Facial structure and achievement drive : evidence from financial analysts

He, Xianjie; Yin, Huifang; Zeng, Yachang; Zhang, Huai; Zhao, Hailong

2018

He, X., Yin, H., Zeng, Y., Zhang, H., & Zhao, H. (2019). Facial structure and achievement drive : evidence from financial analysts. *Journal of Accounting Research*, 57(4), 1013-1057.
doi:10.1111/1475-679X.12259

<https://hdl.handle.net/10356/144447>

<https://doi.org/10.1111/1475-679X.12259>

© 2019 University of Chicago on behalf of the Accounting Research Center. All rights reserved. This paper was published in *Journal of Accounting Research* and is made available with permission of University of Chicago on behalf of the Accounting Research Center.

Downloaded on 26 Apr 2023 17:48:03 SGT

Facial Structure and Achievement Drive: Evidence from Financial Analysts

Xianjie He

Shanghai University of Finance and Economics
he.xianjie@mail.shufe.edu.cn

Huifang Yin

Shanghai University of Finance and Economics
yin.huifang@mail.shufe.edu.cn

Yachang Zeng

Nanyang Technological University
y.zeng@ntu.edu.sg

Huai Zhang

Nanyang Technological University
huaizhang@ntu.edu.sg

Hailong Zhao

Southwestern University of Finance and Economics
zhaohl@swufe.edu.cn

This paper is accepted for publication in *Journal of Accounting Research*.

Accepted by Christian Leuz. We are grateful for the constructive comments by two anonymous reviewers. We also appreciate helpful comments from Margaret Abernethy, Jing He, Xue Jia, Yuping Jia, Igor Kadach, Flora Kuang, Laurence van Lent, Bo Qin, Mohan Venkatachalam, Jin Zhang and seminar participants at Nanyang Technological University and The University of Melbourne. We thank Jiahui Fan, Jingyi Wu, Xue Lai for their help in collecting the data. A previous draft of this paper was titled “Achievement drive and analysts’ performance”. He and Yin acknowledge the financial support from the MOE project of Key Research Institute of Humanities and Social Science in University (No. 16JJD790037). Yin also acknowledges the financial support from the National Natural Science Foundation of China (Grant number: 71302076). Zeng acknowledges financial support from Singapore Ministry of Education and NTU Start-Up Grant (RG59/17 and M4081654). Zhang acknowledges the financial support from Singapore Ministry of Education for providing research funding (RG75/16 and RG163/17). All errors are our own. An Online Appendix to this paper can be downloaded at <http://research.chicagobooth.edu/arc/journal-of-accounting-research/online-supplements>.

Facial Structure and Achievement Drive: Evidence from Financial Analysts

ABSTRACT

We use a large pictorial sample of Chinese financial analysts to test the association between facial width-to-height ratio (*fWHR*) and performance in men. Financial analysts offer an ideal setting for our investigation because we can objectively track individual analysts' behaviors and performance. We find that high-*fWHR* analysts are more likely to conduct corporate site visits and they exhibit better performance. The positive *fWHR*-performance association survives a battery of robustness checks and the association is more pronounced for analysts with lower status, for firms with higher uncertainty, and for analysts facing more intense competition. Our results suggest that the dominant trait predicted by *fWHR* is achievement drive.

Keywords: Facial structure; Achievement drive; Financial analysts

JEL Codes: M20; M40; M41; M50

1. Introduction

A nascent literature explores economic consequences of physical traits that are important to capital market participants. Hobson et al. (2012) and Mayew et al. (2013) find that CEOs' voice pitch can predict their compensation levels and their likelihood to misreport while Addoum et al. (2017) show that individuals' height and obesity influence their portfolio choices. Facial width-to-height ratio (*fWHR*), in particular, has garnered considerable academic and public attention because *men* with higher *fWHR* are shown to exhibit more masculine behaviors, probably due to testosterone influencing both facial structure and behaviors (Carré and McCormick 2008; Haselhuhn and Wong 2012; Haselhuhn et al. 2014; Jia et al. 2014; Kamiya et al. 2018; Stirrat and Perrett 2010). Many prior studies assume that the effects of testosterone on masculine behaviors are through autonomic arousal, resulting in *innate* masculine behaviors (i.e., the *autonomic arousal* hypothesis). Recent research, however, suggests that testosterone is more closely associated with achievement-striving incentives (Eisenegger et al. 2011; Fisk et al. 2017) and hypothesizes that the dominant trait predicted by *fWHR* is *achievement drive*, which manifests behaviorally as *calculated* masculine behavioral traits such as aggression, deception and risk-taking (Lefevre et al. 2014; Lewis et al. 2012; Welker et al. 2015).¹ The achievement drive hypothesis marks an important step forward in our understanding of the association between *fWHR* and behaviors, because it suggests that we may be unable to observe the positive link between *fWHR* and masculine behaviors when these behaviors hinder attaining achievement.

The central testable prediction of the achievement drive hypothesis is that individuals with higher *fWHR* will exert more efforts and perform better due to their higher achievement drive. We use a large pictorial sample of Chinese financial analysts to test this hypothesis in a natural-occurring real business setting. This setting provides a set of key advantages for our research. First, this setting allows us to distinguish between the *autonomic arousal* hypothesis and *achievement drive* hypothesis. If the *autonomic arousal* explanation dominates, we don't expect to observe a positive relation between *fWHR* and analyst performance in China, because aggressive behaviors are socially undesirable within the collaborative workplace culture in East Asia (Re and Rule 2016; Rule et al.

¹ Please refer to Figure 1 for the illustration of the two hypotheses.

2010; Rule et al. 2011). If analysts with high *fWHR* behave aggressively, they would have a lower likelihood of getting information from managers and other capital market participants and consequently their performance would suffer. Second, firms listed on China's Shenzhen Stock Exchange are required to report analysts' corporate site visit records, allowing us to directly and objectively measure analysts' efforts (Cheng et al. 2016). While the *achievement drive* hypothesis predicts a positive association between *fWHR* and corporate site visits, the *autonomic arousal* hypothesis predicts the opposite: since high-*fWHR* analysts are more likely to take a risk and gamble on their own information, they are less likely to conduct a corporate site visit. Third, all Chinese financial analysts are required to register with Securities Association of China (SAC) and provide a license ID photo. This requirement alleviates the concern of *selection bias* and the concern of *measurement errors* due to bad picture quality. Fourth, the link between observable performance and *fWHR* is tighter for our sample because an analyst's *individual* performance can be measured *objectively*. This compares favorably with the setting of corporate executives, where accounting numbers of the firms are used to gauge the executives' performance. These numbers are subject to managerial manipulations and may reflect the collective performance of the management team instead of individual executives.²

We examine a sample of 1,193 *male* analysts, with their license ID photos collected from the website of SAC. Following prior literature, we consider male individuals only in our main analyses because the association between *fWHR* and behavior is valid for *men* only. As a robustness check, we consider female analysts in Section 7.3. While there is computer software that can automatically measure *fWHR* by detecting the position of a certain sense organ on the picture (e.g., the eyes), we follow the predominant approach in this line of literature to manually measure the analysts' *fWHR* for our sample.³ We measure analyst's performance through forecast accuracy (Clement and Tse 2005). Hong and Kubik (2003) and Wu and Zang (2009) show that more accurate forecasters are more likely

² Wong et al. (2011) document a positive association between a CEO's *fWHR* and his financial performance (measured by the return on assets). However, this finding might be explained by Jia et al. (2014) who document a positive association between upward earnings management and *fWHR*.

³ There are two important caveats to the software-based auto-approach. First, the software might falsely detect the organ positions if the individual on the picture wears glasses or has mustache (or heavy beard). Second, the software might not be able to correctly classify the facial expression (e.g., a genuine smile or a fake smile). The failure in classifying a genuine smile will lead to a measurement error in *fWHR* (Jia et al. 2014).

to move up the brokerage house hierarchy and be promoted to research executive positions, while Mikhail et al. (1999) find that an analyst has a greater likelihood of being fired if her/his forecast accuracy declines relative to the peers. These papers suggest that forecast accuracy is used by brokerage houses to judge analysts' performance in promotion and turnover decisions.⁴ Consistent with our prediction, we find that analysts with above-median *fWHR* exhibit higher forecast accuracy than analysts with below-median *fWHR*. The difference is about 2.1%–4.7% of the mean value of the scaled forecast accuracy measure in our sample, which is as big as the economic effect of analysts' experience. Our robustness check shows that an analyst's *fWHR* is also positively associated with other important measures of analyst performance: (1) high-*fWHR* analysts (i.e., analysts with above-median *fWHR*) are about 43 percent more likely to be selected as star analysts, (2) they issue more profitable stock recommendations, (3) they are 2.045 times more likely to bring investment banking deals to their employers, (4) they are 13.2 percent more likely to issue bold forecasts (Clement and Tse 2005),⁵ (5) they display more consistent forecast errors (Hilary and Hsu 2013), and (6) they issue more timely forecasts than low-*fWHR* analysts (Hilary and Hsu 2013). There is no evidence that they are optimistic (and/or overconfident) or they walk down their forecasts to curry favor with the management (Bradshaw et al. 2016; Hilary and Menzly 2006). Our main result is robust to various perturbations to the sample. In untabulated tests (reported in our online appendix), we show that high-*fWHR* analysts are more likely to follow bigger and better-performing firms.⁶

Next, we document the effort channel between *fWHR* and analyst performance. We use an analyst's corporate site visit records to proxy for his efforts because Cheng et al. (2016) show that corporate site visits are an important way for Chinese analysts to gather information. We find that the

⁴ We do not claim that forecast accuracy is the most important performance measure for financial analysts. We only hypothesize that greater efforts from analysts with higher *fWHR* result in more accurate earnings forecasts. Compared with other performance measures (e.g., stock picking ability and forecast consistency), forecast accuracy allows a substantially larger sample. Hence, we use forecast accuracy as our main performance measure and provide robustness tests using alternative analyst performance measures.

⁵ Clement and Tse (2005) provide evidence suggesting that bold forecasts incorporate analysts' private information and are more informative. Therefore, we use forecast boldness to measure analyst performance rather than *imrate* risk-taking.

⁶ It is important to note that our regression variables are largely based on relative rankings among analysts following the *same* firm in the *same* year (Clement and Tse 2005). Therefore, differences in firm characteristics (e.g., size) between the two groups of analysts are unlikely to affect our conclusions on the *fWHR*-performance link. We also find that our main finding continues to hold when we use a random correlated effects model to implement analyst fixed effects for controlling the analyst-level factors unrelated to *fWHR*.

odds of conducting a corporate site visit are about 12%–15% higher for high-*fWHR* analysts than for low-*fWHR* analysts. In addition, we find that corporate site visit mediates the relationship between *fWHR* and forecast accuracy: on average, 8% of the estimated total effect of *fWHR* on forecast accuracy is due to the mediation effect of corporate site visit (with a 95% confidence interval for the mediation effect ranging from 5% to 24%). Given the myriad ways in which the achievement drive can influence analysts' forecasting performance, we deem the mediation effect meaningful and economically substantive.

We next provide further evidence that the trait of achievement drive is likely responsible for the *fWHR*-performance association in the analyst setting. Goetz et al. (2013) and Welker et al. (2015) offer such evidence in the laboratory setting and the professional hockey player setting by examining the moderating effect of status on the *fWHR*-behavior association. Similarly, if the achievement drive hypothesis is responsible for our finding, we expect that the drive of those high-*fWHR* analysts is dampened when a certain achievement, e.g., a higher status, has been attained. In this case, the reduced achievement desire and the limited upward mobility could result in a smaller difference in achievement striving between the high- and low-*fWHR* analysts. We therefore expect the positive *fWHR*-performance relationship to be less (more) pronounced among analysts with high (low) status. Our results show that the positive association between *fWHR* and forecast accuracy is only observed in the subsample of analysts with low-status (i.e., analysts with lower titles and analysts employed by smaller brokerage firms),⁷ suggesting that the *fWHR*-performance relationship in our sample is due to the trait of achievement drive. In addition, the achievement drive hypothesis predicts that the *fWHR*-performance association is more pronounced for firms with more uncertainty, because for these firms, there is more publicly unknown information that can be unraveled by high-*fWHR* analysts' information collection activities, resulting in a greater difference in performance between high- and low-*fWHR* analysts. Our empirical results support this prediction.

⁷ To provide evidence supporting the achievement drive hypothesis, the status measure cannot capture an easy-to-lose status, because the achievement desire is supposed to be dampened only if an analyst obtains a sustainable status. With this understanding in mind, we use brokerage size and analyst title as the status proxy because once an analyst begins to work for a bigger brokerage house, he usually can stay with the firm for a long period of time, and because once a title is given, it typically stays.

Although the *fWHR*-behavior link is well documented, the exact underlying neuro mechanism remains subject to debate in neurosciences. The most discussed hypothesis in the prior literature is the *pubertal* testosterone-based mechanism. This hypothesis conjectures that testosterone exposure (1) plays a role in the development of male secondary sex traits during adolescence (e.g., facial structure) and (2) has a permanent influence on the central nervous system in brain that affects behavior (Carré and McCormick 2008). Hence, evidence on sexual dimorphism in *fWHR* is important for supporting this hypothesis (Lefevre et al. 2013). However, prior studies fail to find consistent evidence on sex difference in *fWHR*, probably due to their small sample size or ethnic diversity in their samples (Geniole et al. 2015; Kramer 2017).⁸ Our setting is ideal for addressing these issues. Since Han Chinese are ethnically homogenous and our sample comprises a large number of Chinese analysts, our tests have high statistical power, and our results are less likely to be affected by variations in *fWHR* due to ethnicity or profession.⁹ We find evidence consistent with the hypothesis of sexual dimorphism in *fWHR*, which is in line with the *pubertal* testosterone-based mechanism.¹⁰

To further study the testosterone-based mechanism, we examine the effect of competition threats on the *fWHR*-performance association. Prior literature provides evidence that the effect of testosterone on achievement-striving is more pronounced when individuals face threats (Derntl et al. 2009; Eisenegger et al. 2011; Goetz et al. 2014; Hermans et al. 2006; Hermans et al. 2008). If the testosterone-based mechanism underlies the association between facial structures and achievement drive, greater competition threats will lead to a stronger link between *fWHR* and achievement drive, and consequently a stronger *fWHR*-performance association. Our findings are consistent with this prediction.

It is worth noting that our findings do not support the self-fulfilling prophesy mechanism, i.e., to the extent that individuals use *fWHR* to guide judgments of personal traits in others, the anticipated

⁸ Although the existing evidence on sexual dimorphism in *fWHR* is not consistent, Lefevre et al. (2013) find a positive correlation between *fWHR* and *circulating* testosterone concentrations in *adult men*, suggesting the role of circulating testosterone explaining the *fWHR*-behavior relationship. However, their finding is challenged by Bird et al. (2016), who fail to obtain such a finding in a large sample, and Bird et al. (2016) call for more research on the *pubertal* testosterone-based hypothesis.

⁹ The downside of the setting is that our findings may not be generalizable to other ethnicity.

¹⁰ Due to the lack of other biometric variables in our sample, we are unable to prove that the documented sexual dimorphism in *fWHR* is independent of the other biometrics (e.g., body size). Therefore, our findings should be interpreted with this caveat in mind.

behaviors of the face owner become *self-fulfilling* (Eckel 2007; Haselhuhn et al. 2013). Prior literature documents that low-*fWHR* individuals are *perceived* to be more cooperative and less aggressive (Short et al. 2012; Stirrat and Perrett 2010). In addition, these warmth-related traits are shown to be associated with success in East Asia (Re and Rule 2016; Rule et al. 2010; Rule et al. 2011). Hence, the self-fulfilling prophesy predicts that low-*fWHR* analysts perform better than high-*fWHR* analysts in China, a prediction unsupported by our empirical results.

Finally, prior literature finds that the *fWHR*-behavior relationship does not hold in women (Carré and McCormick 2008; Carré et al. 2013; Geniole et al. 2014; Goetz et al. 2013; Haselhuhn and Wong 2012; Stirrat and Perrett 2010). We find no *fWHR*-performance relationship among female analysts, although our results can be due to the smaller sample size.

Our paper contributes to the growing literature on *fWHR* in three ways. First, our results support recent finding in neuroscience that *fWHR* is a valid facial cue to the achievement-striving trait that can manifest behaviorally as other masculine behaviors (Lefevre et al. 2014; Lewis et al. 2012). In doing so, our paper answers the call of Welker et al. (2015) for more research investigating the achievement drive hypothesis.¹¹ Second, we provide evidence that potentially deepens our understanding of the underlying mechanism of the *fWHR*-behavior link. Our evidence on sexual dimorphism in *fWHR* is consistent with the *pubertal* testosterone-based mechanism.¹² Third, prior studies offer a puzzling finding that the relation between *fWHR* and aggressive behaviors doesn't exist in certain regions (Özener 2012). By showing evidence that *fWHR* mainly predicts the achievement-striving trait, we offer a plausible explanation that, in regions where aggressive behaviors are socially undesirable, the *fWHR*-aggression relationship might not hold, since aggressive behaviors are not compatible with the achievement-striving incentives.

¹¹ Our setting is useful to test the achievement drive hypothesis because analyst performance depends largely on efforts to acquire information. Lebuda and Karwowski (2016) examine the *fWHR* among nominees and laureates of the Nobel Prize. They find that high-*fWHR* individuals are nominated at a younger age, but *fWHR* does not predict the probability of being nominated for or awarded the Nobel Prize. However, Nobel Prize nominations arguably depends more on creativity than efforts and the authors interpret their findings as suggesting a *fWHR*-creativity link. In addition, unlike our setting, their sample of eminent writers is outside of the business context.

¹² Given that the documented *fWHR*-behavior link is supposed to be due to the effects of testosterone, the ethnic differences should not affect our results unless testosterone affects the behavior and *fWHR* differently among different ethnic groups. To the best of our knowledge, there is no existing study examining the cross-ethnic differences in the testosterone effects yet.

Our study also deepens our understanding of how *fWHR* is associated with individual performance in the business context, answering Jia et al. (2014)'s call for more research in this area. Jia et al. (2014) show that a high-*fWHR* CEO is more likely to misreport and that a CEO's *fWHR* is greater than the average *fWHR* of laboratory participants. Their findings beg the question why firms hire high-*fWHR* individuals as CEOs, given their higher propensity to misreport. Our results suggest that high-*fWHR* individuals have a stronger drive for achievement, thus it may be unsurprising to see them at the top of the corporate hierarchy.

