 participating students registered for a degree in finance may have a greater belief in their investment ability than those who are studying for an unrelated degree. And we describe those participants who have either traded on their own account or have worked in the financial sector as having investment experience.

(v) Trading Volume Variables

Most studies that explore the link between confidence or competence and trading activity define trading activity in terms of the size of trades, either relative to the initial size of the individual's portfolio (turnover) or in absolute value terms (Statman et al. 2006; Barber and Odean 2001; Glaser and Weber 2007). However, Glaser and Weber (2007) also test for a relationship with trading frequency while Graham et al. (2009) tests solely for an effect on trading frequency. In the current study we utilise three measures of trading activity that cover both the value and the trading frequency approaches to measuring activity. These are: the average value of individual trades; the total value of all trades undertaken during the eight month competition; and the total number of transactions carried out.

4. Results and Interpretation

(i) Descriptive Statistics

Our final sample of 146 participants were students from the University of Essex and included 41 female and 105 male participants, coming from a total of 35 countries, with approximately half of our sample British (46) or Chinese (30). The ages of participants ranged from 19 to 42 years old, with the majority being between the ages of 19 and 25. 108 were undergraduate students and 38 postgraduate. Although the competition was open to all students at the University, the largest proportion were majoring in business related subjects, namely, finance (66), economics (28) and business and management (24). As reported in Table 1, other subjects with a significant number of students included law (11) and computer science and engineering (8).

[Table 1 around here]

When asked if they had any investment experience, 121 had none and 12 had less than 6 months. Additionally, 31 participants had worked in a finance related industry, including investment management, auditing, banking, and corporate finance. When asked if they had close friends and family investing in the stock markets, 56.16%

answered yes, 43.84% had none. Hence, the majority of our participants fit the profile of a first time investor with little investment experience but some exposure to the stock markets via friends and families, and an interest in investing in the stock markets, indicated by their voluntary engagement in an eight week stock trading challenge.

Participants were very bullish about their prospects for achieving high returns, with 63% expecting to achieve returns in excess of 15% during the eight week period.[6] In practice only 3% of participants managed returns in excess of 15%. When participants were asked to choose which decile they expected to end up in at the end of 8 weeks according to portfolio returns in comparison with the other participants, 72% of participants said they would end up in the top 5 deciles (better than average). In fact, 57% of participants fared worse than their expected positions, only 29% fared better and 14% realised their expected decile.

Correlation coefficients between the psychometric and volume measures are reported in table 2. Of the three correlations between our direct measures of confidence and competence, that between CONF and COMP is significant at the 1% confidence level and the correlation between BTA and CONF is significant at the 10% level.[7] Participants demonstrating higher competence in the questionnaire also scored very highly on the CONF measure although not, perhaps surprising, for BTA.

[Table 2: around here]

(ii) Results for univariate sorts

Table 3 presents significance tests for the univariate sorts based on the Wilcoxon signed rank test for equality of medians. Participants are categorised according to whether they have above or below median scores for BTA and COMP, positive or negative values of CONF, whether or not they are registered on a degree in finance, and finally whether they have experience either as an investor or in the finance

[6] Although answers to this question will in part reveal the participant's knowledge of the market, in particular with regard with what constitutes a reasonable return for an eight week investment horizon, it is reasonable to surmise that higher predictions are indicators of greater confidence, and indeed over-confidence given the generally accepted evidence that investors do not make abnormal profits. At first glance, the reported predictions are striking given that an expected outcome of 10% may be more realistic for a full year's investment horizon rather than eight weeks. However, these predictions are not so outlandish when one considers media coverage of high performing stocks that may well influence individual predictions more than realistic forecasts of aggregate market performance.

[7] Due to the nature of the BTA measure where Decile 1 indicates higher BTA and Decile 10 indicates lower BTA, a negative correlation in fact indicates that SCOMP and the actual BTA is positively correlated.

industry. The reported tests are for the equality of median values for portfolio performance and trading volume, as measured by total transactions, the average value of transactions, and the number of logins according to these categorisations.