Jacob et al. (1999) and Leone and Wu (2007) have found consistent differences in individual analyst's performance. Prior studies identify several *acquired* traits that affect individual analysts' job performance such as professional qualification, experience, resources, reputation, industry specialization, social network and political contribution (Stickel 1992; Clement 1999; Clement et al. 2007; Ramnath et al. 2008; De Franco and Zhou 2009; Cohen et al. 2010; Jiang et al. 2016). Instead, we focus on a *primitive* measure of analysts' traits (i.e., *fWHR* that predicts achievement drive) and we find this measure significantly associated with analyst performance. Our results therefore offer a solid first step towards understanding analyst aptitude from a novel perspective.

2. Facial width-to-height ratio, testosterone, and behaviors

2.1 Mechanisms linking facial width-to-height ratio to behaviors and sexual dimorphism

Recently, facial width-to-height ratio (*fWHR*) has garnered considerable attention from researchers because it has been established as a remarkably robust predictor of a wide range of masculine behaviors in *men* (Carré and McCormick 2008; Carré et al. 2009; Haselhuhn and Wong 2012; Haselhuhn et al. 2014; Jia et al. 2014; Lefevre and Lewis 2014; Stirrat and Perrett 2010; Třebický et al. 2015; Zilioli et al. 2015). However, the underlying mechanisms linking *fWHR* to behaviors remain subject to debate in neurosciences. Some scholars suggest an *self-fulfilling prophecy* explanation (Haselhuhn et al. 2013), i.e., to the extent that individuals judge others' personal traits by their *fWHR* (Stirrat and Perrett 2010), the face owners behave in ways that fulfil the observers' expectations (Eckel 2007). Nevertheless, the majority of the literature hypothesizes that the link between *fWHR* and behaviors is due to the functioning of the hormone testosterone during

adolescence (i.e., *pubertal* testosterone-based mechanism) (Bird et al. 2016). This mechanism suggests that pubertal testosterone exposure (1) influences the development of male *secondary sex traits* during adolescence (e.g., *fWHR*) (Verdonck et al. 1999), and (2) exerts organizational effects sculpting the brain's neural structures associated with the ability to program activational responses to steroids (e.g., circulating testosterone), which affects behaviors later in life (Sisk et al. 2003; Sisk and Zehr 2005).

The *pubertal* testosterone-based mechanism hypothesizes that the effect of testosterone on pubertal facial growth results in sexual dimorphism in face shape, with men having greater *fWHR* than women (Carré and McCormick 2008). Existing studies find inconsistent evidence of sexual dimorphism in *fWHR*. Some work finds that *fWHR* is larger in males than in females, supporting the sexual dimorphism in *fWHR* (Carré and McCormick 2008; Geniole et al. 2015; Sanchez-Pages et al. 2014; Weston et al. 2007). Some studies, however, report no difference in *fWHR* between men and women (Kramer et al. 2012; Özener 2012; Stirrat et al. 2012). Lefevre et al. (2012) even find that females have higher *fWHR* than males in one of their subsamples, and conclude that *fWHR* does not appear to be sexually dimorphic.

Kramer (2017) points out that the mixed evidence might be due to that the test subjects include populations with different ethnicities or geographical origins. In addition, Alrajih and Ward (2014) find evidence of between-profession differences in *fWHR*. Jointly, the two papers suggest the importance of obtaining a large sample of observations with the same profession and the same ethnicity for testing the sexual dimorphism in *fWHR*.¹³

2.2 Autonomic arousal hypothesis versus achievement drive hypothesis

Jia et al. (2014) classify all the *fWHR*-associated masculine behaviors into four traits, i.e., aggression, egocentrism, risk-taking and status-seeking, without spelling out which specific trait dominates. Prior studies implicitly assume that those are *innate* masculine behaviors characterized by autonomic arousal (the *autonomic arousal* hypothesis). Recent studies in neuroscience however offer

¹³ Alrajih and Ward (2014) find that CEOs have greater *fWHR* than age- and sex-matched individuals.

evidence that *achievement drive* seems to be the first-order trait, which manifests behaviorally as *calculated* masculine behaviors (the *achievement drive* hypothesis).

These studies suggest that the link between testosterone and human social behavior is a situationally-contingent manifestation of a broader motivation of achievement-striving (Eisenegger et al. 2010; Eisenegger et al. 2011; Fisk et al. 2017; van Honk et al. 2011; Boksem et al. 2013; Smeets-Janssen et al. 2015; van Honk et al. 2016). The underlying theory is that testosterone induces a greater concern for achievement. When threat of failure emerges, this concern generates an enhanced predisposition to aggressively seek achievements through masculine behaviors (e.g., aggression, egocentrism, and risk-taking). Consistent with this theory, these studies show that, when aggressive behaviors hinder achievements, high-testosterone individuals tend to behave less aggressively than low-testosterone individuals. Therefore, testosterone's pivotal role in promoting masculine behaviors is inherently triggered by the achievement-driven mentality (van Honk et al. 2011). Since *pubertal* testosterone exposure is shown to have organizational effects on the nervous system related to the ability to program the above activational responses to circulating testosterone (Sisk et al. 2003; Sisk and Zehr 2005), higher *pubertal* testosterone exposure is supposed to result in higher achievement-striving traits later in life (Fisk et al. 2017).

Recent studies provide results consistent with the achievement drive hypothesis. Goetz et al. (2013) find that the positive *fWHR*-aggression association is more pronounced among individuals with low social status, supporting the notion that masculine behaviors are rooted in the desire to achieve high status and the desire is stronger when individuals' current status is low. Similarly, Welker et al. (2015) show that the association between *fWHR* and risk-taking only exists among low-status individuals, while Stirrat and Perrett (2012) demonstrate that the *fWHR*-egocentrism relationship is weakened in situations where being egocentric results in lower perceived social status. Lewis et al. (2012) provide more direct evidence by documenting a positive association between *fWHR* and the trait of achievement-striving, using Simonton (1986)'s psychometric analysis of former U.S. presidents. The preponderance of evidence is consistent with the hypothesis in Lefevre et al. (2014) and Lewis et al. (2012) that the masculine traits predicted by *fWHR* are the behavioral manifestation of achievement striving.

We use the setting of Chinese financial analysts to test the achievement drive hypothesis. The main prediction is that analysts with higher *fWHR* have a stronger desire to strive for achievement and success. Consequently, they put more effort in their jobs and exhibit higher performance. As discussed in the introduction, this prediction is unattainable under the *automatic arousal* hypothesis, since the East Asian culture disparages aggressive behaviors (Re and Rule 2016; Rule et al. 2010; Rule et al. 2011).

3. Variable measurement, sample selection and summary statistics

3.1 Measurement of the facial width-to-height ratio and the endogeneity issue

Following the literature on facial width-to-height ratio (*fWHR*), we measure a given analyst's *fWHR* as the distance between the two zygions (bizygomatic width) relative to the distance between the upper lip and the highest point of the eyelid (height of upper face) from the analyst's photograph. We collect license ID photos of *male* Chinese financial analysts from the website of Securities Association of China (SAC), in which head-and-shoulder views cover around 65% to 75% of the photo. In most of the pictures, the facial outline, eyes, lip and ears of an individual can be identified clearly. Therefore, the SAC website provides us a unique pictorial database that can help to accurately measure the *fWHR* of financial analysts in China.

Following the procedure outlined in Jia et al. (2014), two research assistants independently measure each analyst's *fWHR* using the *ImageJ* software provided by the National Institutes of Health (Rasband 2015). The *fWHR* variable is coded as the average of the two measures if the difference between them is less than five percent. Otherwise, a third independent research assistant measures the facial structure again. If the difference between the third measurement and either of the first two measurements remains greater than five percent, we classify the picture as "low quality" because the substantial measurement error might be due to low picture resolution. If the difference is less than five percent, the picture is categorized as "good quality" and the *fWHR* value is the average value of the

two closest measurements.¹⁴ We also code pictures with a genuine smile as low-quality pictures.¹⁵ About 83% of the analysts' photos are categorized as "good quality". Consistent with prior literature (e.g., Jia et al. 2014) and for the ease of interpretation, our empirical tests are based on the dummy, $D(fWHR > median)$, an indicator variable which is equal to one if an analyst's $fWHR$ is above the median in the analyst-level sample, and zero otherwise.¹⁶

Similar to other empirical studies, we entertain the concern that our main variable of interest, $fWHR$, is endogenous. We take comfort in that an individual's facial structure is a predetermined biogenetic measure. Hence, the reverse causality issue is less severe in our setting. In addition, the $fWHR$ measure is unlikely determined by economic-wide factors, brokerage-level characteristics and post-adolescent analyst-level characteristics (e.g., experience), largely eliminating the possibility of confounding effects (i.e., a confounder exists determining both $fWHR$ and performance). Finally, we control for the lagged performance measure. To the extent that potential confounding factors are likely to influence lagged performance, our regression specification helps to tease out their effects. While we can't completely rule out the endogeneity concern, the "born with" nature of $fWHR$ renders this concern unlikely applicable to our study.¹⁷

3.2 Measure of forecast characteristics

The achievement drive hypothesis predicts that analysts with higher $fWHR$ perform better. In our main test, we measure analyst performance via forecast accuracy. Prior studies suggest that forecast accuracy is one important dimension in analysts' performance evaluation. Stickel (1992) and Leone and Wu (2007) find that all-star analysts' earnings forecasts are more accurate. Mikhail et al.

¹⁴ Alternatively, we divide the average of the two research assistants' measures of bizygomatic width by the average of their measures of height of upper face to calculate the ratio. Then, we repeat our main tests by applying this alternative $fWHR$ measure in Online Appendix 11. Results yield identical conclusions.

¹⁵ To obtain a non-biased $fWHR$ measure, the face in the picture should not be tilted by more than 45 degree and a genuine smile should not appear on the face (Jia et al. 2014). All the license photos in our sample meet the first criterion. However, there are a few photos that do not meet the second criterion.

¹⁶ Following Jia et al. (2014), we use the indicator variable transformation of $fWHR$ to reduce measurement error in $fWHR$ (Greene 2012, p. 379) and to better interpret the economic significance of the estimated effects. Similar results are obtained when we use the untransformed $fWHR$ variable in the regressions (Online Appendix 10).

¹⁷ Body Mass Index (BMI) is found to be associated with both $fWHR$ and cognitive abilities. We do not control for BMI in our analyses for two reasons. First, our sample does not have this biometric data. Second, since $fWHR$ is shown to be positively correlated with BMI, and BMI is found to be negatively associated with cognitive abilities (Azurmendi et al. 2005; Mayew 2013), omitting BMI from our regression biases against our finding by forcing the coefficient on $fWHR$ to be negative, assuming that performance is positively correlated with cognitive abilities.

(1999), Hong and Kubik (2003), and Wu and Zang (2009) document that brokerage firms use analysts' forecast accuracy to evaluate their performance in promotion and turnover decisions. In addition, forecast accuracy allows much more observations than other performance measures. We therefore use forecast accuracy as our main performance measure.¹⁸ Specifically, we follow Clement and Tse (2005) in measuring forecast accuracy.

$$Accuracy_{ijt} = \frac{AFE\ max_{jt} - AFE_{ijt}}{AFE\ max_{jt} - AFE\ min_{jt}} \quad (1)$$

where AFE_{ijt} is absolute forecast error (deflated by stock price two trading days before the earnings forecast) of analyst i following firm j in year t ; $AFE\ max_{jt}$ and $AFE\ min_{jt}$ are the maximum and minimum absolute forecast errors of all analysts following firm j in year t . The value of $Accuracy_{ijt}$ ranges between 0 and 1, and a higher value of $Accuracy_{ijt}$ indicates better forecasting performance. It's fairly straightforward to see that $Accuracy_{ijt}$ equals 1 (0) for the most (least) accurate analyst.

Following Clement and Tse (2005), we include the following control variables: *Lagaccuracy* (forecast accuracy of the same analyst's prior forecast for the same firm),¹⁹ *Dayselap* (the days elapsed since the last forecast issued by any analyst for the same firm), *Frequency* (the forecasting frequency of the analyst for the firm in the current fiscal year), *Horizon* (the time from the forecast date to the fiscal year-end), *Brokersize* (the number of analysts employed by the brokerage firm), *Experience* (the number of years that the analyst has followed the firm), *Companies* (the number of companies that the analyst follows), and *Industries* (the number of industries that the analyst follows). To facilitate comparisons and interpretations of coefficient estimates, following Clement and Tse (2005), we scale all the control variables to range from 0 to 1. The scaled control variables for analyst i following firm j in year t is

$$Characteristics_{ijt} = \frac{Raw\ Characteristic_{ijt} - Raw\ Characteristic\ min_{jt}}{Raw\ Characteristic\ max_{jt} - Raw\ Characteristic\ min_{jt}} \quad (2)$$

¹⁸ Instead of claiming that forecast accuracy is the most important performance measure for financial analysts, we hypothesize that greater efforts exerted by high-*fWHR* analysts result in more accurate earnings forecasts. In the robustness tests, we find evidence supporting the achievement drive hypothesis using alternative analyst performance measures.

¹⁹ We control for the lagged value of forecast accuracy because, to the extent that the confounding factors (e.g., education background and social connections) are likely to influence lagged performance, our regression specification helps to tease out their effects.

where $Raw\ Characteristic_{ijt}$ is the original forecast characteristics for analyst i following firm j in year t ; $Raw\ Characteristic\ max_{jt}$ and $Raw\ Characteristic\ min_{jt}$ are the original maximum and minimum values of a forecast characteristic of all analysts following firm j in year t .

We scale these variables because of the following reasons. First, this approach is highly popular and has been widely used in the analyst literature (e.g., De Franco and Zhou 2009; Kumar 2010; Clement and Law 2014; Yin and Zhang 2014). Using the same approach allows us to yield results that are comparable to prior studies. Second, this approach helps to control for unobservable firm-level characteristics and time-variant factors in our regressions. By indicating the relative ranking among all analysts following the same firm in the same year, the association between the scaled dependent variable and the scaled independent variable (e.g., analyst experience) is less likely affected by firm characteristics and year fixed effects, which may confound our conclusion when we use unscaled variables (Kumar 2010). For example, a positive association between the unscaled forecast accuracy and the unscaled experience variable might be because more experienced analysts are more likely to follow firms with better information environment, and it's better information environment but not higher analyst experience, that explains the higher forecast accuracy we observe. Third, this approach offers an intuitive and easy interpretation. For example, the coefficient on this scaled experience variable indicates the change in the dependent variable when we go from the least experienced to the most experienced analyst. Lastly, this approach allows comparisons in the impact of independent variables. For example, in a regression with the scaled forecast accuracy as the dependent variable, if the coefficient on the scaled experience variable is 0.5, while the coefficient on the scaled forecast frequency variable is 0.2, we can say that analyst experience has a greater impact on forecast accuracy than forecast frequency, because the scaled forecast accuracy improves by 0.5, when we go from the least experienced to the most experienced analyst, while it improves only by 0.2, when we go from the least frequent analyst to the most frequent analyst.

3.3 Sample selection and summary statistics

Table 1 reports our sample selection procedure. We start with a sample of analysts whose license ID photos are displayed on the SAC website in July 2015. We then draw data from CBAS

(Chinese Brokerage and Analyst System) to obtain these analysts' annual earnings forecasts between 2004 and 2014.²⁰ This procedure yields an initial sample of 2,108 analysts. We delete 249 male analysts and 115 female analysts due to the low picture quality issue.²¹ Next, we eliminate 551 female analysts because *fWHR* predicts behaviors in *men* only. This leads to a sample of 1,193 male analysts (296,244 forecasts) for constructing the final sample. We remove 159,148 forecasts that are issued by a group of analysts instead of individual analysts to assure that the forecasts in our sample indeed capture individual performance. Following Clement and Tse (2005), we delete 3,264 forecasts issued less than 30 days before the earnings announcement date, which were issued with extreme certainty. Consistent with prior studies, we retain the most recent forecast if an analyst issues multiple forecasts for the same firm-year combination. We further remove the forecasts for firms in the financial industry due to their distinct firm characteristics (Cheng et al. 2016). Finally, we retain observations with non-missing data for the regression analyses, which yields a final sample of 19,324 forecasts. We obtain firm-level variables from CSMAR, WIND, iFinD database or Shenzhen Stock Exchange. The Appendix provides data source(s) of each variable.

Table 2 presents the summary statistics of our sample. Panel A shows that the distribution of the *fWHR* measure at the analyst level is comparable to statistics reported in the prior literature (Kramer 2017), with a mean value of 2.166 (2.092) and a standard deviation of 0.181 (0.130) for male (female) analysts.²² The variables of analyst characteristics shown in Panel B, including the forecast accuracy, are scaled to range from 0 to 1 as described in Section 3.2. The means (medians) of *Accuracy* and *LagAccuracy* are, respectively, 0.677 and 0.603 (0.788 and 0.693). *Dayselap* has the mean of 0.267 and the median of 0.167. The mean and median values of *Frequency* are 0.300 and

²⁰ CBAS is an academic database on Chinese brokerage firm and financial analyst research. CBAS contains data on analysts' forecasts and recommendations. It also provides data on analyst and brokerage house characteristics, which are gathered from the brokerage firms' annual reports posted in the SAC website. The database is available on the platform of CNRDS (Chinese Research Data Services).

²¹ We also follow Jia et al. (2014)'s approach, using the full sample in the regressions and including the indicator variables that capture the potential measurement errors due to low picture quality. The results still hold.

²² We conduct a comprehensive survey and list the *fWHR* distributions of the samples examined by prior studies in Online Appendix 14. The distribution of *fWHR* in our sample is similar to the Chinese subsample examined by Kramer (2017). The mean value in our sample is 2.166 (2.092) for males (females), close to the mean of 2.140 (2.110) in Kramer (2017).

0.000 respectively.²³ The means (medians) of *Horizon* and *Brokersize* are, respectively, 0.308 and 0.431 (0.228 and 0.422). *Experience* has the mean (median) value of 0.398 (0.333). *Companies* and *Industries* average 0.280 and 0.304, with the median values of 0.168 and 0.200 respectively. The means (medians) of *AnalystAge* and *Postgraduate* are, respectively, 0.421 and 0.877 (0.364 and 1.000).²⁴

4. Research design and main empirical results

4.1 Empirical model specification and the *fWHR*-performance association

We employ the following regression model to examine the association between an analyst's *fWHR* and his performance.

$$\begin{aligned} Accuracy_{ijt} = & a_0 + a_1 D(fWHR > median)_i + a_2 LagAccuracy_{ijt} + a_3 Dayselap_{ijt} + a_4 Frequency_{ijt} \\ & + a_5 Horizon_{ijt} + a_6 Brokersize_{ijt} + a_7 Experience_{ijt} + a_8 Companies_{ijt} + a_9 Industries_{ijt} \\ & + a_{10} AnalystAge_{it} + a_{11} Postgraduate_i + \varepsilon_{ijt} \end{aligned} \quad (3)$$

The dependent variable, *Accuracy_{ijt}*, is the scaled forecast accuracy of analyst *i* following firm *j* in year *t* defined in Section 3.2. Our variable of interest is *D(fWHR > median)_i*, an indicator variable equal to one if analyst *i*'s *fWHR* is above the median, and zero otherwise. The control variables are discussed in Section 3.2 and the Appendix. Robust standard errors are clustered at the analyst level.²⁵ As discussed before, the achievement drive hypothesis (autonomic arousal hypothesis) predicts that *a₁*, the coefficient estimate on *D(fWHR > median)*, is positive (negative).