[Table 3: around here]

Both CONF and BTA play some role in influencing trading activity that accord with theory. Participants with positive values of CONF undertook significantly more transactions than those with negative values with a *p*-value of 0.01, while those with above median BTA[8] carried out transactions of marginally larger size than those with below median BTA with a *p*-value of 0.10.[9] There is no evidence of a relationship between COMP and trading activity.

Participants who had investment experience, either from investing on their own account or from having worked in the finance industry, undertook less trades and logged on less frequently, although the difference is insignificant. However, those taking finance degrees logged on more frequently and undertook significantly more transactions than those who were not studying finance, but the average size of trades carried out by finance students was no larger than that of non-finance students. While there may be a confidence story underlying this finding, it is also consistent with finance students being more involved in the contest due to their interest and knowledge of the subject rather.

Turning to the results for performance reported in the final column, these are consistent across all our measures and indicators of confidence with no evidence of a significant difference in portfolio return for high and low measures of BTA and COMP or positive and negative CONF or whether or not the participant was registered for a finance degree. Although this is not a surprising result, it is nevertheless important since it allows us to discount differences in trading activity reflecting differential investing abilities.

Studies in the field of psychology have found that men tend to be more overconfident than women. Our results reported in Table 4 are supportive of this with significantly higher median values of CONF and BTA for men relative to women. Although there is no significant difference between the number of transactions made by men relative to women, there is a very striking difference between the values of

[8] The cut-off point for BTA was 3 with 62 participants predicting that they would finish in the third decile or higher.

[9] In unreported hypothesis tests based on the equality of means, the difference between average transaction values according to the BTA sort was significant with a p-value of 0.087.

transactions with the median value of those undertaken by men more than twice that of those made by women, a difference that is significant at the 1% level. Finally, although men expected to outperform women there are no significant gender differences in actual portfolio returns.

[Table 4: around here]

(iii) Empirical Results: Cross Sectional Regressions

In table 5 we present regressions for three measures of trading activity: number of trades, average transaction value, and total value of trades. Since our measures of confidence and competence are highly correlated we estimate separate regressions for each measure of trading activity with BTA, CONF and COMP as regressors in the results reported in columns (1), (2) and (3) respectively. For each regression we include three binary control indicators GENDER, FINANCE and INVEXP which take the value one respectively for participants who are male, registered for a degree in finance and have prior investment experience.

[Table 5: around here]

Our favoured measure of confidence, CONF, does not play a significant role in any of the regressions. However, for the regression explaining total number of trades reported in column (2) of Panel A, the *p*-value for CONF is 0.111. The results reported in Table 5 demonstrated that a sort of total number of trades based on a binary version of this variable produced a significant difference in number of trades made by overconfident traders relative to underconfident traders. In view of this and in acknowledgement that CONF is a noisy measure of confidence, in column (2’) of each Panel we report results replacing CONF with a dummy variable, CONFDUM, taking the value 1 for values of CONF>0 and zero for negative values of CONF. The results are intriguing with a significant positive coefficient in column (2’) of Panel A but a negative coefficient with a *p*-value of 0.058 in the corresponding column in Panel B. This suggests overconfident participants carry out more trades but they tend to be smaller than those carried out by underconfident participants with the net result that there is no significant difference in the total value of trades undertaken by over and underconfident participants.

Both BTA and COMP are significant at the 5% level in their respective regressions explaining the average transaction value but are not significant in the regressions for total number of trades or the total value of trades. This indicates, for example, that for every higher decile a participant places their self relative to the average, the average size of their individual transactions is approximately 11% higher and a one standard deviation increase in COMP causes a 26% increase in the average transaction size.