Consistent with the achievement drive hypothesis, the indicators, *D(fWHR > median)*, load positively and significantly for both columns 1 and 2 in Table 3, Panel A. The coefficient estimate on *D(fWHR > median)* in column 1 is 0.032 (with *p* < 0.01), indicating that the forecast accuracy of an analyst with above-median *fWHR* is 0.032 higher than that of an analyst with below-median *fWHR*,

²³ Since the scaled variables are calculated based on the sample including both male and female analysts, it is possible for analyst characteristics to have a median value of zero when we consider male analysts only in the summary statistics.

²⁴ We provide details on the variable constructions in the Appendix.

²⁵ We cluster standard errors by analysts to account for potential within analyst correlation because our variable of interest (i.e., *D(fWHR > median)*) is fixed at the analyst level (Angrist and Pischke 2009; Jia et al. 2014). We do not cluster standard errors by year because the number of clusters in the year-level dimension is below 40, which would increase bias if the standard errors are clustered along this dimension (Petersen 2009).

which is 4.7% of the mean value of the scaled forecast accuracy in our sample. Column 2 reports the multivariate regression with the control variables included. The coefficient estimate on $D(fWHR > median)$ is positive and significant at the 0.01 level. The magnitude of the coefficient shows that the economic effect of $D(fWHR > median)$ is as big as that of analysts' experience.

Consistent with prior literature (Cheng et al. 2016; Kumar 2010), we find that lagged forecast accuracy, forecast experience and the number of companies covered by an analyst are positively associated with forecast accuracy, while forecast frequency and forecast horizon are negatively associated with forecast accuracy.

4.2 Is the superior performance of high-fWHR analysts due to their extra efforts?

This section tests the effort channel linking $fWHR$ to analyst performance. In particular, we examine (1) whether analysts with higher $fWHR$ exert more efforts to gather information and (2) to what extent effort explains the $fWHR$ -performance relationship.

Analysts' efforts are generally difficult to be objectively captured by archival data. We take advantage of Shenzhen Stock Exchange's requirement that listed firms disclose analysts' corporate site visits. Using the corporate site visit data, Cheng et al. (2016) find that such visits increase analysts' forecast accuracy. In addition, the impact of the visits is stronger when the site visit is an analyst-only visit, when the current visit is preceded by fewer visits and when visiting analysts are based far from the visited firms. Their results suggest that a corporate site visit is an important information acquisition activity for analysts and it can serve as a reasonable proxy for analysts' efforts. If indeed analysts with higher $fWHR$ exert more information acquisition efforts, we expect that these analysts are more likely to conduct corporate site visits.

Specifically, we use a logistic regression to test this prediction. The dependent variable, $SiteVisit_{ijt}$, equals to one if analyst i conducts at least one corporate site visit to firm j in year t , and zero otherwise. We use the same set of control variables as used in Cheng et al. (2016), capturing the firm-, brokerage- and analyst-level characteristics. The variable construction is described in the Appendix. Following Cheng et al. (2016), we also control for industry-fixed effects and cluster the standard errors at the firm level. Column 1 (column 2) of Table 3, Panel B reports that the coefficient

on $D(fWHR > median)$ is 0.136 with $p < 0.01$ (0.117 with $p < 0.05$), suggesting that the odds of conducting a corporate site visit are about 12%–15% higher for a high- $fWHR$ analyst than for a low- $fWHR$ analyst. The signs of control variables are generally consistent with those in Cheng et al. (2016). Overall, our results indicate that high- $fWHR$ analysts exert more efforts to acquire information. This finding also helps to rule out the autonomic arousal hypothesis, which predicts that high- $fWHR$ analysts are less likely to conduct a corporate site visit because they are more likely to take a risk and gamble on their own information.

We next estimate the mediation effect of analyst effort on the $fWHR$ -performance relationship, using a path analysis depicted in Figure 2.²⁶ Path a (path b) estimates the direct effect of $D(fWHR > median)$ on *Accuracy* (*SiteVisit*), and path c estimates the direct effect of *SiteVisit* on *Accuracy*. The indirect effect of $fWHR$ on performance via efforts is estimated through path $b \times$ path c. Table 3, Panel C reports that, on average, 8% of the estimated total effect of $fWHR$ on forecast accuracy is via corporate site visit (with a 95% confidence interval for the mediation effect ranging from 5% to 24%). This effect is both statistically and economically significant. Furthermore, it is worth noting that we might underestimate the mediation effect of efforts on the $fWHR$ -performance relationship because *SiteVisit* cannot fully capture analysts' efforts. Nevertheless, the records of corporate site visit provide us an objective and observable proxy of forecast efforts, allowing us to document clear evidence on the effort channel between $fWHR$ and forecast accuracy.

4.3 Subsample analyses related to the achievement drive hypothesis

In this section, we provide further evidence on the achievement drive hypothesis by examining the role of status and operating uncertainty in moderating the relationship between $fWHR$ and performance.²⁷

4.3.1 Moderating effect of status

²⁶ We employ the mediation test using the “Medeff” function in Stata as it allows a binary dependent variable (Hicks and Tingley 2011).

²⁷ As discussed in Section 4.2, *SiteVisit* cannot fully capture analysts' efforts. Hence, we do not conduct subsample analyses for the association between $fWHR$ and *SiteVisit*, because the prediction will be unclear if the $fWHR$ -performance association is through other unobservable effort measures.

We first follow prior studies and examine the moderating effect of status on the association between *fWHR* and performance (Goetz et al. 2013; Welker et al. 2015). If the *fWHR*-performance association is explained by the achievement drive hypothesis, we would expect the association to be weaker among analysts with high status, because these analysts have fulfilled their achievement drive to a greater extent and they face restricted upward mobility.

To test our prediction, we use measures that capture sustainable status because it is hard to imagine that analysts become complacent after they receive a recognition that can be easily stripped off. In fact, they may be motivated to work even harder to protect their easy-to-lose status. With this understanding in mind, we use the following status proxies: (1) the size of an analyst's brokerage firm (*BrokerAssets*, measured as the previous year's total assets) and (2) the title of an analyst's position (*Title*). Analysts from a big brokerage firm earn higher compensation and working for a big brokerage house is deemed desirable in the analyst community (Hong and Kubik 2003). Once an analyst begins to work for a bigger brokerage house, he usually can stay with the firm for a long period of time. As for *Title*, carrying a higher title (e.g., department director, manager, president or chief analyst) is a clear indicator of status; and once a title is given, it typically stays.

We construct two subsamples, using the *Title* dummy and a median split based on *BrokerAssets*. We then run the same regression as reported in Table 3, Panel A for the two subsamples. Our empirical results are reported in Table 4. Consistent with our expectation, we find that $D(fWHR > median)$ is positively and significantly correlated with forecast accuracy only when analysts are classified as low-status analysts ($p < 0.01$). In addition, the coefficient on $D(fWHR > median)$ is more positive in the low-status subsample than in the high-status subsample. The differences in the coefficients from the two subsamples are significant at the 0.01 (0.05) level when *BrokerAssets* (*Title*) is the proxy for status.

4.3.2 Moderating effect of operating uncertainty

The noisy rational expectations models (Grossman and Stiglitz 1980; Diamond and Verrecchia 1981) posit that securities prices of firms with more uncertain operating environment may fail to fully reveal all private information, suggesting that information acquisition efforts are likely to

be more effective in improving analysts' performance. If the achievement drive hypothesis holds, the extra efforts exerted by high-*fWHR* analysts will lead to a greater difference in performance between the high- and low-*fWHR* analysts for firms with more uncertainty.

Following prior literature, the uncertainty proxies are (1) the operating cash flow volatility, i.e., the standard deviation of the operating cash flow deflated by total assets in the past five years (*CF_Volatility*), and (2) research and development expenditure (*R&D*). We partition our sample into firms with high and low uncertainty, using the *R&D* dummy and the median of *CF_Volatility*, respectively.

We find in Table 5 that the coefficient on $D(fWHR > median)$ is more positive in the high uncertainty subsample (i.e., the subsample in column 1 (column 3) consisting of firms with higher cash flow volatility (with non-zero research and development expenditure)). The difference in coefficients on $D(fWHR > median)$ from the two subsamples is significant at the 0.05 (0.10) level in columns 1 and 2 (columns 3 and 4). In addition, the positive association between $D(fWHR > median)$ and forecast accuracy is significant only in the high uncertainty subsample when we use *CF_Volatility* to separate the sample. Our findings lend support to the prediction that greater information acquisition efforts by high-*fWHR* analysts result in more dramatic performance advantage for these analysts when the covered firm has higher uncertainties.

Overall, results of the subsample analyses are consistent with the achievement drive hypothesis.

5. Tests related to the mechanisms linking facial width-to-height ratio to behaviors

5.1 Sexual dimorphism in facial width-to-height ratio

We test sexually dimorphic in *fWHR* in our sample of Chinese financial analysts, which consists of a large number of individuals with the same ethnicity and the same profession. Table 2, Panel A shows that the mean value of *fWHR* is 3.5% higher for males than for females ($p < 0.01$), suggesting the sexual dimorphism in *fWHR* (at least in our sample). Our finding is consistent with the *pubertal* testosterone-based mechanism for the *fWHR*-behavior link. It is important to note that we are unable to document that our finding is independent of the other biometrics (e.g., body size) due to the

lack of biometrics data in our sample. Therefore, our conclusions should be interpreted with this caveat in mind.

5.2 Testosterone, competition, and *fWHR*-performance association

This section examines the moderating effect of competition threats on the *fWHR*-performance association to provide further evidence that is consistent with the testosterone-based mechanism. As discussed in Section 2, the *pubertal* testosterone-based mechanism for the association between *fWHR* and achievement drive is that *fWHR* is indicative of pubertal testosterone exposure, which exerts organizational effects on the neural circuits associated with the ability to program activational responses to circulating testosterone that affects achievement drive later in life. Prior literature also provides evidence that the effect of testosterone on achievement-striving is more pronounced when individuals face threats (Derntl et al. 2009; Eisenegger et al. 2011; Goetz et al. 2014; Hermans et al. 2006; Hermans et al. 2008). Therefore, greater threats lead to a stronger link between *fWHR* and achievement drive, and consequently a stronger *fWHR*-performance association. If the positive association between an analyst's *fWHR* and his performance is due to testosterone, we predict that the association is stronger when the competition among analysts is more intense, because analysts are more likely to feel threatened when they face intense competition.

Competition among analysts is more intense when the number of analyst following increases (Frankel et al. 2006; O'Brien and Bhushan 1990). Therefore, our first competition measure is the number of analysts following the same firm (*NumAnalyst*), and we obtain two subsamples based on a median-split of this measure. In addition, Groysberg et al. (2011) find that the pay differentials among financial analysts are greater during periods of higher investor sentiment, incentivizing analysts to compete more intensively with their peers. We follow Fabozzi and Francis (1977), and identify years with high (low) sentiment as those in which Shanghai Composite Index return exceeds (falls behind) one-year risk-free rate.

Table 6 reports a significant and positive association between *fWHR* and forecast accuracy only when the competition among analysts is more intense. In addition, the coefficient on $D(fWHR > median)$ is more positive in the high-competition subsample than in the low-competition

subsample. The difference in coefficients from the two subsamples is significant at the 0.05 level when *NumAnalyst* is used to proxy for competition intensity. Overall, our findings are consistent with the testosterone-based hypothesis and suggest that the positive association between *fWHR* and performance is stronger when the competition among analysts is more intense.

6. Comprehensive analyses of dimensions of analyst performance and analyst forecast behaviors

6.1 Facial width-to-height ratio and dimensions of analyst performance

In this section, we consider dimensions of analyst performance other than forecast accuracy. Specifically, we examine whether analysts with higher *fWHR* are more likely to be selected as star analysts,²⁸ issue more profitable stock recommendations, are more likely to generate investment banking business for their employers, and issue bolder forecasts, forecasts with more consistent forecast errors, and timelier forecasts. For brevity, we present the summary of the results in Table 7. A full set of results are presented in Online Appendixes 1-9.

Similar to *Institutional Investor* All-star election in the U.S., star analysts in China are selected each year through surveys conducted by the magazine, *New Fortune*. Being selected as star analysts represents a remarkable career achievement and it substantially increases the analyst's reputation and bargaining power in the labor market. Our empirical analysis starts with the construction of an indicator variable (*Star*), which is coded as one if an analyst is ranked as a *New Fortune* star analyst during the year, and zero otherwise. We then estimate the time elapsed from the start of an analyst's career until he is selected as a star analyst (i.e., time until *Star*=1), using the Cox proportional hazard model. Since not all analysts are selected as the star analyst during their career, our data are right-censored. The dependent variable is the hazard rate, defined as the probability of an analyst experiencing the event (i.e., *Star*=1) in the current year, given that the analyst "has survived" without such an event until then. We include the same set of control variables as Leone and Wu (2007). The exponentiated coefficient on *D(fWHR>median)* is 1.427, significant at the 5% level. Our

²⁸ As discussed in Section 4.3.1, we are concerned about using star analysts to proxy for status because the election of star analysts is held annually and the star recognition can change every year.

findings indicate that the odds of being selected as star analysts are about 43% higher for high *fWHR* analysts than for low *fWHR* analysts.

Stock recommendation profitability is another important dimension of analysts' performance (Groysberg et al. 2011). We follow Wu and Zang (2009) and use the buy-and-hold market-adjusted returns to measure the stock recommendation profitability (i.e., *BHAR*[0,2] defined in the Appendix). The OLS regression in Table 7, column 2, relates an analyst's stock recommendation profitability to his *fWHR*. The positive and significant coefficient on *D(fWHR > median)* suggests that analysts with higher *fWHR* are better stock pickers than those with lower *fWHR* ($p < 0.05$).

An analyst's contribution to his bank's investment-banking operations can also be considered his performance measure (Groysberg et al. 2011). Similar to Groysberg et al. (2011), we calculate this performance measure as an indicator variable that equals one, if the analyst's bank underwrites an initial public offering or a seasoned equity offering of the firm covered by him during the current year, and zero otherwise (*Investment_Banking*). The odds ratio statistics in column 3 of Table 7 indicate that the odds of generating investment-banking business are 2.045 times higher for a high-*fWHR* analyst than for a low-*fWHR* analyst ($p < 0.01$).

Clement and Tse (2005) provide evidence suggesting that bold forecasts impound more relevant private information about earnings,²⁹ Hilary and Hsu (2013) find that earnings forecasts with more consistent forecast errors are more informative, while Cooper et al. (2001) suggest that forecast timeliness (measured by *LFR*, the leader-follower-ratio) reflects analysts' ability to influence the capital market. We therefore have three alternative dimensions of analysts' performance: forecast boldness, forecast consistency, and forecast timeliness. Consistent with the achievement drive hypothesis, columns 4-6 of Table 7 find that, compared to low-*fWHR* analysts, high-*fWHR* analysts are 13.2% more likely to issue bold forecasts, they issue forecasts with more consistent forecast errors and their forecasts are timelier.

6.2 Facial width-to-height ratio and other analyst forecast behaviors

²⁹ Since bold forecasts impound private information about earnings, we use forecast boldness to measure analyst performance rather than *innate* risk-taking.

In Section 6.1, we find that the positive association between an analyst's *fWHR* and his performance is robust to alternative analyst performance measures, which is consistent with the achievement drive hypothesis. In this section, we investigate whether *fWHR* also predicts other analyst forecast behaviors: forecast optimism, analyst overconfidence and forecast walkdowns.

Forecast optimism (*Optimism*) is an indicator variable, which equals one if the focal analyst's forecast is greater than the consensus forecast and zero otherwise. In constructing the variable of analyst overconfidence (*Overconfidence*), we follow Hilary and Menzly (2006) and form a measure capturing bold but inaccurate forecasts. In particular, *Overconfidence* is calculated as the absolute difference between the focal analyst's forecast and the consensus forecast (scaled into the range between 0 and 1 following Clement and Tse (2005)'s approach) if his forecast accuracy is lower than the sample median, and zero otherwise.³⁰ Columns 7-8 of Table 7 show that *Optimism* and *Overconfidence* are not significantly correlated with $D(fWHR > median)$, consistent with Jia et al. (2014)'s findings that *fWHR* measures a distinct personality trait different from optimism and overconfidence.

A forecast walkdown refers to the pattern of earnings forecasts within a fiscal year, which begins with optimistic forecasts followed by pessimistic forecasts when the earnings announcement date is close. Forecast walkdowns are shown to be attributed to analysts' incentives to curry favor with managers in order to obtain better access to management-provided information (Ke and Yu 2006; Bradshaw et al. 2016). We follow Bradshaw et al. (2016)'s approach to measure the walkdown pattern. If the *fWHR*–performance link is due to high-*fWHR* analysts currying favor with management, we expect a positive association between *fWHR* and the forecast walkdown pattern. However, the results reported in Table 7, column 9 are inconsistent with the expectation.

7. Robustness tests

7.1 Robustness tests with sample perturbations

³⁰ It is important to note that overconfidence results from attribution bias rather than achievement drive (Hilary and Menzly 2006). Hence, our overconfidence measure captures bold but inaccurate forecasts.

Table 8 reports a battery of robustness tests to examine whether our main result is robust to various perturbations to the sample. First, we randomly drop 1/3 of our sample and repeat this 1,000 times to form 1,000 samples. Then, we rerun the main regression (presented in Equation (3)) for each of the 1,000 samples. Panel A of Table 8 reports the distribution of the coefficients on $D(fWHR > median)$, their T-values, and their standard errors. The mean T-value suggests that in general, the coefficient on $D(fWHR > median)$ is statistically significant. Both the mean and median values for these coefficients are 0.014, and they are equal to the coefficient on $D(fWHR > median)$ in our main regression (reported in column 2 of Table 3, Panel A), suggesting that our main findings are unlikely driven by potential outliers.