Taking the results from the regressions and univariate sorts as a whole it is notable that the direct measures of confidence, CONFDUM and BTA, play different roles in influencing trading volume with CONFDUM associated with greater trading activity (the number of transactions) and BTA influencing the magnitude of individual trades, but surprisingly there is no suggestion of BTA influencing trading activity while CONFDUM has an unexpected sign for explaining the size of trades. It is clear that these measures relate to different aspects of confidence, with BTA perhaps relating as much to optimism than confidence.

Each of the remaining binary explanatory variables play a role in explaining trading activity. The variable that has the least explanatory power is INVEXP which is significant at the 10% level for three out of four regressions reported in Panel A but is not significant in Panels B and C. The coefficient in Panel A is negative, implying participants with investment experience undertake less trades than those without experience and is consistent with the corresponding univariate sort reported in Table 3. This result is neither consistent with Barber and Odean’s (1998) confidence effect nor with Heath and Tversky’s (1991) competence effect. We could instead speculate on this being a consequence of experienced investors being more likely to favour a buy-and-hold strategy or maybe being less enthusiastic due to the game not being able to compete with the ‘real thing’. However, given the low level of significance combined with the broad definition of the meaning of ‘experienced’ it is perhaps best not to place too much importance on this variable.

The results for FINANCE and GENDER are more interesting. The coefficients for FINANCE are positive and significant at the 5% level for all but two regressions, for which they are significant at the 10% level. FINANCE plays its strongest role in the results for total value of trades, reported in Panel C, with coefficients that are significant at the 1% level. The positive coefficients indicate that participants who were studying finance undertook more trades of larger value resulting in a total value traded between 78% and 88% higher than those who were not studying finance. GENDER has an even larger impact, with highly significant coefficients in both Panels B and C demonstrating that both the average transaction value and the total value of all trades undertaken by men are more than twice that of women. But the insignificant coefficients for GENDER reported in Panel A indicate that men are not inclined to carry out more trades than women implying that the gender effect is restricted to monetary magnitudes.

Before we interpret the results for our dichotomous proxy variables as supportive of the overconfidence hypothesis we should consider potential alternative interpretations. The finance dummy may be capturing an interest or knowledge effect whereby participants in the contest who were also registered for a finance degree may be more likely to devote more time to the competition simply because they are more interested in the activity of investing. This interpretation is consistent with the coefficient for finance being significant for the number of transactions regressions but not for regressions explaining the value of transactions.

To explore these alternative explanations we run additional regressions with the number of logins and the number of trades per login as the dependent variables. If the significant coefficient for the finance dummy reflects greater interest and/or knowledge of finance students, we would expect to see evidence of finance students participating more actively with the game without necessarily making more trades. The results reported in table 6 show that finance students do indeed log in more frequently but do not make more trades per login, suggesting the significance of finance reported in Table 5 does reflect greater interest as opposed to confidence in ability.

[Table 6: around here]

A similar pattern of results is found for men, with a positive coefficient for logins but negative for trades per log in. However, given that men are associated with high value trades rather than large numbers of trades, the interpretation of the login results differs to that for finance. Men logon more frequently but are less likely to trade when they log on, and when they do trade they make substantially larger trades than those made by women.[10] Moreover, these results are robust to the inclusion of our direct measures of confidence and are both more significant and have higher economic impact than BTA, CONF and CONFDUM. This leads us to conclude that an alternative explanation for the trading behaviour of men is required. One such

[10] A closer inspection of the underlying data reveals that 117 of the 146 participants had a ratio of trades/logins less than one, implying they logged on more frequently than they traded. Of these, 74% were men while 62% of those with trade/login ratios greater than one were men.

explanation is that it reflects greater assertiveness, even aggressiveness, on the part of the male players. Recent studies have found a negative relationship between testosterone and risk aversion (Sapienza et al. 2009; Apicella et al. 2008) and performance on the trading floor (Coates and Herbert 2008; Coates et al. 2010). However, Sapienza et al. (2009) report a gender effect that is independent of testosterone which suggests gender related socio-cultural factors may as important as hormonal differences in explaining gender difference in trading activity.