To ensure that our main findings are not due to the false positive issue (Ohlson 2017), we use a permutation test to examine whether the positive association between $D(fWHR > median)$ and forecast accuracy disappears when we shuffle the $D(fWHR > median)$ measure randomly in the sample. We first randomly select 50% of the analysts from the sample. We code $D(fWHR > median)$ as one for the selected analysts, and zero for the remaining analysts. We shuffle $D(fWHR > median)$ 1,000 times, resulting in 1,000 pseudo samples with permuted $D(fWHR > median)$. We then rerun the main regression for each of the 1,000 pseudo samples, and report the distribution of the summary statistic from these regressions in Panel B of Table 8. We find that the mean and median values for the coefficients on $D(fWHR > median)$ from these 1,000 regressions are close to zero, implying that the probability of obtaining a positive (negative) coefficient is about 50%. The T-value distribution shows that these coefficients are, in general, statistically insignificant at the 10% level. Panel C (Panel D) presents the statistics distribution for the regressions with a positive (negative) coefficient on $D(fWHR > median)$. The mean T-value suggests that the coefficient estimates are economically and statistically insignificant in these regressions as well. Collectively, our evidence suggests a low likelihood of false positives in our main findings.

Finally, we test whether our main finding is more pronounced when we reduce the measurement error of $D(fWHR > median)$ by eliminating observations with $fWHR$ surrounding the

sample median.³¹ We therefore drop observations within a 5%, 10%, and 15% band around the median of $fWHR$ and rerun the main regression, respectively. Panel E of Table 8 reports that the coefficient estimates on $D(fWHR > median)$ are all positive and significant at the 1% level. The effect size increases by about 20% when we drop observations within a 10% band, and then remains at the higher level when we eliminate the observations within a 15% band. Our results suggest that some measurement errors exist for the observations within a 10% band around the median of $fWHR$.

7.2 Survivorship and backfilling bias

Our sample is limited to the analysts employed in 2015 and we gather the forecast data for these analysts from 2004 to 2014. Therefore, we might suffer from the survivorship (or backfilling) bias. We address this concern in two ways, following Jia et al. (2014). First, we repeat the main regression presented in Equation (3), using the 2014 data only. This ensures that each analyst enters the sample only once. Untabulated findings reveal that the positive and significant association between $D(fWHR > median)$ and forecast accuracy continues to hold. Second, we run a Kolmogorov-Smirnov test, comparing the low- $fWHR$ and high- $fWHR$ groups with respect to the distribution of the total number of observations for each analyst in the sample. We find no significant differences in the distributions between the two groups (corrected p -value = 0.403), indicating that our sample is not influenced disproportionately by unobservable factors.

7.3 Association between facial width-to-height ratio and achievement drive in females

To ensure that the $fWHR$ -behavior mechanism examined by prior studies works similarly in our Asia setting, we test the association between $fWHR$ and performance (corporate site visit) for female analysts only in Panel A (Panel B) of Table 9. If a similar mechanism works in our sample, we expect to find results consistent with prior literature, i.e., the relationship between $fWHR$ and performance (corporate site visit) is not significant for female analysts.

Panel A shows that in the univariate regression, the coefficient on $D(fWHR > median)$ is 0.016, while it is 0.004, small in terms of magnitude, in the regression where we include control variables.

³¹ We expect some measurement errors when $fWHR$ is close to the sample median because the misclassification of high- $fWHR$ (low- $fWHR$) analysts is more likely to happen in this interval.

The coefficient is not significant at the 10% level in both regressions. Panel B indicates that the coefficient is 0.461 and 0.423 respectively in the univariate and multivariate regressions, and it is not significant at the 10% level in both cases. The odds ratio statistics are however larger than those that we obtain when analyzing male analysts.

Although results reported in Table 9 show a lack of significant results when we analyze female analysts, we can't categorically rule out the alternative explanation that the insignificant results are due to that the reduced number of observations in our sample of female analysts lowers the power of our statistical test.

8. Conclusions

Prior literature has offered robust evidence that men with high-*fWHR* are more likely to exhibit masculine behaviors (Carré and McCormick 2008; Haselhuhn and Wong 2012; Haselhuhn et al. 2014; Jia et al. 2014; Stirrat and Perrett 2010). Recent advances in biology pinpoint achievement drive as the main trait predicted by *fWHR*, which manifests in *calculated* masculine behaviors, such as aggression, deception and risk-taking (Lefevre et al. 2014; Lewis et al. 2012; Welker et al. 2015). This marks a new milestone in our understanding of the association between facial structure and human behaviors because it implies that the positive link between *fWHR* and masculine behaviors will break, when such behaviors obstruct achievement attainment.

The central prediction of the achievement drive hypothesis is that individuals with higher *fWHR* are more motivated in seeking achievement, they put more efforts in their jobs, and consequently, they exhibit better performance.

We use a large pictorial sample of Chinese financial analysts to test this prediction. This setting offers a large number of observations of individuals who are homogeneous in ethnicity and profession, and whose individual performance can be measured directly and objectively. Furthermore, because aggressive behaviors are socially undesirable in East Asia (Re and Rule 2016; Rule et al. 2010; Rule et al. 2011) and corporate site visit data are available in China, this setting allows us to differentiate from the *autonomic arousal* explanation.

We find that high-*fWHR* analysts are more likely to conduct corporate site visits, and they perform better in multiple dimensions. Specifically, they issue more accurate earnings forecasts, they are about 43 percent more likely to be selected as star analysts, their stock recommendations are more profitable, they are about two times more likely to bring investment banking deals to their employers, they are 13.2% more likely to issue bold forecasts, their forecast errors are more consistent, and their forecasts are more timely. There is no evidence that they are optimistic (and/or overconfident) or their superior performance is due to their currying favor with the management (Ke and Yu 2006; Bradshaw et al. 2016). Our main result is robust to various perturbations to the sample.

Our further investigations show that the positive association between *fWHR* and performance is more pronounced for analysts with lower status and for firms with higher uncertainty. These findings buttress the notion that the association is likely due to the achievement drive hypothesis. We also find that the association is more pronounced for analysts facing more intense competition. Together with the evidence supportive of sexual dimorphism in *fWHR*, this finding highlights the important role played by testosterone, especially pubertal testosterone, in explaining the *fWHR*-behavior link.

Overall, the preponderance of our results lends support to the achievement drive hypothesis and suggests that pubertal hormone shapes the association between facial structure and individual behaviors. However, our conclusion is subject to two important caveats. First, since the trait of achievement drive reflects complex phenotype, it is possible that more than one hormone is at play. Second, conceptually, we can't completely rule out the possibility that our finding is due to *fWHR* being a proxy for constructs other than achievement drive, such as perceived competence, sociability, morality and attractiveness. We however take comfort in that existing evidence lends no support to the alternative explanations. Specifically, prior research either shows no association between *fWHR* and these perceived traits (Carré et al. 2009; Geniole et al. 2012; Ormiston et al. 2017) or documents an association that is logically incapable of explaining the positive correlation between *fWHR* and performance (Alrajih and Ward 2014; Carré, et al. 2009; Carré et al. 2010; Geniole et al. 2014; Geniole et al. 2015; Geniole and McCormick 2015; Haselhuhn et al. 2013; Lieberz et al. 2017; Mileva

et al. 2014; Ormiston et al. 2017; Short et al. 2012; Stirrat and Perrett 2010).³² These two limitations of our study, in our view, represent potential opportunities for future research.

³² What's more, conceptually, facial features are unlikely to play an important role in the equilibrium-level perceptions of competence, sociability and morality because true personal traits will be revealed ultimately after multi-period interactions.

References:

- Addoum, J. M., G. Korniotis, and A. Kumar. 2017. Stature , Obesity , and Portfolio Choice. *Management Science* 63 (10): 3393–3413.
- Alrajih, S., and J. Ward. 2014. Increased facial width-to-height ratio and perceived dominance in the faces of the UK's leading business leaders. *British Journal of Psychology* 105: 153–161.
- Angrist, J. D., and J.-S. Pischke. 2009. *Mostly Harmless Econometrics. An Empiricist's Companion. An empiricist's companion*. Princeton: Princeton University Press.
- Azurmendi, A., F. Braza, A. Sorozaal, A. García, P. Braza, M. R. Carreras, J. M. Muñoz, J. Cardas, and J. R. Sánchez-Martín. 2005. Cognitive abilities, androgen levels, and body mass index in 5-year-old children. *Hormones and Behavior* 48 (2): 187–195.
- Bird, B. M., V. S. Cid Jofré, S. N. Geniole, K. M. Welker, S. Zilioli, D. Maestripieri, S. Arnocky, and J. M. Carré. 2016. Does the facial width-to-height ratio map onto variability in men's testosterone concentrations? *Evolution and Human Behavior* 37 (5): 392–398.
- Boksem, M. a S., P. H. Mehta, B. Van den Bergh, V. van Son, S. T. Trautmann, K. Roelofs, A. Smidts, and A. G. Sanfey. 2013. Testosterone Inhibits Trust But Promotes Reciprocity. *Psychological Science* 24 (11): 2306–2314.
- Bradshaw, M. T., L. F. Lee, and K. Peterson. 2016. The interactive role of difficulty and incentives in explaining the annual earnings forecast walkdown. *Accounting Review* 91 (4): 995–1021.
- Carré, J. M., and C. M. McCormick. 2008. In Your Face: Facial Metrics Predict Aggressive Behaviour in the Laboratory and in Varsity and Professional Hockey Players. *Proceedings of the Royal Society B: Biological Sciences* 275: 2651–2656.
- Carré, J. M., C. M. McCormick, and C. J. Mondloch. 2009. Facial Structure is a Reliable Cue of Aggressive Behavior. *Psychological Science* 20 (10): 1194–1198.
- Carré, J. M., M. D. Morrissey, C. J. Mondloch, and C. M. McCormick. 2010. Estimating aggression from emotionally neutral faces: Which facial cues are diagnostic? *Perception* 39 (3): 356–377.
- Carré, J. M., K. R. Murphy, and A. R. Hariri. 2013. What Lies Beneath the Face of Aggression? *Social Cognitive and Affective Neuroscience* 8 (2): 224–229.
- Carré, J. M., S. K. Putnam, and C. M. McCormick. 2009. Testosterone Responses to Competition

- Predict Future Aggressive Behaviour at a Cost to Reward in Men. *Psychoneuroendocrinology* 34 (4): 561–570.
- Cheng, Q., F. Du, X. Wang, and Y. Wang. 2016. Seeing is Believing: Analysts' Corporate Site Visits. *Review of Accounting Studies* 21: 1245–1286.
- Clement, M. B. 1999. Analyst Forecast Accuracy: Do Ability, Resources, and Portfolio Complexity Matter? *Journal of Accounting and Economics* 27 (3): 285–303.
- Clement, M. B., L. Koonce, and T. J. Lopez. 2007. The Roles of Task-specific Forecasting Experience and Innate Ability in Understanding Analyst Forecasting Performance. *Journal of Accounting and Economics* 44 (3): 378–398.
- Clement, M. B., and K. K. F. Law. 2014. Recession Analysts and Conservative Forecasting. SSRN eLibrary.
- Clement, M. B., and S. Y. Tse. 2005. Financial Analyst Characteristics and Herding Behavior in Forecasting. *The Journal of Finance* 60 (1): 307–341.
- Cohen, L., A. Frazzini, and C. Malloy. 2010. Sell-side School Ties. *The Journal of Finance* 65 (4): 1409–1437.
- Cooper, R. A., T. E. Day, and C. M. Lewis. 2001. Following the leader: A study of individual analysts' earnings forecasts. *Journal of Financial Economics* 61 (3): 383–416.
- Derntl, B., C. Windischberger, S. Robinson, I. Kryspin-Exner, R. C. . Gur, E. Moser, and U. Habel. 2009. Amygdala Activity to Fear and Anger in Healthy Young Males is Associated with Testosterone. *Psychoneuroendocrinology* 34: 687–693.
- Diamond, D. W., and R. E. Verrecchia. 1981. Information Aggregation in a Noisy Rational Expectations Economy. *Journal of Financial Economics* 9 (3): 221–235.
- Eckel, C. C. 2007. People Playing Games: The Human Face of Experimental Economics. *Southern Economic Journal* 73 (4): 840–857.
- Eisenegger, C., J. Haushofer, and E. Fehr. 2011. The Role of Testosterone in Social Interaction. *Trends in Cognitive Sciences* 15 (6): 263–271.
- Eisenegger, C., M. Naef, R. Snozzi, M. Heinrichs, and E. Fehr. 2010. Prejudice and Truth About the Effect of Testosterone on Human Bargaining Behaviour. *Nature* 463 (7279): 356–359.

- Fabozzi, F. J., and J. C. Francis. 1977. Stability Tests for Alphas and Betas Over Bull and Bear Market Conditions. *The Journal of Finance* 32 (4): 1093–1099.
- Fisk, S. R., B. J. Miller, and J. Overton. 2017. Why social status matters for understanding the interrelationships between testosterone, economic risk-taking, and gender. *Sociology Compass* 11 (3): 1–18.
- De Franco, G., and Y. Zhou. 2009. The Performance of Analysts with a CFA® Designation: The Role of Human-capital and Signaling Theories. *The Accounting Review* 84 (2): 383–404.
- Frankel, R., S. P. Kothari, and J. Weber. 2006. Determinants of the Informativeness of Analyst Research. *Journal of Accounting and Economics* 41 (1–2): 29–54.
- Geniole, S. N., T. F. Denson, B. J. Dixon, J. M. Carré, and C. M. McCormick. 2015. Evidence from meta-analyses of the facial width-to-height ratio as an evolved cue of threat. *PLoS ONE* 10 (7): 1–18.
- Geniole, S. N., A. E. Keyes, J. M. Carré, and C. M. McCormick. 2014. Fearless Dominance Mediates the Relationship between the Facial Width-to-height Ratio and Willingness to Cheat. *Personality and Individual Differences* 57: 59–64.
- Geniole, S. N., A. E. Keyes, C. J. Mondloch, J. M. Carré, and C. M. McCormick. 2012. Facing aggression: Cues differ for female versus male faces. *PLoS ONE* 7 (1): 1–12.
- Geniole, S. N., and C. M. McCormick. 2015. Facing our ancestors: Judgements of aggression are consistent and related to the facial width-to-height ratio in men irrespective of beards. *Evolution and Human Behavior* 36 (4): 279–285.
- Geniole, S. N., D. S. Molnar, J. M. Carré, and C. M. McCormick. 2014. The facial width-to-height ratio shares stronger links with judgments of aggression than with judgments of trustworthiness. *Journal of Experimental Psychology: Human Perception and Performance* 40 (4): 1526–1541.
- Goetz, S. M. M., K. S. Shattuck, R. M. Miller, J. A. Campbell, E. Lozoya, G. E. Weisfeld, and J. M. Carré. 2013. Social Status Moderates the Relationship Between Facial Structure and Aggression. *Psychological science* 24 (11): 2329–34.
- Goetz, S. M. M., L. Tang, M. E. Thomason, M. P. Diamond, A. R. Hariri, and J. M. Carré. 2014. Testosterone Rapidly Increases Neural Reactivity to Threat in Healthy Men: A Novel Two-step

- Pharmacological Challenge Paradigm. *Biological Psychiatry* 76 (4): 324–331.
- Greene, W. W. H. . 2012. *Econometric Analysis*. Prentice Hall.
- Grossman, S. J., and J. E. Stiglitz. 1980. On the Impossibility of Informationally Efficient Markets. *The American Economic Review* 70 (3): 393–408.
- Groysberg, B., P. M. Healy, and D. A. Maber. 2011. What Drives Sell-Side Analyst Compensation at High-Status Investment Banks? *Journal of Accounting Research* 49 (4): 969–1000.
- Haselhuhn, M. P., and E. M. Wong. 2012. Bad to the Bone: Facial Structure Predicts Unethical Behaviour. *Proceedings of the Royal Society B: Biological Sciences* 279 (1728): 571–576.
- Haselhuhn, M. P., E. M. Wong, and M. E. Ormiston. 2013. Self-Fulfilling Prophecies as a Link between Men’s Facial Width-to-Height Ratio and Behavior. *PLoS ONE* 8 (8): doi:10.1371/journal.pone.0072259.
- Haselhuhn, M. P., E. M. Wong, M. E. Ormiston, M. E. Inesi, and A. D. Galinsky. 2014. Negotiating Face-to-face: Men’s Facial Structure Predicts Negotiation Performance. *Leadership Quarterly* 25 (5): 835–845.
- Hermans, E. J., P. Putman, J. M. Baas, H. P. Koppeschaar, and J. Van Honk. 2006. A Single Administration of Testosterone Reduces Fear-Potentiated Startle in Humans. *Biological Psychiatry* 59: 872–874.
- Hermans, E. J., N. F. Ramsey, and J. Van Honk. 2008. Exogenous Testosterone Enhances Responsiveness to Social Threat in the Neural Circuitry of Social Aggression in Humans. *Biological Psychiatry* 63: 263–270.
- Hicks, R., and D. Tingley. 2011. Causal mediation analysis. *Stata Journal* 11 (4): 605–619.
- Hilary, G., and C. Hsu. 2013. Analyst Forecast Consistency. *Journal of Finance* 68 (1): 271–297.
- Hilary, G., and L. Menzly. 2006. Does Past Success Lead Analysts to Become Overconfident? *Management Science* 52 (4): 489–500.
- Hobson, J. L., W. J. Mayew, and M. Venkatachalam. 2012. Analyzing Speech to Detect Financial Misreporting. *Journal of Accounting Research* 50 (2): 349–392.
- Hong, H., and J. D. Kubik. 2003. Analyzing the Analysts: Career Concerns and Biased Earnings Forecasts. *The Journal of Finance* 58 (1): 313–351.

- van Honk, J., D. Terburg, and P. A. Bos. 2011. Further Notes on Testosterone as a Social Hormone. *Trends in Cognitive Sciences* 15 (7): 291–292.
- van Honk, J., G. Will, D. Terburg, W. Raub, C. Eisenegger, and V. Buskens. 2016. Effects of Testosterone on Strategic Gambling in Poker Play. *Scientific Reports*: doi: 10.1038/srep18096.
- Jacob, J., T. Z. Lys, and M. A. Neale. 1999. Expertise in Forecasting Performance of Security Analysts. *Journal of Accounting and Economics* 28 (1): 51–82.
- Jia, Y., L. Van Lent, and Y. Zeng. 2014. Masculinity, Testosterone, and Financial Misreporting. *Journal of Accounting Research* 52 (5): 1195–1246.
- Jiang, D., A. Kumar, and K. K. F. Law. 2016. Political Contributions and Analyst Behavior. *Review of Accounting Studies* 21 (1): 37–88.
- Kamiya, S., Y. H. Kim, and S. Park. 2018. The Face of Risk: CEO Facial Masculinity and Firm Risk. *European Financial Management, Forthcoming*.
- Ke, B., and Y. Yu. 2006. The effect of issuing biased earnings forecasts on analysts' access to management and survival. *Journal of Accounting Research* 44 (5): 965–999.
- Kramer, R. S. S. 2017. Sexual dimorphism of facial width-to-height ratio in human skulls and faces: A meta-analytical approach. *Evolution and Human Behavior* 38 (3): 414–420.
- Kramer, R. S. S., A. L. Jones, and R. Ward. 2012. A lack of sexual dimorphism in width-to-height ratio in white european faces using 2D photographs, 3D scans, and anthropometry. *PLoS ONE* 7 (8): e42705.
- Kumar, A. 2010. Self-selection and the Forecasting Abilities of Female Equity Analysts. *Journal of Accounting Research* 48 (2): 393–435.
- Lebuda, I., and M. Karwowski. 2016. Written on the Writer's Face: Facial Width-to-Height Ratio among Nominees and Laureates of the Nobel Prize in Literature. *Creativity Research Journal* 28 (2): 207–211.
- Lefevre, C. E., P. J. Etchells, E. C. Howell, A. P. Clark, and I. S. Penton-Voak. 2014. Facial Width-to-height Ratio Predicts Self-reported Dominance and Aggression in Males and Females, But a Measure of Masculinity Does Not. *Biology Letters* 10 (1): 20140729–20140729.
- Lefevre, C. E., and G. J. Lewis. 2014. Perceiving Aggression from Facial Structure: Further Evidence

- for a Positive Association with Facial Width-to-Height Ratio and Masculinity, but not for Moderation by Self-Reported Dominance. *European Journal of Personality* 28 (6): 530–537.
- Lefevre, C. E., G. J. Lewis, T. C. Bates, M. Dzhelyova, V. Coetzee, I. J. Deary, and D. I. Perrett. 2012. No Evidence for Sexual Dimorphism of Facial Width-to-height Ratio in Four Large Adult Samples. *Evolution and Human Behavior* 33 (6): 623–627.
- Lefevre, C. E., G. J. Lewis, D. I. Perrett, and L. Penke. 2013. Telling Facial Metrics: Facial Width is Associated with Testosterone Levels in Men. *Evolution and Human Behavior* 34 (4): 273–279.
- Leone, A., and J. S. Wu. 2007. What Does It Take to Become a Superstar? Evidence from Institutional Investor Rankings of Financial Analysts. *Working Paper*. Penn State University and University of Rochester. SSRN eLibrary.
- Lewis, G. J., C. E. Lefevre, and T. C. Bates. 2012. Facial Width-to-height Ratio Predicts Achievement Drive in US Presidents. *Personality and Individual Differences* 52 (7): 855–857.
- Lieberz, K. A., S. Windmann, S. N. Geniole, C. M. McCormick, M. Mueller-Engelmann, F. Gruener, P. Bornefeld-Ettmann, and R. Steil. 2017. The facial width-to-height ratio determines interpersonal distance preferences in the observer. *Aggressive Behavior* 43 (5): 460–470.
- Mayew, W. J. 2013. Reassessing the association between facial structure and baseball performance: a comment on Tsujimura & Banissy. *Biology letters* 9 (5): 20130538.
- Mayew, W. J., C. A. Parsons, and M. Venkatachalam. 2013. Voice pitch and the labor market success of male chief executive officers. *Evolution and Human Behavior* 34 (4): 243–248.
- Mikhail, M. B., B. R. Walther, and R. H. Willis. 1999. Does Forecast Accuracy Matter to Security Analysts? *The Accounting Review* 74 (2): 185–200.
- Mileva, V. R., M. L. Cowan, K. D. Cobey, K. K. Knowles, and a. C. Little. 2014. In the face of dominance: Self-perceived and other-perceived dominance are positively associated with facial-width-to-height ratio in men. *Personality and Individual Differences* 69: 115–118.
- O'Brien, P. C., and R. Bhushan. 1990. Analyst Following and Institutional Ownership. *Journal of Accounting Research* 28: 55–76.
- Ohlson, J. A. 2017. Researchers' Data Analysis Choices: An Excess of False Positives? SSRN eLibrary.

- Ormiston, M. E., E. M. Wong, and M. P. Haselhuhn. 2017. Facial-width-to-height ratio predicts perceptions of integrity in males. *Personality and Individual Differences* 105: 40–42.
- Özener, B. 2012. Facial Width-to-height Ratio in a Turkish Population is Not Sexually Dimorphic and is Unrelated to Aggressive Behavior. *Evolution and Human Behavior* 33 (3): 169–173.
- Petersen, M. A. 2009. Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. *Review of Financial Studies* 22 (1): 435–480.
- Ramnath, S., S. Rock, and P. Shane. 2008. The Financial Analyst Forecasting Literature: A Taxonomy with Suggestions for Further Research. *International Journal of Forecasting* 24 (1): 34–75.
- Rasband, W. 2015. ImageJ. U. S. National Institutes of Health, Bethesda, Maryland, USA: [//imagej.nih.gov/ij/](http://imagej.nih.gov/ij/).
- Re, D. E., and N. O. Rule. 2016. Predicting Firm Success From the Facial Appearance of Chief Executive Officers of Non-Profit Organizations. *Perception* 45 (10): 1137–1150.
- Rule, N. O., N. Ambady, R. B. Adams, H. Ozone, S. Nakashimi, S. Yoshikawa, and M. Watabe. 2010. Polling the face: Prediction and consensus across cultures. *Journal of Personality and Social Psychology* 98 (1): 1–15.
- Rule, N. O., K. Ishii, and N. Ambady. 2011. Cross-cultural impressions of leaders' faces: Consensus and predictive validity. *International Journal of Intercultural Relations* 35 (6): 833–841.
- Sanchez-Pages, S., C. Rodriguez-Ruiz, and E. Turiegano. 2014. Facial masculinity: How the choice of measurement method enables to detect its influence on behaviour. *PLoS ONE* 9 (11) : e112157.
- Short, L. a., C. J. Mondloch, C. M. McCormick, J. M. Carré, R. Ma, G. Fu, and K. Lee. 2012. Detection of Propensity for Aggression Based on Facial Structure Irrespective of Face Race. *Evolution and Human Behavior* 33 (2): 121–129.
- Simonton, D. K. 1986. Presidential personality: Biographical use of the Gough Adjective Check List. *Journal of Personality and Social Psychology* 51 (1): 149–160.
- Sisk, C. L., K. M. Schulz, and J. L. Zehr. 2003. Puberty: A Finishing School for Male Social Behavior. *Annals of the New York Academy of Sciences* 1007: 189–198.
- Sisk, C. L., and J. L. Zehr. 2005. Pubertal hormones organize the adolescent brain and behavior.

- Frontiers in Neuroendocrinology* 26 (3–4): 163–174.
- Smeets-Janssen, M. M. J., K. Roelofs, J. Van Pelt, P. Spinhoven, F. G. Zitman, B. W. J. H. Penninx, and E. J. Giltay. 2015. Salivary Testosterone is Consistently and Positively Associated with Extraversion: Results from the Netherlands Study of Depression and Anxiety. *Neuropsychobiology* 71 (2): 76–84.
- Stickel, S. E. 1992. Reputation and Performance Among Security Analysts. *The Journal of Finance* 47 (5): 1811–1836.
- Stirrat, M., and D. I. Perrett. 2010. Valid Facial Cues to Cooperation and Trust: Male Facial Width and Trustworthiness. *Psychological science* 21 (3): 349–354.
- Stirrat, M., and D. I. Perrett. 2012. Face Structure Predicts Cooperation: Men With Wider Faces Are More Generous to Their In-Group When Out-Group Competition Is Salient. *Psychological Science* 23 (7): 718–722.
- Stirrat, M., G. Stulp, and T. V. Pollet. 2012. Male facial width is associated with death by contact violence: Narrow-faced males are more likely to die from contact violence. *Evolution and Human Behavior* 33 (5): 551–556.
- Třebický, V., J. Fialová, K. Kleisner, S. C. Roberts, A. C. Little, and J. Havlíček. 2015. Further Evidence for Links between Facial Width-to-height Ratio and Fighting Success: Commentary on Zilioli et al. (2014). *Aggressive behavior* 41: 331–334.
- Verdonck, a., M. Gaethofs, C. Carels, and F. De Zegher. 1999. Effect of Low-dose Testosterone Treatment on Craniofacial Growth in Boys with Delayed Puberty. *European Journal of Orthodontics* 21 (2): 137–143.
- Welker, K. M., S. M. M. Goetz, and J. M. Carré. 2015. Perceived and Experimentally Manipulated Status Moderates the Relationship between Facial Structure and Risk-taking. *Evolution and Human Behavior* 36 (6): 423–429.
- Weston, E. M., A. E. Friday, and P. Liò. 2007. Biometric Evidence that Sexual Selection has Shaped the Hominin Face. *PLoS ONE* 2 (8): doi:10.1371/journal.pone.0000710.
- Wong, E. M., M. E. Ormiston, and M. P. Haselhuhn. 2011. A Face Only an Investor Could Love: CEOs' Facial Structure Predicts Their Firms' Financial Performance. *Psychological Science* 22

(12): 1478–1483.

Wu, J. S., and A. Y. Zang. 2009. What Determine Financial Analysts' Career Outcomes During Mergers? *Journal of Accounting and Economics* 47: 59–86.

Yin, H., and H. Zhang. 2014. Tournaments of financial analysts. *Review of Accounting Studies* 19 (2): 573–605.

Zilioli, S., A. N. Sell, M. Stirrat, J. Jagore, W. Vickerman, and N. V. Watson. 2015. Face of a Fighter: Bizygomatic Width as a Cue of Formidability. *Aggressive Behavior* 41: 322–330.

Figure 1

Autonomic arousal hypothesis:

Achievement drive hypothesis:

Figure 2: Direct and indirect effects of $fWHR$ on forecast accuracy

Path a is the direct effect of $D(fWHR > median)$ on $Accuracy$.

Path b is the direct effect of $D(fWHR > median)$ on $SiteVisit$.

Path c is the direct effect of $SiteVisit$ on $Accuracy$.

Path $b \times$ Path c is the indirect effect of $D(fWHR > median)$ on $Accuracy$ via $SiteVisit$.

Table 1 Sample Selection

This table reports the sample selection procedure. We obtain the initial analyst-level pictorial sample through the website of Securities Association of China (SAC) in July 2015 (N=2,108). We then collect those analysts' annual earnings forecasts between 2004 and 2014 from CBAS (Chinese Brokerage and Analyst System). The final analyst-level sample comprises 1,193 *male* analysts with good quality pictures. The final sample at the analyst-firm-year level comprises 19,324 observations for the analyses.

Sampling Procedure	Number of analysts	Number of forecasts (2004-2014)
Initial analyst-level sample with data available in SAC website and CBAS in July 2015	2,108	
Less: Low quality pictures for males	249	
Low quality pictures for females	115	
Female analysts	551	
Male analysts with good quality pictures	1,193	296,244
Less: Forecasts issued by a group of analysts		159,148
Forecasts issued less than 30 days before earnings announcement		3,264
Multiple forecasts per year except the most recent forecasts		77,901
Forecasts of firms from the financial industry		2,069
Missing data for variables required for the regressions		34,538
Final sample		19,324

Table 2 Summary statistics and sexual dimorphism in *fWHR*

Panel A of this table reports the analysts' facial width-to-height ratio (*fWHR*) across the analyst-level sample for both male and female analysts. Panel B presents descriptive statistics for the analyst-level characteristics across the analyst-firm-year level sample (N=19,324). The variables of analyst characteristics, including the forecast accuracy, are scaled to range from 0 to 1 as described in Section 3.2. Appendix provides a detailed description of variable construction.

Panel A: Sexual dimorphism in *fWHR*

<i>fWHR</i>	N	Mean	Median	Q1	Q3	SD	T-test	Z-test
<i>Male</i>	1,193	2.166	2.156	2.057	2.261	0.181	0.074***	0.066***
<i>Female</i>	551	2.092	2.090	1.997	2.173	0.130	(8.571)	(9.495)
<i>Full Sample</i>	1,744	2.143	2.135	2.038	2.236	0.170		

Panel B: Descriptive for the analyst-level characteristics

	N	Mean	Median	Q1	Q3	SD
<i>Accuracy</i>	19,324	0.677	0.788	0.464	0.948	0.315
<i>LagAccuracy</i>	19,324	0.603	0.693	0.319	0.926	0.348
<i>Dayselap</i>	19,324	0.267	0.167	0.012	0.417	0.295
<i>Frequency</i>	19,324	0.300	0.000	0.000	0.500	0.368
<i>Horizon</i>	19,324	0.308	0.228	0.089	0.505	0.267
<i>Brokersize</i>	19,324	0.431	0.422	0.268	0.567	0.235
<i>Experience</i>	19,324	0.398	0.333	0.000	0.667	0.347
<i>Companies</i>	19,324	0.280	0.168	0.077	0.373	0.286
<i>Industries</i>	19,324	0.304	0.200	0.097	0.412	0.290
<i>AnalystAge</i>	19,324	0.421	0.364	0.176	0.619	0.303
<i>Postgraduate</i>	19,324	0.877	1.000	1.000	1.000	0.328

Table 3 Analysts' *fWHR*, forecast accuracy, and corporate site visits

Panel A reports results for OLS regressions relating an analyst's forecast accuracy to his *fWHR* (measured by $D(fWHR > median)$). The dependent variable is the scaled forecast accuracy measure based on Clement and Tse (2005). Panel B reports results for logistic regressions relating an analyst's corporate site visit probability to his *fWHR*. The number of observations is reduced to 12,081 analyst-firm-years, because only firms listed on Shenzhen Stock Exchange are required to report analysts' corporate site visit records. The dependent variable (*SiteVisit*) equals one if the analyst conducts at least one corporate site visit to the firm in the current year, and zero otherwise. Panel C reports the mediation effect of corporate site visit on the relationship between an analyst's *fWHR* and his forecast accuracy. As indicated in Figure 2, path a (direct effect) is the standardized regression coefficient of $D(fWHR > median)$ on forecast accuracy; path b is the standardized regression coefficient of $D(fWHR > median)$ on corporate site visit; path c is the standardized regression coefficient of corporate site visit on forecast accuracy. Appendix provides a detailed description of the construction of all the variables. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Panel A: OLS regressions relating an analyst's forecast accuracy to his *fWHR*

Variables	Accuracy (1)	Accuracy (2)
<i>D(fWHR > median)</i>	0.032*** (0.013)	0.014*** (0.004)
<i>LagAccuracy</i>		0.435*** (0.009)
<i>Dayselap</i>		0.002 (0.006)
<i>Frequency</i>		-0.023*** (0.005)
<i>Horizon</i>		-0.454*** (0.011)
<i>Brokersize</i>		-0.002 (0.004)
<i>Experience</i>		0.014** (0.006)
<i>Companies</i>		0.026** (0.012)
<i>Industries</i>		-0.014 (0.013)
<i>AnalystAge</i>		-0.010 (0.008)
<i>Postgraduate</i>		0.007 (0.006)
<i>Constant</i>	0.661*** (0.011)	0.544*** (0.012)
<i>N</i>	19,324	19,324
<i>Adj R²</i>	0.003	0.544

Table 3 *continued***Panel B: Logistic regressions relating an analyst's corporate site visit probability to his *fWHR***

Variables	<i>SiteVisit</i>		<i>SiteVisit</i>	
	(1)		(2)	
	Coefficient	Odds Ratio	Coefficient	Odds Ratio
<i>D(fWHR>median)</i>	0.136*** (0.048)	1.146	0.117** (0.051)	1.124
<i>Distance</i>			-0.016 (0.030)	0.984
<i>Brokern</i>			-0.022 (0.038)	0.979
<i>Tangidum</i>			-0.043 (0.067)	0.958
<i>HHIdum</i>			0.088 (0.106)	1.092
<i>Strongbuy</i>			0.550*** (0.047)	1.733
<i>Drating</i>			0.414*** (0.138)	1.512
<i>Coverage</i>			-0.019*** (0.006)	0.981
<i>SOE</i>			-0.160* (0.096)	0.852
<i>MV</i>			-0.210*** (0.053)	0.811
<i>ROA</i>			0.054 (0.809)	1.056
<i>FirmAge</i>			-0.041 (0.062)	0.960
<i>BM</i>			-0.152 (0.165)	0.859
<i>Leverage</i>			0.037 (0.224)	1.038
<i>AnalystAge</i>			-0.707*** (0.089)	0.493
<i>Postgraduate</i>			0.122 (0.095)	1.130
<i>Constant</i>	-2.117*** (0.212)	0.121	1.164 (0.811)	3.202
Industry Fixed effects	<i>Included</i>		<i>Included</i>	
<i>N</i>	12,081		12,081	
<i>Pseudo R²</i>	0.030		0.132	

Table 3 *continued***Panel C: Mediation effect of corporate site visit on the relationship between an analyst's *fWHR* and his forecast accuracy**

Effect	Mean	[95% Conf. Interval]	
Direct effect (path a, Figure 2)	0.0144		
Indirect effect (path b $\times$ path c, Figure 2)	0.0013		
Total effect (direct + indirect effects)	0.0157		
% of total effect mediated	8.2802%	4.8294%	23.5447%

Table 4 Analysts' *fWHR* and forecast accuracy—sample split by the level of status