The results in Panel B of Table 6 provide an alternative way of looking at the effect of confidence on trading activity by focusing on the times when the participants are actually engaged with the game. The highly significant, positive relationship between both CONF and CONFDUM with trades per login suggesting that more confident participants are more decisive when trading in terms of the decision to trade although this is not reflected in the magnitude of the trade.

(iv) Portfolio Performance and Overconfidence

To determine whether high trading volume or high levels confidence are associated with poorer portfolio performance, returns are regressed on measures of overconfidence and trading volume along with the gender and finance dummy variables. The results reported in Table 7 indicate that participants who scored highly on the finance questionnaire did perform significantly better. However, performance is not positively related to either of the overconfidence measures or trading volume. Indeed, the coefficients for CONF and average transaction values both indicate a significant negative relationship with portfolio performance.

[Table 7: around here]

5. Conclusion

A review of the limited empirical literature testing for a relationship between overconfidence and trading activity shows mixed support for the theory with most corroborating evidence coming from studies that use rather crude proxies for overconfidence. Studies that utilise purpose made, direct measures of confidence and overconfidence provide at best patchy support. The poor performance of direct measures could be due to problems with the experimental design or the measures themselves may lack accuracy. The success of the proxies in explaining trading activity may be due to them being better indicators of what is a complex cognitive state, or alternatively it is a consequence of them capturing other phenomenon that influence trading.

We take a number of steps to reduce problems that are often associated with the experimental framework and the construction of the direct measures of confidence, including using real financial data; giving participants maximum freedom to choose when and how often they trade; and deriving our primary measure of confidence from forty finance related questions. Our results provide some support for the central hypothesis linking confidence to trading activity. In support of the theory, participants who were identified as overconfident by our favoured miscalibration measure traded more frequently than those who were underconfident, but contrary to the theory, the size of their trades were significantly smaller. By contrast, the average size of trades was significantly high for participants who believed they were better than average measure, but they traded less frequently.

The specific hypothesis proposed by Graham et al. (2009) that trading frequency increases with competence is not supported by our evidence. However, although our measure of competence is negatively related to the number of transactions, albeit with an insignificant coefficient, it is positively related to the average value of transactions undertaken.

Perhaps our most striking results are associated with the role of gender. Men log in more often but trade less than women, both in absolute terms and relative to the number of log-ins, but when they do trade the size of those trades are significantly larger, both in statistical and economic terms. This pattern is robust to the inclusion of each of our direct measures of confidence and competence which, along with the evidence that gender-related hormonal differences influence risk aversion, leads us to conclude that overconfidence can at best be only part of the explanation for a relationship between gender and trading volume.

Bibliography

Apicelli, C.L., Dreber, A., Campbell, B., Gray, P.B., Hoffman, M. and Little, A.C. (2008) “Testosterone and financial risk preferences”, Evolution and Human Behavior, 20, 384-390.

Barber, B. and Odean, T. (1999), “The Courage of Misguided Convictions”, Financial Analysts Journal, Special Issues on Behavioural Finance, 41-55

Barber, B. and Odean, T. (2000), “Trading is Hazardouz to Your Wealth: The Common Stock Investment Performance of Individual Investors”, Journal of Finance, Vol. 55, 773-806.

Barber, B. and Odean, T. (2001), “Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment”, Quarterly Journal of Economics, Vol. 116, 261-92.

Brier, G. W. (1950), “Verification of Forecasts Expressed in Terms of Probability”, Monthly Weather Review, Vol. 78, 1-3.

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

Camerer, C.F. and Hogarth, R. M. (1999), “The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework”, Journal of Risk and Uncertainty, Vol. 19, No. 1-3, 7-42.

Coates, J.M., and M. Gurnell (2008) “Endogenous steroids and financial risk taking on a London trading floor”, Proceedings of the National Academy of Sciences, 104, 6167-6172.

Coates, J.M., M. Gurnell, Z Sarnyai (2010) ‘From molecule to market: steroid hormones and financial risk taking’, Philosophical Transactions of the Royal Society B, 365, 331-43

Deaves, R., Lüders, E. and Luo, G.Y. (2009), “An Experimental Test of the Impact of Overconfidence and Gender on Trading Activity”, Review of Finance, 13, 555-575.

Daniel, K., Hirshleifer, D. and Subrahmanyam, A. (2001), “Overconfidence, Arbitrage and Equilibrium Asset Pricing”, Journal of Finance, Vol. 56 (3), 921-65.

De Bondt, W. and Thaler, R. H. (1995), “Financial Decision-Making in Markets and Firms: a Behavioral Perspective”, in Jarrow, R.A., Maksimovic, V. and Ziemba, W.T. (eds), “Finance, Handbooks in Operations Research and Management Science”, Vol. 9, 385-410, Amsterdam, North Holand.

Ekholm, A. (2006) “How Do Different Types of Investors React to New Earnings Information?”, Journal of Business Finance and Accounting, 33, 127-144

Ekholm, A. and D. Pasternack (2007) “Overconfidence and Investor Size”, European Financial Management, 14. 82-98.

Heath, C., and Tversky, A. (1991) “Preference and belief: Ambiguity and competence in choice under uncertainty”, Journal of Risk and Uncertainty, 4, 5-28.

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

Graham, J. R., Huang, H., and Harvey, C. (2009), Investor Competence, Trading Frequency, and Home Bias, Management Science 55, 1094-1106.

Griffin, D. and L. Brenner (2004) “Perspectives on probability judgment calibration”, in D.J. Koehler and N. Harvey (eds.) Blackwell Handbook of Judgment and Decision Making, Blackwell

Heath, C. and A. Tversky (1991) ‘Preference and belief: ambiguity and competence in choice and uncertainty, Journal of Risk and Uncertainty, 4, 5-28.

Klayman, J., Soll, J.B. and González-Vallejo, C. and Barlas, S. (1999), “Overconfidence: It Depends on How, What, and Whom You Ask”, Organizational Behavior and Human Decision Processes, Vol. 79, No. 3, 216-247.

Lichtenstein, S., Fischhoff, B., and Phillips, L.D. (1982), “Calibration of Probabilities: The State of the Art to 1980”, in Kahneman, D., Slovic, A., and Tversky, A. (eds.), “Judgement Under Uncertainty: Heuristics and Biases”, Cambridge University Press, Cambridge.

McClelland, A. and Bolger, F. (1994), “The Calibration of Subjective Probabilities: Theories and Models 1980-93”, in Wright, G., and Ayton, P. (eds.), Subjective Probability, Wiley, Chichester.

Moore, D.A. and P.J. Healy (2008) “The trouble with overconfidence”, Psychological Review, 115(2), 502-517.

Russo, J.E. and P.J.H. Schoemaker, Managing overconfidence, *Sloan Management Review* 33 (1992), pp. 7–17.

Sapienza, P., L. Zingales, D. Maetripier (2009) ‘Gender differences in financial risk aversión and career choices are affected by testosterone’, Proceedings of the National Academy of Sciences of the United States of America, 106: 36, 15268-15273.

Statman, M., Thorley, S. and Vorkink, K. (2006), “Investor Overconfidence and Trading Volume”, The Review of Financial Studies, 19(4), 1531-1565.

Svenson, O. (1981), “Are We All Less Risky and More Skillful than Our Fellow Drivers?”, Acta Psychologica, Vol. 47, 143-48.