This table reports results from OLS regressions relating an analyst's forecast accuracy to his *fWHR* (measured by $D(fWHR > median)$) for subsamples formed based on the status measure. The dependent variable is the scaled forecast accuracy measure based on Clement and Tse (2005). To measure status, columns 1 and 2 use brokerage size (*BrokerAssets*), and columns 3 and 4 use the title of an analyst's position (*Title*). We use the sample median of *BrokerAssets* (the *Title* dummy indicating an analyst carrying a higher title) to separate the sample in columns 1 and 2 (columns 3 and 4). Appendix provides a detailed description of the construction of all the variables. The number of observations reduced to 18,176 analyst-firm-years for the sample in columns 1 and 2 because we further require total assets of the brokerage firm employing the analyst to be non-missing. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Variables	Accuracy	Accuracy	Accuracy	Accuracy
	<i>BrokerAssets</i>		<i>Title</i>	
	High (1)	Low (2)	High (3)	Low (4)
<i>D(fWHR > median)</i>	0.007 (0.006)	0.024*** (0.006)	-0.015 (0.019)	0.015*** (0.004)
<i>LagAccuracy</i>	0.433*** (0.013)	0.442*** (0.013)	0.403*** (0.028)	0.439*** (0.009)
<i>Dayselap</i>	-0.012 (0.008)	0.015 (0.009)	-0.011 (0.023)	0.005 (0.006)
<i>Frequency</i>	-0.021*** (0.007)	-0.021*** (0.008)	0.004 (0.018)	-0.027*** (0.005)
<i>Horizon</i>	-0.454*** (0.016)	-0.440*** (0.014)	-0.423*** (0.032)	-0.458*** (0.011)
<i>Brokersize</i>	0.001 (0.006)	-0.008 (0.007)	0.013 (0.015)	-0.004 (0.004)
<i>Experience</i>	0.019** (0.007)	0.004 (0.009)	0.015 (0.017)	0.012** (0.006)
<i>Companies</i>	0.014 (0.018)	0.030* (0.018)	0.107*** (0.041)	0.024* (0.012)
<i>Industries</i>	-0.004 (0.018)	-0.011 (0.017)	-0.087** (0.038)	-0.002 (0.013)
<i>AnalystAge</i>	-0.003 (0.009)	-0.013 (0.012)	-0.061** (0.026)	0.002 (0.008)
<i>Postgraduate</i>	0.003 (0.009)	0.011 (0.009)	0.037* (0.022)	-0.006 (0.006)
<i>Constant</i>	0.555*** (0.017)	0.523*** (0.017)	0.551*** (0.035)	0.551*** (0.012)
Chow test statistics of coefficient difference in $D(fWHR > median)$	7.154***		5.040**	
<i>p</i> -value	0.008		0.025	
<i>N</i>	9,071	9,105	2,452	16,872
<i>Adj R2</i>	0.535	0.540	0.487	0.552

Table 5 Analysts' *fWHR* and forecast accuracy—sample split by uncertainty

This table reports results from OLS regressions relating an analyst's forecast accuracy to his *fWHR* (measured by $D(fWHR > median)$) for subsamples formed based on the uncertainty proxy. The dependent variable is the scaled forecast accuracy measure based on Clement and Tse (2005). To measure operating uncertainty, columns 1 and 2 use operating cash flow volatility (*CF_Volatility*), and columns 3 and 4 use research and development expenditure (*R&D*). We use the sample median of *CF_Volatility* (the *R&D* dummy indicating the use of research and development expenditure) to separate the sample in columns 1 and 2 (columns 3 and 4). Appendix provides a detailed description of the construction of all the variables. The number of observations reduced to 11,915 analyst-firm-years for the sample in columns 1 and 2 because the calculation of cash flow volatility requires non-missing operating cash flow deflated by total assets in the past five years. The number of observations reduced to 19,246 analyst-firm-years for the sample in columns 3 and 4 because we further require research and development expenditure to be non-missing. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Variables	Accuracy	Accuracy	Accuracy	Accuracy
	<i>CF_Volatility</i>		<i>R&D</i>	
	High (1)	Low (2)	High (3)	Low (4)
<i>D(fWHR > median)</i>	0.024*** (0.007)	0.001 (0.007)	0.027*** (0.008)	0.011** (0.005)
<i>LagAccuracy</i>	0.477*** (0.014)	0.448*** (0.014)	0.409*** (0.020)	0.440*** (0.009)
<i>Dayselap</i>	0.001 (0.011)	0.003 (0.011)	0.029** (0.013)	-0.004 (0.006)
<i>Frequency</i>	-0.029*** (0.008)	-0.029*** (0.009)	-0.016* (0.010)	-0.025*** (0.005)
<i>Horizon</i>	-0.439*** (0.019)	-0.403*** (0.016)	-0.487*** (0.021)	-0.449*** (0.012)
<i>Brokersize</i>	0.003 (0.008)	-0.004 (0.008)	0.005 (0.010)	-0.004 (0.005)
<i>Experience</i>	0.031** (0.012)	0.008 (0.010)	0.029** (0.014)	0.012** (0.006)
<i>Companies</i>	0.043** (0.021)	0.015 (0.021)	0.038* (0.022)	0.024* (0.014)
<i>Industries</i>	-0.019 (0.023)	-0.007 (0.021)	-0.013 (0.022)	-0.015 (0.014)
<i>AnalystAge</i>	-0.014 (0.013)	-0.006 (0.012)	-0.021 (0.013)	-0.008 (0.008)
<i>Postgraduate</i>	0.004 (0.011)	0.007 (0.009)	0.027** (0.012)	0.004 (0.006)
<i>Constant</i>	0.495*** (0.019)	0.531*** (0.017)	0.544*** (0.022)	0.544*** (0.013)
Chow test statistics of coefficient difference in $D(fWHR > median)$	6.314**		2.950*	
<i>p</i> -value	0.012		0.086	
<i>N</i>	5,948	5,967	3,623	15,623
<i>Adj R2</i>	0.537	0.512	0.567	0.541

Table 6 Analysts' *fWHR* and forecast accuracy—sample split by competition

This table reports results from OLS regressions relating an analyst's forecast accuracy to his *fWHR* (measured by $D(fWHR > median)$) for subsamples formed based on the competition intensity proxy. The dependent variable is the scaled forecast accuracy measure based on Clement and Tse (2005). To measure competition intensity, columns 1 and 2 use the number of analysts following the same firm (*NumAnalyst*), and columns 3 and 4 use a market sentiment indicator (*MKTSentiment*). We use the sample median to separate the sample. Appendix provides a detailed description of the construction of all the variables. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Variables	Accuracy	Accuracy	Accuracy	Accuracy
	<i>NumAnalyst</i>	<i>NumAnalyst</i>	<i>MKTSentiment</i>	<i>MKTSentiment</i>
	High (1)	Low (2)	High (3)	Low (4)
<i>D(fWHR > median)</i>	0.024*** (0.006)	0.006 (0.006)	0.018*** (0.006)	0.010 (0.006)
<i>LagAccuracy</i>	0.474*** (0.012)	0.404*** (0.011)	0.429*** (0.013)	0.440*** (0.012)
<i>Dayselap</i>	0.011 (0.009)	-0.001 (0.008)	-0.001 (0.009)	0.004 (0.008)
<i>Frequency</i>	-0.028*** (0.007)	-0.017*** (0.006)	-0.012* (0.006)	-0.036*** (0.007)
<i>Horizon</i>	-0.450*** (0.014)	-0.454*** (0.014)	-0.446*** (0.014)	-0.463*** (0.015)
<i>Brokersize</i>	0.009 (0.006)	-0.011* (0.006)	-0.004 (0.007)	0.001 (0.006)
<i>Experience</i>	0.013* (0.008)	0.014* (0.008)	0.016** (0.008)	0.014* (0.008)
<i>Companies</i>	-0.005 (0.018)	0.046*** (0.016)	0.048** (0.021)	-0.004 (0.021)
<i>Industries</i>	0.019 (0.023)	-0.035** (0.015)	-0.015 (0.017)	-0.005 (0.019)
<i>AnalystAge</i>	0.001 (0.013)	-0.018** (0.008)	-0.019* (0.011)	-0.003 (0.009)
<i>Postgraduate</i>	0.010 (0.009)	0.006 (0.008)	0.001 (0.009)	0.010 (0.008)
<i>Constant</i>	0.507*** (0.015)	0.569*** (0.015)	0.553*** (0.016)	0.541*** (0.015)
Chow test statistics of coefficient difference in $D(fWHR > median)$	5.161**		0.982	
<i>p</i> -value	0.023		0.322	
<i>N</i>	9,481	9,843	9,168	10,156
<i>Adj R2</i>	0.579	0.517	0.571	0.520

Table 7 Analysts' *fWHR*, alternative analyst performance measures, and other analyst forecast behaviors

This table reports the summary of the regression results relating an analyst's alternative performance measures and other forecast behaviors to his *fWHR* (measured by $D(fWHR > median)$). Column 1 reports a COX proportional hazard regression relating the hazard rate of an analyst being selected as a star analyst to his *fWHR*. Columns 2-9 report OLS regressions (logit regressions) if the dependent variable is continuous (discrete). The dependent variables in columns 1-6 are the alternative analyst performance measures: (1) *Hazard Rate (Star=1)*, the probability of an analyst being ranked as a *New Fortune* star analyst in a given year, conditional upon the analyst "having survived" without such an event until then; (2) *BHAR[0,+2]*, stock picking ability calculated as the buy-and-hold market-adjusted returns in the window of [0, +2] for the "buy" recommendations (for the "sell" recommendations, we multiply the return number by -1); (3) *Investment_Banking*, investment-banking business generating ability coded as one if an analyst's bank underwrites an initial public offering or a seasoned equity offering of the firm covered by him during the current year and zero otherwise; (4) *Bold*, forecast boldness coded as one if an analyst's forecast is above or below both the analyst's prior forecast and the consensus forecast immediately before his forecast and zero otherwise; (5) *Consistency*, forecast consistency measure following Hilary and Hsu (2013); and (6) *LFR*, analyst leader-follower-ratio following Cooper et al. (2001). The dependent variables in columns 7-9 are the other analyst forecast behaviors: (1) *Optimism*, forecast optimism coded as one if an analyst forecast is greater than the consensus forecast and zero otherwise; (2) *Overconfidence*, analyst overconfidence capturing the bold but inaccurate forecasts according to Hilary and Menzly (2006); and (3) *Walkdown*, forecast walkdown measure following Bradshaw et al. (2016). For the continuous dependent variables, we follow Clement and Tse (2005) and scale the measures so that the values range from 0 to 1. Appendix provides a detailed description of the construction of all the variables. Observations in column 1 (column 5) are at the analyst-year level (analyst-firm level). Accordingly, the control variables in column 1 (column 5) are the mean values of the control variables, calculated across firms for each analyst-year observation (across years for each analyst-firm observation). For brevity, the coefficient estimates on the control variables are not tabulated. The complete model specifications of the regressions are presented in Online Appendixes 1-9. All continuous variables are winsorized at 1% and 99%. The number of observations for *BHAR[0,+2]* reduced to 5,446 because the number of analysts who make stock recommendations is much lower than the number of analysts who issue earnings forecasts. We lose 357 observations for *Bold* and *Optimism* because, to calculate the consensus forecast, we require at least two other analysts' forecasts issued before the focal analyst's forecast for each firm-year. We lose 62% of the analyst-firm level observations for *Consistency* because of the calculation of standard deviation of the forecast errors during our sample period. We further lose 48 (2,647) observations for *LFR (Walkdown)* due to the calculation of the dependent variable. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Variables	<i>Hazard Rate (Star=1) (1)</i>	<i>BHAR[0,+2] (2)</i>	<i>Investment_Banking (3)</i>	<i>Bold (4)</i>	<i>Consistency (5)</i>	<i>LFR (6)</i>	<i>Optimism (7)</i>	<i>Overconfidence (8)</i>	<i>Walkdown (9)</i>
<i>D(fWHR > median)</i>	0.355** (0.178)	0.022** (0.011)	1.114*** (0.278)	0.124** (0.054)	0.040** (0.017)	0.034* (0.019)	-0.059 (0.067)	0.012 (0.016)	0.006 (0.008)
Hazard Rate	1.427								
Odds ratio			3.045	1.132			0.943		
Control variables	Included	Included	Included	Included	Included	Included	Included	Included	Included
<i>N</i>	2,670	5,446	19,324	18,967	3,302	19,276	18,967	19,324	16,677
<i>Adj R2</i>	0.045	0.008	0.216	0.045	0.023	0.047	0.020	0.017	0.092

Table 8 Robustness tests with sample perturbations

Panel A of this table reports the distribution of summary statistics (i.e., coefficients on $D(fWHR > median)$, T-values, and standard errors) of the OLS regressions (relating an analyst's forecast accuracy to $D(fWHR > median)$), respectively, applied to 1,000 samples formed by randomly dropping 1/3 of the final sample (i.e., 19,324 analyst-firm-years). Panel B reports the distribution of summary statistics of the OLS regressions (relating an analyst's forecast accuracy to $D(fWHR > median)$), respectively, applied to 1,000 pseudo samples with permuted $D(fWHR > median)$. From the 1,000 regressions reported in Panel B, Panel C (Panel D) reports the distribution for the regressions with a positive (negative) coefficient on $D(fWHR > median)$. Panel E reports the OLS regressions (relating an analyst's forecast accuracy to $D(fWHR > median)$), using the whole sample (column 1) and using the samples after dropping observations within a 5%, 10%, and 15% band around the median $fWHR$, respectively (columns 2, 3, and 4 respectively).

Panel A: Distribution of summary statistics of the regressions, respectively, applied to 1,000 samples formed by randomly dropping 1/3 of the final sample

Variable	Mean	P5	Q1	Median	Q3	P95	SD	N
$D(fWHR > median)$	0.014	0.009	0.012	0.014	0.016	0.020	0.003	1,000
T-value	2.737	1.711	2.341	2.716	3.123	3.795	0.615	1,000
S.E.	0.005	0.005	0.005	0.005	0.005	0.006	0.000	1,000

Panel B: Distribution of summary statistics of the regressions, respectively, applied to 1,000 pseudo samples with permuted $D(fWHR > median)$

Variable	Mean	P5	Q1	Median	Q3	P95	SD	N
$D(fWHR > median)$	0.000	-0.008	-0.003	0.000	0.003	0.008	0.005	1,000
T-value	0.017	-1.706	-0.722	0.003	0.716	1.785	1.090	1,000
S.E.	0.004	0.004	0.004	0.004	0.005	0.005	0.000	1,000

Panel C: Distribution of summary statistics of the regressions with a positive coefficient on $D(fWHR > median)$ (from the 1,000 regressions in Panel B)

Variable	Mean	P5	Q1	Median	Q3	P95	SD	N
$D(fWHR > median)$	0.004	0.000	0.001	0.003	0.006	0.009	0.003	501
T-value	0.879	0.070	0.309	0.716	1.335	2.105	0.683	501
S.E.	0.004	0.004	0.004	0.005	0.005	0.005	0.000	501

Panel D: Distribution of summary statistics of the regressions with a negative coefficient on $D(fWHR > median)$ (from the 1,000 regressions in Panel B)

Variable	Mean	P5	Q1	Median	Q3	P95	SD	N
$D(fWHR > median)$	-0.004	-0.010	-0.005	-0.003	-0.002	-0.000	0.003	499
T-value	-0.848	-2.148	-1.245	-0.731	-0.338	-0.052	0.647	499
S.E.	0.004	0.004	0.004	0.004	0.005	0.005	0.000	499

Panel E: Regressions using the whole sample and the samples after dropping observations within a 5%, 10%, and 15% band around the median $fWHR$, respectively

Variables	Accuracy (1)	Accuracy Drop 5% (2)	Accuracy Drop 10% (3)	Accuracy Drop 15% (4)
$D(fWHR > median)$	0.014*** (0.004)	0.015*** (0.005)	0.017*** (0.005)	0.017*** (0.006)
Controls	Yes	Yes	Yes	Yes
N	19,324	15,313	12,575	9,835
Adj R ²	0.544	0.546	0.551	0.559

Table 9 Female analysts' *fWHR*, forecast accuracy, and corporate site visits

This table uses the sample of 551 female analysts only. Panel A comprises 7,534 forecasts (issued by the 551 female analysts between 2004 and 2014) and reports results for OLS regressions relating an analyst's forecast accuracy to her *fWHR* (measured by $D(fWHR > median)$). The dependent variable is the scaled forecast accuracy measure based on Clement and Tse (2005). Panel B reports results for logistic regressions relating an analyst's corporate site visit probability to her *fWHR*. The number of observations is reduced to 3,472 analyst-firm-years, because only firms listed on Shenzhen Stock Exchange are required to report analysts' corporate site visit records. The dependent variable (*SiteVisit*) equals one if the analyst conducts at least one corporate site visit to the firm in the current year, and zero otherwise. Appendix provides a detailed description of the construction of all the variables. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Panel A: Female analyst's *fWHR* and forecast accuracy

Variables	Accuracy (1)	Accuracy (2)
<i>D(fWHR > median)</i>	0.016 (0.015)	0.004 (0.007)
<i>LagAccuracy</i>		0.411*** (0.013)
<i>Dayselap</i>		-0.004 (0.011)
<i>Frequency</i>		-0.030*** (0.008)
<i>Horizon</i>		-0.465*** (0.017)
<i>Brokersize</i>		-0.012* (0.007)
<i>Experience</i>		0.004 (0.008)
<i>Companies</i>		0.041** (0.019)
<i>Industries</i>		-0.033** (0.015)
<i>AnalystAge</i>		-0.005 (0.013)
<i>Postgraduate</i>		-0.003 (0.009)
<i>Constant</i>	0.692*** (0.012)	0.595*** (0.016)
<i>N</i>	7,534	7,534
<i>Adj R2</i>	0.001	0.537

Table 9 *continued*Panel B: Female analyst's *fWHR* and corporate site visit

Variables	Site Visit (1)		Site Visit (2)	
	Coefficient	Odds Ratio	Coefficient	Odds Ratio
<i>D(fWHR>median)</i>	0.461 (0.313)	1.586	0.423 (0.321)	1.526
<i>Distance</i>			0.177 (0.219)	1.193
<i>Brokern</i>			-0.066 (0.246)	0.936
<i>Tangidum</i>			0.021 (0.357)	1.021
<i>HHIdum</i>			-0.209 (0.449)	0.812
<i>Strongbuy</i>			0.829* (0.255)	2.290
<i>Drating</i>			1.061 (0.749)	2.889
<i>Coverage</i>			-0.003 (0.027)	0.997
<i>SOE</i>			0.219 (0.357)	1.245
<i>MV</i>			-0.682* (0.259)	0.506
<i>ROA</i>			6.650** (3.122)	772.768
<i>FirmAge</i>			-0.224 (0.269)	0.799
<i>BM</i>			-0.106 (0.708)	0.899
<i>Leverage</i>			2.169** (0.997)	8.746
<i>AnalystAge</i>			-1.649* (0.561)	0.192
<i>Postgraduate</i>			1.946*** (1.074)	7.001
<i>Constant</i>	-4.978* (0.848)	0.007	1.443 (3.797)	4.232
Industry Fixed effects	<i>Included</i>		<i>Included</i>	
<i>N</i>	3,472		3,472	
<i>Pseudo R2</i>	0.073		0.149	