Table 1: Profile of participants

The final sample of 146 participants included 41 female and 105 male participants, from a total of 35 countries, with approximately half of our sample British (46) or Chinese (30). The ages of participants ranged from 19 to 42 years old, with the majority being between the age of 19 to 25. 108 were undergraduate students and 38 postgraduate with the largest proportion were majoring in business related subjects

	Undergraduate	Postgraduate	Total
Finance	42	24	66
Business & Management	23	1	24
Computer Science and Engineering	4	4	8
Economics	22	6	28
Law	10	1	11
Others	7	2	9
Total	108	38	146

Table 2: Correlation of variables

This table displays the correlation coefficient between the variables obtained from the questionnaires and from the trading challenge. Trading activity variables include total trade, average transaction value and the total value of trades. Measures of overconfidence include the BTA and CONF, and an additional measure of competence, COMP. The dummy variables are gender (male/Female), type of degree studies (finance/non finance) and investment experience, INV EXP. P-values of the correlation coefficient presented in parentheses.

	Total Trades	Average Transaction Value	Total Value of Trades	BTA	CONF	COMP	Gender	Finance	INV EXP	Logins	Portfolio Returns
Total Trades	1										
Average Transaction Value	-0.2521	1									
	(0.0021)										
Total Value of Trades	0.5694	0.2983	1								
	(0.000)	(0.0003)									
BTA	-0.0061	-0.1579	-0.0952	1							
	(0.9414)	(0.057)	(0.2529)								
CONF	0.146	-0.0603	0.0707	-0.1391	1						
	(0.0787)	(0.4695)	(0.3968)	(0.094)							
COMP	-0.084	0.1755	-0.0334	-0.1192	-0.6715	1					
	(0.3132)	(0.0341)	(0.6894)	(0.1517)	(0.000)						
Gender	0.0355	0.2374	0.1843	-0.1847	0.1481	-0.1319	1				
	(0.6705)	(0.0039)	(0.0259)	(0.0256)	(0.0744)	(0.1125)					
Finance	0.0943	0.0991	0.1236	0.0524	0.0186	0.1762	-0.1368	1			
	(0.2575)	(0.234)	(0.1372)	(0.5303)	(0.8239)	(0.0334)	(0.0998)				
INV EXP	0.0152	0.2039	0.0369	-0.0886	0.0739	0.1336	0.036	0.017	1		
	(0.8552)	(0.0135)	(0.6587)	(0.2876)	(0.3757)	(0.1079)	(0.6666)	(0.839)			
Logins	0.5009	-0.0837	0.3983	0.0729	-0.0167	0.045	-0.0553	0.1444	-0.1467	1	
	(0.0000)	(0.3154)	(0.0000)	(0.3816)	(0.8411)	(0.5897)	(0.5076)	(0.0821)	(0.0772)		
Portfolio Returns	-0.0306	-0.1968	-0.2025	-0.0846	-0.1628	0.1726	0.0416	-0.1039	-0.0977	-0.0881	1
	(0.7141)	(0.0173)	(0.0143)	(0.3099)	(0.0496)	(0.0373)	(0.6181)	(0.2118)	(0.2409)	(0.2904)	

Table 3: Wilcoxon tests for equality of median

This table presents the results of the significance tests for univariate sorts based on the Wilcoxon signed rank test of the null hypothesis that trading volume, as measured by total transactions and the average value of transactions, the frequency of logins and portfolio performance are identical for different grouping of investors. Investors are categorised according to whether they have above or below median scores for Finance or non-finance degrees, BTA scores and COMP and positive or negative values of CONF. P-values for differences have been presented in paranthesis.

	Total Transactions	Average Value of Transactions	Logins	Final return
CONF>0 (94)	14.5	1356.702	34.5	-0.38%
CONF<0 (52)	7.5	1530.257	26.5	-0.65%
Difference	7	-173.555	8	0.27%
	(0.010)	(0.242)	(0.535)	(0.891)
Above median BTA*	14.5	1575.364	28.5	-0.12%
Below median BTA*	9	1255.821	28	-0.66%
Difference	5.5	319.543	0.5	0.54%
	(0.698)	(0.101)	(0.998)	(0.368)
Above median COMP	10	1474.958	29	-0.34%
Below median COMP	11	1385.870	28	-0.53%
Difference	-1	89.088	1	0.20%
	(0.151)	(0.382))	(0.415)	(0.627)
Above median exp Return	9.5	1623.64	35	-0.53%
Below median exp Return	11	1165.61	24.5	-0.43%
Difference	-1.5	458.03	10.5	-0.10%
	(0.428)	(0.280)	(0.073)	(0.656)
Finance degree	13.5	1563.632	40	-0.95%
Non-finance degree	7.5	1238.673	25	-0.20%
Difference	6	324.959	15	-0.78%
	(0.045)	(0.433)	(0.064)	(0.225)
Investment experience	7	1657.80	25	0.00%
No investment experience	12	1169.735	32	-0.79%
Difference	-5	488.74	-7	0.78%
	(0.156)	(0.418)	(0.122)	(0.733)

* BTA=top 3 deciles (62 participants)

Table 4: Gender differences

The table below reports the median value for male and female participants for overconfidence measures (CONF, BTA and SCOMP) and trading variables (login, total transactions, average transaction values, total value of trades) and portfolio returns. The p-values for the differences between median report reported in the table below.

	Male	Female	Difference	p-value
Sample	105	41		
Overconfidence Measures:				
CONF	0.055	0.0173	0.0378	0.0664
BTA	4	5	-1	0.0276
COMP	0.675	0.700	-0.025	0.0785
Trading Variables:				
Logins	32	22	10	0.0645
Total Transactions	10	10	0	0.9809
Average Transaction Value	1659.45	750.67	908.78	0.0001
Total Value of Trades	16,572.50	8,279.63	8,292.87	0.0008
Portfolio returns	-0.82%	-0.02%	-0.80%	0.579

Table 5: OLS regressions – various measures of trading volume

The table presents the coefficients and p-values (in parentheses) for twelve OLS regressions in which the number and value of trades is regressed on the two overconfidence variables (BTA and CONF), competence (COMP) and three dummy variables for gender, finance degrees and prior investment experience (INV EXP). Note that lower the values of BTA, the more confident the participant (i.e. Decile 1 indicates the top 10%). In addition, for the following dummy variables, GENDER = 1 indicates male participants, FINANCE = 1 indicates that the participant is studying for a finance related degree and INV EXP = 1 indicates prior experience in the financial markets.

	PANEL A Total number of trades (log)				PANEL B Average transaction value (log)				PANEL C Total value of trades (log)			
	(1)	(2)	(2')	(3)	(1)	(2)	(2')	(3)	(1)	(2)	(2')	(3)
BTA	0.013				-0.108				-0.096			
	(0.768)				(0.028)				(0.106)			
CONF		1.423				-1.214				0.209		
		(0.111)				(0.238)				(0.866)		
CONFDUM			0.448				-0.435				0.012	
			(0.025)				(0.058)				(0.964)	
COMP				-1.432				2.693				1.262
				(0.157)				(0.020)				(0.365)
GENDER	0.111	0.048	0.025	0.064	1.102	1.242	1.270	1.265	1.213	1.290	1.295	1.329
	(0.609)	(0.821)	(0.905)	(0.763)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
FINANCE	0.394	0.384	0.343	0.439	0.484	0.479	0.520	0.387	0.878	0.863	0.864	0.826
	(0.042)	(0.045)	(0.073)	(0.024)	(0.027)	(0.030)	(0.019)	(0.079)	(0.001)	(0.001)	(0.001)	(0.002)
INVEXP	-0.363	-0.390	-0.384	-0.327	-0.011	0.053	0.049	-0.044	-0.373	-0.338	-0.335	-0.371
	(0.084)	(0.060)	(0.062)	(0.118)	(0.964)	(0.825)	(0.837)	(0.854)	(0.190)	(0.239)	(0.242)	(0.199)
constant	2.158	2.198	2.013	-1.432	6.484	5.978	6.160	2.693	8.642	8.176	8.173	1.262
	(0.000)	(0.000)	(0.000)	(0.157)	(0.000)	(0.000)	(0.000)	(0.020)	(0.000)	(0.000)	(0.000)	(0.365)
R^2	0.0225	0.0659	0.0824	0.0624	0.1864	0.1661	0.1790	0.1898	0.1809	0.1656	0.1655	0.1703