Appendix: Variables definition

Variables	Definition	Data Source
-----------	------------	-------------

<i>fWHR_i</i>	Analyst <i>i</i> 's facial width-to-height ratio, measured as the distance between the left and right zygon (bizygomatic width) relative to the distance between the upper lip and the highest point of the eyelids (upper facial height).	SAC
<i>D(fWHR>median)_i</i>	An indicator variable coded as one if analyst <i>i</i> 's <i>fWHR</i> is above the median in the analyst-level sample, and zero otherwise.	SAC
<i>Accuracy_{ijt}</i>	Forecast accuracy for analyst <i>i</i> 's annual earnings forecast of firm <i>j</i> in year <i>t</i> , calculated as the absolute forecast error (deflated by stock price two days before the forecast). Then we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst <i>i</i> following firm <i>j</i> in year <i>t</i> . A higher value of <i>Accuracy_{ijt}</i> indicates a more accurate forecast.	CBAS CSMAR
<i>LagAccuracy_{ijt}</i>	Forecast accuracy for analyst <i>i</i> 's prior earnings forecast of firm <i>j</i> .	CBAS CSMAR
<i>Dayselap_{ijt}</i>	Day elapsed since the analyst's last forecast for firm <i>j</i> in year <i>t</i> , measured as the difference in days between analyst <i>i</i> 's forecast of firm <i>j</i> 's earnings in the year and the most recent preceding forecast of firm <i>j</i> 's earnings by any analyst. Then we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst <i>i</i> following firm <i>j</i> in year <i>t</i> . A higher value of <i>Dayselap_{ijt}</i> indicates more days elapsed.	CBAS
<i>Frequency_{ijt}</i>	Number of forecasts made by analyst <i>i</i> for firm <i>j</i> in year <i>t</i> . Then we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst <i>i</i> following firm <i>j</i> in year <i>t</i> . A higher value of <i>Frequency_{ijt}</i> indicates a higher forecast frequency.	CBAS
<i>Horizon_{ijt}</i>	Number of days between the forecast date of analyst <i>i</i> following firm <i>j</i> and the fiscal year-end of firm <i>j</i> in year <i>t</i> . Then we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst <i>i</i> following firm <i>j</i> in year <i>t</i> . A higher value of <i>Horizon_{ijt}</i> indicates a greater forecast horizon.	CBAS
<i>Brokersize_{ijt}</i>	Number of analysts employed by the brokerage employing analyst <i>i</i> following firm <i>j</i> in year <i>t</i> . Then we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst <i>i</i> following firm <i>j</i> in year <i>t</i> . A higher value of <i>Brokersize_{ijt}</i> indicates a larger brokerage size.	CBAS
<i>Experience_{ijt}</i>	Number of years of firm-specific experience for analyst <i>i</i> following firm <i>j</i> in year <i>t</i> . Then we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst <i>i</i> following firm <i>j</i> in year <i>t</i> . A higher value of <i>Experience_{ijt}</i> indicates more extensive firm-specific experiences.	CBAS
<i>Companies_{ijt}</i>	Number of companies analyst <i>i</i> follows in year <i>t</i> . Then we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst <i>i</i> following firm <i>j</i> in year <i>t</i> . A higher value of <i>Companies_{ijt}</i> indicates a	CBAS

	greater number of companies followed by analyst i in year t .	
<i>Industries_{ijt}</i>	Number of industries analyst i follows in year t . Then we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst i following firm j in year t . A higher value of <i>Industries_{ijt}</i> indicates a greater number of industries followed by analyst i in year t .	CBAS
<i>AnalystAge_{ijt}</i>	Age of analyst i in year t . Then we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst i following firm j in year t .	CBAS
<i>Postgraduate_i</i>	An indicator variable coded as one if analyst i has a Master's or Phd Degree, and zero otherwise.	CBAS
<i>Sitevisit_{ijt}</i>	An indicator variable coded as one if analyst i conducts at least one corporate site visit for firm j in year t , and zero otherwise.	CBAS SZSE
<i>Distance_{ijt}</i>	The natural logarithm of distance in kilometers between analyst i 's work location and the forecasted firm's headquarter for firm j in the year preceding year t .	CSMAR WIND iFinD
<i>Brokern_{ijt}</i>	The natural logarithm of the number of analysts employed by the brokerage employing analyst i following firm j for the year preceding year t .	CBAS
<i>Tangidum_{jt}</i>	An indicator variable coded as one if property, plant and equipment (PPE) deflated by total assets at the end of fiscal year preceding t is greater than the median value of the sample, and zero otherwise.	CSMAR
<i>HHIdum_{jt}</i>	An indicator variable coded as one if the Herfindahl-Hirschman index (HHI) of firm j 's sales in the year preceding t is greater than the median value of the sample, and zero otherwise.	CSMAR
<i>Strongbuy_{ijt}</i>	An indicator variable coded as one if analyst i issued at least one "strong buy" recommendation for firm j in the year preceding t , and zero otherwise.	CBAS
<i>Drating_{jt}</i>	An indicator variable coded as one if the rating of information disclosure quality assigned by the Shenzhen Stock Exchange to firm j in the year preceding t is A or B, and zero otherwise.	SZSE
<i>Coverage_{jt}</i>	The natural logarithm of analyst coverage, measured by the number of analyst following firm j in the year preceding t plus one.	CBAS
<i>SOE_{jt}</i>	An indicator variable coded as one if firm j is a state-owned enterprise in the year preceding t , and zero otherwise.	CSMAR
<i>MV_{jt}</i>	The natural logarithm of market value of firm j at the end of fiscal year preceding t .	CSMAR
<i>ROA_{jt}</i>	Net income deflated by total assets for firm j in the year preceding t .	CSMAR
<i>FirmAge_{jt}</i>	The natural logarithm of firm age, measured by the number of years between firm j 's establishment date and firm j 's fiscal year-end date of the year preceding t plus one.	CSMAR
<i>BM_{jt}</i>	Book value of equity divided by market value of equity of	CSMAR

	firm j at the end of fiscal year preceding t .	
<i>Leverage_{jt}</i>	Total liabilities divided by total assets for firm j at the end of fiscal year preceding t .	CSMAR
<i>BrokerAssets_{it}</i>	The previous year's total assets of the brokerage firm that analyst i works for in year t .	CBAS
<i>Title_{it}</i>	An indicator variable coded as one if analyst i holds one of the following titles in year t : department director, manager, president or chief analyst, and zero otherwise.	CBAS
<i>CF_Volatility_{jt}</i>	The standard deviation of the operating cash flow deflated by total assets in the past five years for firm j in the year preceding t .	CSMAR
<i>R&D_{jt}</i>	An indicator variable coded as one if firm j has research and development expenditure incurred in year t , and zero otherwise. We collect R&D from details of management expenses disclosed in footnote.	CSMAR
<i>NumAnalyst_{jt}</i>	The number of analyst following firm j in the year preceding t .	CBAS
<i>MKTSentiment_t</i>	An indicator variable coded as one if Shanghai Composite Index return exceeding one-year risk-free rate in year t , and equals zero otherwise.	CSMAR
<i>Star_{it}</i>	An indicator variable coded as one if analyst i is ranked as <i>New Fortune</i> star analysts in year t , and zero otherwise.	CBAS
<i>BHAR[0,2]_{ijt}</i>	Market-adjusted buy-and-hold return in the window of [0, +2] (where Day 0 indicates the recommendation announcement date) for the "buy" recommendations. For the "sell" recommendations, we multiply the return number by -1. Then we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst i following firm j in year t . A higher value of <i>BHAR[0,2]_{ijt}</i> indicates higher stock picking abilities.	CBAS CSMAR
<i>Investment_Banking_{ijt}</i>	An indicator variable coded as one if analyst i 's bank underwrites an initial public offering or a seasoned equity offering of firm j covered by him in year t , and zero otherwise.	CBAS CSMAR
<i>Bold_{ijt}</i>	An indicator variable which is coded as one if analyst i 's forecast is above or below both the analyst's prior forecast and the consensus forecast immediately before analyst i 's forecast, and coded as zero otherwise. The consensus forecast is based on forecasts issued in the 90 days prior to analyst i 's forecast.	CBAS
<i>Consistency_{ij}</i>	We construct forecast consistency for analyst i firm j in our sample period following Hillary and Hsu (2013). First, we calculate analyst i 's forecast error for firm j in year t as the actual earnings minus the analyst forecast. We calculate the standard deviation of the forecast errors over the sample period for each analyst-firm pair, requiring that analyst i has issued forecasts for firm j at least three times in the sample period. Finally, we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst i following firm j . A higher value of <i>Consistency_{ij}</i> indicates	CBAS

forecasts with more consistent forecast errors.

<i>LFR_{ijt}</i>	The leader-follower-ratio for analyst <i>i</i> following firm <i>j</i> in year <i>t</i> , calculated as the cumulative lead-time (i.e., the cumulative number of days by which the two preceding forecasts issued by other analysts lead analyst <i>i</i> 's forecast for firm <i>j</i>) divided by the cumulative follow-time (i.e., the cumulative number of days by which the two subsequent forecasts issued by other analysts follow analyst <i>i</i> 's forecast for firm <i>j</i>). Since other analysts are likely to issue forecasts soon after a leader analyst, the cumulative lead-time is likely to be larger relative to the cumulative follow-time, for a leader analyst. Therefore, a higher value of this ratio reflects a higher likelihood that this analyst is a leader analyst. We then scale this ratio to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst <i>i</i> following firm <i>j</i> in year <i>t</i> . A higher value of <i>LFR_{ijt}</i> indicates a greater leader-follower-ratio.	CBAS
<i>Optimism_{ijt}</i>	Forecast optimism for analyst <i>i</i> following firm <i>j</i> in year <i>t</i> , an indicator variable coded as one if analyst <i>i</i> 's forecast for firm <i>j</i> in year <i>t</i> is greater than the consensus forecast and zero otherwise.	CBAS
<i>Overconfidence_{ijt}</i>	Overconfidence for analyst <i>i</i> following firm <i>j</i> in year <i>t</i> , calculated as the absolute difference between analyst <i>i</i> 's forecast for firm <i>j</i> in year <i>t</i> and the consensus forecast (scaled into range between 0 and 1 following Clement and Tse (2005)'s approach) if his forecast accuracy is lower than the sample median, and zero otherwise. A higher value of <i>Overconfidence_{ijt}</i> indicates a higher degree of overconfidence.	CBAS
<i>Walkdown_{ijt}</i>	The forecast walkdown measure is calculated as the ending forecast minus the beginning forecast scaled by the absolute value of the ending forecast for analyst <i>i</i> following firm <i>j</i> in year <i>t</i> . Then we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst <i>i</i> following firm <i>j</i> in year <i>t</i> . A higher value of <i>Walkdown_{ijt}</i> indicates a higher degree of forecast walkdown.	CBAS

Online Appendix to:
“Facial Structure and Achievement Drive: Evidence from Financial Analysts”

1. Alternative performance measures

In this section, we test whether the *fWHR*-performance association is robust towards alternative performance measures, including the probability of being selected as star analysts, stock recommendation profitability, investment-banking business generating ability, forecast boldness, the consistency of forecast errors, and the timeliness of analyst forecasts.

1.1. Likelihood of being selected as star analysts

Being selected as star analysts represents a remarkable career achievement and it substantially increases the analyst’s reputation and bargaining power in the labor market. Our empirical analysis starts with the construction of an indicator variable (*Star*), which is coded as one if an analyst is ranked as a *New Fortune* star analyst during the year, and zero otherwise. We then estimate the time elapsed from the start of an analyst’s career until he is selected as a star analyst (i.e., time until *Star*=1), using the Cox proportional hazard model. Since not all analysts are selected as the star analyst during their career, our data are right-censored. The dependent variable is the hazard rate, defined as the probability of an analyst experiencing the event (i.e., *Star*=1) in the current year, given that the analyst “has survived” without such an event until then. We include the same set of control variables as Leone and Wu (2007). Detailed variable definitions are provided in the Appendix of the manuscript and Online Appendix 15.¹ The exponentiated coefficient on $D(fWHR > median)$ in of Table OA1 is 1.427, significant at the 5% level. Our findings indicate that analysts with above-median *fWHR* face an about 43 percent higher chance of being selected as star analysts than analysts with below-median *fWHR*.

1.2. Stock recommendation profitability

Similar to Wu and Zang (2009), we use the buy-and-hold market-adjusted returns in the window of [0, +2] (where Day 0 is the recommendation announcement date) to measure the stock

¹ Since analysts might follow more than one firm per year, we follow Leone and Wu (2007) and average the firm-level characteristics across firms for each analyst-year when we compute the control variables.

recommendation profitability for the “buy” recommendations. For the “sell” recommendations, we multiply the return value by -1. The higher the value, the better the analyst is at picking winning stocks. We estimate the association between an analyst’s stock recommendation profitability and his *fWHR* using the OLS regression model reported in Table OA2. We calculate the dependent variable by scaling the stock recommendation profitability figure to range from 0 to 1, following the approach in Clement and Tse (2005). Detailed variable definitions of the control variables are provided in the Appendix of the manuscript and Online Appendix 15. The positive and significant coefficients on $D(fWHR > median)$ in Table OA2 suggest that analysts with higher *fWHR* are better stock pickers than those with lower *fWHR*.

1.3. Investment-banking business generating ability

Table OA3 reports the logistic regression relating an analyst’s investment-banking business generating ability (measured by *Investment_Banking*) to his *fWHR*. Detailed variable definitions of the dependent and independent variables are provided in the Appendix of the manuscript. We find that the coefficient on $D(fWHR > median)$ is 1.114, significant at the 1% level. The odds ratio statistics indicate that the odds of generating investment-banking business are above two times higher for a high-*fWHR* analyst than for a low-*fWHR* analyst. Our results therefore suggest that high-*fWHR* analysts have higher investment-banking business generating abilities.

1.4. Likelihood of issuing bold forecasts

Following Clement and Tse (2005), forecast boldness (*Bold*) is an indicator variable coded as one if an analyst’s forecast is above or below both the analysts’ prior forecast and the consensus forecast immediately before his forecast, and zero otherwise. The consensus forecast is based on forecasts issued in the 90 days prior to the focal analyst’s forecast. We expect a positive association between $D(fWHR > median)$ and *Bold*. Table OA4 reports the logistic regression relating an analyst’s likelihood of issuing a bold forecast to his *fWHR* (measured by $D(fWHR > median)$). Detailed variable definitions of the control variables are provided in the Appendix of the manuscript. The coefficient on $D(fWHR > median)$ is 0.124, significant at the 5% level. The corresponding odds ratios suggest that the odds of issuing bold forecasts are 13% higher for a high-*fWHR* analyst than for a low-*fWHR* analyst.

1.5. Forecast consistency

We construct the forecast consistency variable ($Consistency_{ij}$) following Hilary and Hsu (2013). First, we calculate analyst i 's forecast error for firm j in year t as the actual earnings minus the analyst forecast. We calculate the standard deviation of the forecast errors over the sample period for each analyst-firm pair, requiring that analyst i has issued at least three forecasts for firm j in the sample period. Finally, we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst i following firm j . A higher value of $Consistency_{ij}$ indicates forecasts with more consistent forecast errors. Table OA5 examines the association between an analyst's $fWHR$ and his forecast consistency. Detailed variable definitions of the control variables are provided in the Appendix of the manuscript.² The positive and significant coefficient estimate on $D(fWHR > median)$ indicates that high- $fWHR$ analysts' forecast errors are more consistent than low- $fWHR$ analysts' forecast errors ($p < 0.05$).

1.6. Leading forecasts

Our measure of leader-follower-ratio (LFR) follows Cooper et al. (2001). Specifically, LFR_{ijt} for analyst i following firm j in year t is calculated as the cumulative lead-time (i.e., the cumulative number of days by which the two preceding forecasts issued by other analysts lead analyst i 's forecast for firm j) divided by the cumulative follow-time (i.e., the cumulative number of days by which the two subsequent forecasts issued by other analysts follow analyst i 's forecast for firm j). Since other analysts are likely to issue forecasts soon after a leader analyst, the cumulative lead-time is likely to be larger relative to the cumulative follow-time, for a leader analyst. Therefore, a higher value of this ratio reflects a higher likelihood that this analyst is a leader analyst. Then we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst i following firm j in year t . The positive and significant coefficient on $D(fWHR > median)$ tabulated in Table OA6 indicate that high- $fWHR$ analysts issue more timely forecasts than low- $fWHR$ analysts. Detailed variable definitions of the control variables are provided in the Appendix of the manuscript.

² Since the sample is at the analyst-firm level, accordingly, the control variables are the mean values of the variables defined in the Appendix of the manuscript, calculated across years for each analyst-firm observation.

2. Other analyst forecast behaviors

In this section, we investigate the association between analysts' *fWHR* and other forecasting behaviors: forecast optimism, analyst overconfidence and forecast walkdowns.

2.1. Forecast optimism and analyst overconfidence

Forecast optimism for analyst *i* following firm *j* in year *t* (*Optimism_{ijt}*) is an indicator variable coded as one if analyst *i*'s forecast for firm *j* in year *t* is greater than the consensus forecast and zero otherwise. Following Hilary and Menzly (2006), analyst overconfidence for analyst *i* following firm *j* in year *t* (*Overconfidence_{ijt}*) is calculated as the absolute difference between analyst *i*'s forecast for firm *j* in year *t* and the consensus forecast based on forecasts issued in 90 days prior to analyst *i*'s forecast (scaled into range between 0 and 1 following Clement and Tse (2005)'s approach) if his forecast accuracy is less than the sample median, and zero otherwise. Table OA7 (Table OA8) reports results of the logistic regression (OLS regression) relating an analyst's forecast optimism (overconfidence) to his *fWHR* measured by $D(fWHR > median)$. Detailed variable definitions of the control variables are provided in the Appendix of the manuscript. The insignificant coefficients on $D(fWHR > median)$ reported in both tables are in line with Jia et al. (2014)'s findings that *fWHR* measures a distinct personality trait different from optimism and/or overconfidence.

2.2. Forecast walkdown

The forecast walkdown measure for analyst *i* following firm *j* in year *t* (*Walkdown_{ijt}*) is calculated as the ending forecast minus the beginning forecast scaled by the absolute value of the ending forecast for analyst *i* following firm *j* in year *t*. Then we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst *i* following firm *j* in year *t*. A higher value of *Walkdown_{ijt}* indicates forecasts with greater level of forecast walkdown. Our results reported in Table OA9 show an insignificant relation between $D(fWHR > median)$ and forecast walkdown. The results are inconsistent with the conjecture that the *fWHR*-performance link is through the channel of strategic forecast behaviors. Detailed variable definitions of the control variables are provided in the Appendix of the manuscript.

3. Robustness test of our main results

For easier interpretation of our results, we use the indicator variable based on $fWHR$ (i.e., $D(fWHR > median)$) in the analyses reported in the manuscript. Table OA10 repeats the main test by replacing $D(fWHR > median)$ with the continuous $fWHR$ measure. Our conclusion continues to hold.