Table 6: Logins

The table presents the coefficients and p-values (in parentheses) for OLS regressions in which the number of logins and trades per login are regressed on the two overconfidence variables (BTA and CONF), competence (COMP) and three dummy variables for gender, finance degrees and prior investment experience (INV EXP). Note that lower the values of BTA, the more confident the participant (i.e. Decile 1 indicates the top 10%). In addition, for the following dummy variables, GENDER = 1 indicates male participants, FINANCE = 1 indicates that the participant is studying for a finance related degree and INV EXP = 1 indicates prior experience in the financial markets.

	PANEL A Number of Logins				PANEL B Trades per Login			
	(1)	(2)	(2')	(3)	(1)	(2)	(2')	(3)
BTA	0.000				0.013			
	(0.995)				(0.737)			
CONF		-0.680				2.102		
		(0.427)				(0.008)		
CONFDUM			-0.052				0.499	
			(0.788)				(0.005)	
COMP				1.152				-2.584
				(0.233)				(0.004)
GENDER	0.485	0.510	0.494	0.514	-0.374	-0.461	-0.468	-0.450
	(0.020)	(0.014)	(0.017)	(0.012)	(0.055)	(0.015)	(0.014)	(0.017)
FINANCE	0.431	0.436	0.437	0.396	-0.037	-0.052	-0.094	0.044
	(0.020)	(0.018)	(0.019)	(0.034)	(0.830)	(0.755)	(0.579)	(0.796)
INV EXP	-0.375	-0.364	-0.373	-0.407	0.012	-0.027	-0.011	0.080
	(0.061)	(0.068)	(0.062)	(0.042)	(0.949)	(0.884)	(0.952)	(0.660)
Constant	2.876	2.885	2.899	1.152	-0.718	-0.688	-0.886	-2.584
	(0.000)	(0.000)	(0.000)	(0.233)	(0.008)	(0.000)	(0.000)	(0.004)
R^2	0.0847	0.0888	0.0852	0.0939	0.0291	0.0775	0.0815	0.0840

Table 7: Portfolio Returns
The table presents the coefficients and p-values (in parentheses) for OLS regressions in which portfolio returns are regressed on the number of transactions and the average value of transactions, the two overconfidence variables (BTA and CONF), competence (COMP) and three dummy variables for gender, finance degrees and prior investment experience (INV EXP). Note that lower the values of BTA, the more confident the participant (i.e. Decile 1 indicates the top 10%). In addition, for the following dummy variables, GENDER = 1 indicates male participants, FINANCE = 1 indicates that the participant is studying for a finance related degree and INV EXP = 1 indicates prior experience in the financial markets.

	(1)	(2)	(2')	(3)
Log transactions	0.003	0.005	0.005	0.005
	(0.675)	(0.481)	(0.531)	(0.470)
Log transactions value	-0.016	-0.015	-0.015	-0.018
	(0.018)	(0.020)	(0.023)	(0.006)
BTA	-0.006			
	(0.144)			
CONF		-0.177		
		(0.023)		
CONFDUM			-0.022	
			(0.221)	
COMP				0.283
				(0.001)
GENDER	0.021	0.031	0.028	0.034
	(0.306)	(0.125)	(0.170)	(0.081)
FINANCE	-0.013	-0.013	-0.012	-0.022
	(0.459)	(0.439)	(0.492)	(0.191)
INV EXP	-0.022	-0.016	-0.018	-0.027
	(0.231)	(0.374)	(0.317)	(0.132)
Constant	0.101	0.068	0.076	-0.109
	(0.066)	(0.154)	(0.125)	(0.124)
R^2	0.0715	0.0915	0.0672	0.1241