In the manuscript, $fWHR$ is computed as the average of two independent research assistants' measures of $fWHR$. To test the robustness of our finding, we calculate an alternative measure for $fWHR$ by dividing the average of the two research assistants' measures of bizygomatic width by the average of their measures of upper facial height. Results reported in Table OA11 show that our main results are robust towards this alternative measure.

In addition, Table OA12 repeats our main test using the correlated random effects approach (Wooldridge 2012). Our results continue to hold under this approach.

4. Portfolio of firms followed by analysts

We examine the differences in analyst following decisions between the two groups of analysts by testing the differences in the means of firm characteristics between firms followed by high- $fWHR$ analysts and those followed by low- $fWHR$ analysts. The high- $fWHR$ (low- $fWHR$) samples in Table OA13 comprise analysts with above (below) median $fWHR$ in the analyst-level sample. The examined firm characteristics include information visibility (measured by firm size, number of analyst followings, and institutional ownership (Bushee et al. 2010)), firm performance (measured by return on assets, return on equity, and stock returns), growth (measured by sales growth and earnings growth), equity risk (measured by return volatility), and the riskiness of corporate policies (measured by financial leverage, cash holdings, and accrual management (Chava and Purnanandam 2010)). Detailed variable definitions of the variables are provided in the Appendix of the manuscript.

Table OA13 reports no significant differences in these variables between the two groups of analysts, except for firm size (MV), unadjusted annual buy-and-hold stock returns ($BHAR$) and market-adjusted annual buy-and-hold stock returns ($BHAR_mktadj$). Our results suggest that high- $fWHR$ analysts tend to follow more prominent firms that are larger and perform better in terms of

stock returns. It is important to note that our regression variables are largely based on relative rankings among analysts following the *same* firm in the *same* year. Therefore, differences in firm characteristics (e.g., size) between the two groups of analysts does not affect our conclusions on the *fWHR*-performance link.

*5. Distribution of *fWHR* reported in the prior literature*

We conduct a comprehensive survey and detail distributions of *fWHR* as reported in the prior literature in Table OA14. Two points stand out. First, Asians tend to have higher *fWHR* ratios than Caucasians. The average ratio for Asians (except for celebrity endorsers) is slightly above 2 while most studies find that the average ratio for Caucasians (except for CEOs)³ is below 2. Second, the distribution of *fWHR* in our sample is similar to Kramer (2017), a subsample of which is also based on Chinese subjects. The mean value in our sample is 2.166 for males, close to the mean of 2.140 in Kramer (2017).

³ The average *fWHR* of the Caucasian CEOs is slightly above 2 (Alrajih and Ward 2014; Jia et al. 2014).

References:

- Alrajih, S., and J. Ward. 2014. Increased facial width-to-height ratio and perceived dominance in the faces of the UK's leading business leaders. *British Journal of Psychology* 105: 153–161.
- Anderl, C., T. Hahn, A. K. Schmidt, H. Moldenhauer, K. Notebaert, C. C. Clément, and S. Windmann. 2016. Facial Width-to-Height Ratio Predicts Psychopathic Traits in Males. *Personality and Individual Differences* 88: 99–101.
- Bird, B. M., V. S. Cid Jofré, S. N. Geniole, K. M. Welker, S. Zilioli, D. Maestripieri, S. Arnocky, and J. M. Carré. 2016. Does the Facial Width-to-Height Ratio Map onto Variability in Men's Testosterone Concentrations? *Evolution and Human Behavior* 37 (5): 392–398.
- Blankespoor, E., B. E. Hendricks, and G. S. Miller. 2017. Perceptions and Price: Evidence from CEO Presentations at IPO Roadshows. *Journal of Accounting Research* 55 (2): 275–327.
- Bradshaw, M. T., L. F. Lee, and K. Peterson. 2016. The interactive role of difficulty and incentives in explaining the annual earnings forecast walkdown. *Accounting Review* 91 (4): 995–1021.
- Bushee, B. J., J. E. Core, W. Guay, and S. J. W. Hamm. 2010. The role of the business press as an information intermediary. *Journal of Accounting Research* 48 (1): 1–19.
- Chava, S., and A. Purnanandam. 2010. CEOs versus CFOs: Incentives and corporate policies. *Journal of Financial Economics* 97 (2): 263–278.
- Carré, J. M., and C. M. McCormick. 2008. In Your Face: Facial Metrics Predict Aggressive Behaviour in the Laboratory and in Varsity and Professional Hockey Players. *Proceedings of the Royal Society B: Biological Sciences* 275: 2651–2656.
- Carré, J. M., K. R. Murphy, and A. R. Hariri. 2013. What Lies Beneath the Face of Aggression? *Social Cognitive and Affective Neuroscience* 8 (2): 224–229.
- Clement, M. B., and S. Y. Tse. 2005. Financial Analyst Characteristics and Herding Behavior in Forecasting. *The Journal of Finance* 60 (1): 307–341.
- Cooper, R. A., T. E. Day, and C. M. Lewis. 2001. Following the leader: A study of individual analysts' earnings forecasts. *Journal of Financial Economics* 61 (3): 383–416.
- Geniole, S. N., D. S. Molnar, J. M. Carré, and C. M. McCormick. 2014. The facial width-to-height ratio shares stronger links with judgments of aggression than with judgments of trustworthiness. *Journal of Experimental Psychology: Human Perception and Performance* 40 (4): 1526–1541.
- Goetz, S. M. M., K. S. Shattuck, R. M. Miller, J. A. Campbell, E. Lozoya, G. E. Weisfeld, and J. M. Carré. 2013. Social Status Moderates the Relationship Between Facial Structure and Aggression. *Psychological science* 24 (11): 2329–34.
- Haselhuhn, M. P., and E. M. Wong. 2012. Bad to the Bone: Facial Structure Predicts Unethical Behaviour. *Proceedings of the Royal Society B: Biological Sciences* 279 (1728): 571–576.
- Hilary, G., and C. Hsu. 2013. Analyst Forecast Consistency. *Journal of Finance* 68 (1): 271–297.
- Hilary, G., and L. Menzly. 2006. Does Past Success Lead Analysts to Become Overconfident? *Management Science* 52 (4): 489–500.
- Huh, H., D. Yi, and H. Zhu. 2014. Facial width-to-height ratio and celebrity endorsements. *Personality and Individual Differences* 68: 43–47.
- Jia, Y., L. Van Lent, and Y. Zeng. 2014. Masculinity, Testosterone, and Financial Misreporting. *Journal of Accounting Research* 52 (5): 1195–1246.
- Kramer, R. S. S. 2017. Sexual dimorphism of facial width-to-height ratio in human skulls and faces: A meta-analytical approach. *Evolution and Human Behavior* 38 (3): 414–420.
- Kramer, R. S. S., A. L. Jones, and R. Ward. 2012. A lack of sexual dimorphism in width-to-height ratio in white european faces using 2D photographs, 3D scans, and anthropometry. *PLoS ONE* 7 (8).
- Lebuda, I., and M. Karwowski. 2016. Written on the Writer's Face: Facial Width-to-Height Ratio among Nominees and Laureates of the Nobel Prize in Literature. *Creativity Research Journal* 28 (2): 207–211.
- Lefevre, C. E., P. J. Etchells, E. C. Howell, A. P. Clark, and I. S. Penton-Voak. 2014. Facial Width-to-height Ratio Predicts Self-reported Dominance and Aggression in Males and Females, But a Measure of Masculinity Does Not. *Biology Letters* 10 (1): 20140729–20140729.
- Lefevre, C. E., G. J. Lewis, T. C. Bates, M. Dzhelyova, V. Coetzee, I. J. Deary, and D. I. Perrett. 2012. No Evidence for Sexual Dimorphism of Facial Width-to-height Ratio in Four Large Adult

- Samples. *Evolution and Human Behavior* 33 (6): 623–627.
- Leone, A., and J. S. Wu. 2007. What Does It Take to Become a Superstar? Evidence from Institutional Investor Rankings of Financial Analysts. *Working Paper*. Penn State University and University of Rochester. SSRN eLibrary.
- Lewis, G. J., C. E. Lefevre, and T. C. Bates. 2012. Facial Width-to-height Ratio Predicts Achievement Drive in US Presidents. *Personality and Individual Differences* 52 (7): 855–857.
- Mileva, V. R., M. L. Cowan, K. D. Cobey, K. K. Knowles, and a. C. Little. 2014. In the face of dominance: Self-perceived and other-perceived dominance are positively associated with facial-width-to-height ratio in men. *Personality and Individual Differences* 69: 115–118.
- Özener, B. 2012. Facial Width-to-height Ratio in a Turkish Population is Not Sexually Dimorphic and is Unrelated to Aggressive Behavior. *Evolution and Human Behavior* 33 (3): 169–173.
- Sanchez-Pages, S., C. Rodriguez-Ruiz, and E. Turiegano. 2014. Facial masculinity: How the choice of measurement method enables to detect its influence on behaviour. *PLoS ONE* 9 (11).
- Stirrat, M., and D. I. Perrett. 2012. Face Structure Predicts Cooperation: Men With Wider Faces Are More Generous to Their In-Group When Out-Group Competition Is Salient. *Psychological Science* 23 (7): 718–722.
- Welker, K. M., S. M. M. Goetz, and J. M. Carré. 2015. Perceived and Experimentally Manipulated Status Moderates the Relationship between Facial Structure and Risk-taking. *Evolution and Human Behavior* 36 (6): 423–429.
- Welker, K. M., S. M. M. Goetz, S. Galicia, J. Liphardt, and J. M. Carré. 2015. An Examination of the Associations Between Facial Structure, Aggressive Behavior, and Performance in the 2010 World Cup Association Football Players. *Adaptive Human Behavior and Physiology* 1 (1): 17–29.
- Weston, E. M., A. E. Friday, and P. Liò. 2007. Biometric Evidence that Sexual Selection has Shaped the Hominin Face. *PLoS ONE* 2 (8): doi:10.1371/journal.pone.0000710.
- Wong, E. M., M. E. Ormiston, and M. P. Haselhuhn. 2011. A Face Only an Investor Could Love: CEOs' Facial Structure Predicts Their Firms' Financial Performance. *Psychological Science* 22 (12): 1478–1483.
- Wooldridge, J. M. 2012. *Introductory Econometrics: A Modern Approach*. 5th Editio. South-Western College Pub.
- Wu, J. S., and A. Y. Zang. 2009. What Determine Financial Analysts' Career Outcomes During Mergers? *Journal of Accounting and Economics* 47: 59–86.

Online Appendix 1 Analysts' *fWHR* and star recognition

This table reports results from a COX proportional hazard regression relating the hazard rate of an analyst being selected as a star analyst to his *fWHR* (measured by $D(fWHR > median)$). The star recognition is calculated as an indicator variable (*Star*) coded as one if an analyst is ranked as a *New Fortune* star analyst during the year and zero otherwise. The dependent variable is the hazard rate, defined as the probability of an analyst being ranked as a *New Fortune* star analyst (i.e., $Star=1$) in a given year, conditional upon the analyst "having survived" without such an event until then. Our sample includes 2,670 analyst-years from 2004 to 2014. Appendix of the manuscript and Online Appendix 15 provide a detailed description of the construction of all the variables. The control variables are the mean values of the variables, calculated across firms for each analyst-year observation. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Variables	<i>Hazard Rate (Star=1)</i>	
	Coefficient	Hazard Rate
<i>D(fWHR > median)</i>	0.355** (0.178)	1.427
<i>Accuracy</i>	2.369*** (0.704)	10.684
<i>Bias</i>	-0.985 (0.823)	0.373
<i>Frequency</i>	2.301*** (0.446)	9.983
<i>Companies</i>	-0.204 (0.480)	0.815
<i>Brokersize</i>	0.425 (0.318)	1.529
<i>Bold</i>	-0.417 (0.827)	0.659
<i>First-mover</i>	-0.246 (0.887)	0.782
<i>LFR</i>	0.025* (0.015)	1.026
<i>Experience</i>	0.695 (0.577)	2.004
<i>AnalystAge</i>	-0.033* (0.019)	0.968
<i>Postgraduate</i>	0.092 (0.294)	1.097
<i>N</i>	2,670	
<i>Pseudo R2</i>	0.045	

Online Appendix 2 Analysts' *fWHR* and stock picking abilities

This table reports results from an OLS regression relating an analyst's stock picking ability to his *fWHR* (measured by $D(fWHR > median)$). The stock recommendation profitability is measured by the buy-and-hold market-adjusted returns in the window of $[0, +2]$ for the "buy" recommendations. For the "sell" recommendations, we multiply the return number by -1. We follow Clement and Tse (2005) and scale the profitability measure into a range between 0 and 1 to construct the dependent variable, $BHAR[0, +2]$. A higher value of $BHAR[0, +2]$ indicates higher stock picking abilities. The number of observations is reduced to 5,446 because the number of analysts who make stock recommendations is much lower than the number of analysts who issue earnings forecasts. Appendix of the manuscript and Online Appendix 15 provide a detailed description of the construction of all the variables. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Variables	<i>BHAR[0,+2]</i>
<i>D(fWHR>median)</i>	0.022** (0.011)
<i>LagBHAR[0,+2]</i>	-0.050*** (0.012)
<i>Recomms</i>	-0.013 (0.012)
<i>Dayselap</i>	-0.045*** (0.013)
<i>Frequency</i>	0.050*** (0.019)
<i>Brokersize</i>	-0.018 (0.015)
<i>Experience</i>	-0.013 (0.012)
<i>Horizon</i>	0.046*** (0.017)
<i>Companies</i>	-0.040* (0.022)
<i>Industries</i>	0.065*** (0.023)
<i>AnalystAge</i>	0.007 (0.015)
<i>Postgraduate</i>	0.016 (0.018)
<i>Constant</i>	0.447*** (0.026)
<i>N</i>	5,446
<i>Adj R2</i>	0.008

Online Appendix 3 Analysts' *fWHR* and investment-banking business generating ability

This table reports results from a logistic regression relating an analyst's investment-banking generating ability to his *fWHR* (measured by $D(fWHR > median)$). The dependent variable (*Investment_Banking*) is an indicator variable that equals one if the analyst's bank underwrites an initial public offering or a seasoned equity offering of the firm covered by him during the current year, and zero otherwise. Appendix of the manuscript provides a detailed description of the construction of all the variables. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Variables	<i>Investment_Banking</i>	
	Coefficient	Odds Ratio
<i>D(fWHR > median)</i>	1.114*** (0.278)	3.045
<i>LagAccuracy</i>	-1.219*** (0.313)	0.296
<i>Dayselap</i>	0.967*** (0.259)	2.631
<i>Frequency</i>	-0.948*** (0.348)	0.387
<i>Horizon</i>	-1.505*** (0.446)	0.222
<i>Brokersize</i>	1.604*** (0.214)	4.972
<i>Experience</i>	3.200*** (0.387)	24.522
<i>Companies</i>	-0.398 (0.777)	0.672
<i>Industries</i>	-3.474*** (0.796)	0.031
<i>AnalystAge</i>	-0.946** (0.372)	0.388
<i>Postgraduate</i>	0.940* (0.518)	2.560
<i>Constant</i>	-6.760*** (0.679)	0.001
<i>N</i>	19,324	
<i>Adj R2</i>	0.216	

Online Appendix 4 Analysts' *fWHR* and likelihood of issuing bold forecasts

This table reports results from a logistic regression relating an analyst's likelihood of issuing bold forecasts to his *fWHR* (measured by $D(fWHR > median)$). The dependent variable (*Bold*) is an indicator variable coded as one if an analyst's forecast is above or below both the analyst's prior forecast and the consensus forecast immediately before his forecast and zero otherwise. The consensus forecast is based on forecasts issued in the 90 days prior to analyst *i*'s forecast. We lose 357 observations for *Bold* because, to calculate the consensus forecast, we require at least two other analysts' forecasts issued before the focal analyst's forecast for each firm-year. Appendix of the manuscript provides a detailed description of the construction of all the variables. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Variables	<i>Bold</i>	
	Coefficient	Odds Ratio
<i>D(fWHR > median)</i>	0.124** (0.054)	1.132
<i>LagAccuracy</i>	-1.501*** (0.071)	0.223
<i>Dayselap</i>	0.326*** (0.063)	1.385
<i>Frequency</i>	-0.415*** (0.054)	0.661
<i>Horizon</i>	-1.444*** (0.094)	0.236
<i>Brokersize</i>	-0.087* (0.049)	0.917
<i>Experience</i>	0.427*** (0.071)	1.532
<i>Companies</i>	0.263** (0.122)	1.301
<i>Industries</i>	-0.112 (0.118)	0.894
<i>AnalystAge</i>	-0.047 (0.082)	0.954
<i>Postgraduate</i>	0.067 (0.087)	1.069
<i>Constant</i>	1.075*** (0.124)	2.931
<i>N</i>	18,967	
<i>Adj R2</i>	0.045	

Online Appendix 5 Analysts' *fWHR* and forecast consistency

This table reports results from an OLS regression relating the consistency of an analyst's forecast errors to his *fWHR* (measured by $D(fWHR > median)$). To construct the dependent variable (*Consistency_{ij}*), we first calculate the standard deviation of the forecast errors over the sample period for each analyst-firm pair, requiring that analyst *i* has issued forecasts for firm *j* at least three times in the sample period (Hilary and Hsu 2013). We then follow Clement and Tse (2005) and scale this measure into a range between 0 and 1 to construct *Consistency_{ij}*. A higher value of *Consistency_{ij}* indicates forecasts with more consistent forecast errors. Our sample includes 3,302 analyst-firms from 2004 to 2014. We lose 62% of the analyst-firm level observations for *Consistency* because of the calculation of standard deviation of the forecast errors during our sample period. Appendix of the manuscript provides a detailed description of the construction of all the variables. The control variables are the mean values of the variables, calculated across years for each analyst-firm observation. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Variables	Consistency
<i>D(fWHR > median)</i>	0.040** (0.017)
<i>LagAccuracy</i>	0.243*** (0.059)
<i>Dayselap</i>	-0.086 (0.059)
<i>Frequency</i>	0.010 (0.070)
<i>Horizon</i>	-0.232** (0.097)
<i>Brokersize</i>	-0.040 (0.039)
<i>Experience</i>	-0.045 (0.049)
<i>Companies</i>	-0.028 (0.099)
<i>Industries</i>	-0.038 (0.099)
<i>AnalystAge</i>	0.078 (0.056)
<i>Postgraduate</i>	0.024 (0.027)
<i>Constant</i>	0.426*** (0.078)
<i>N</i>	3,302
<i>Adj R2</i>	0.023

Online Appendix 6 Analysts' *fWHR* and forecast timeliness

This table reports results from an OLS regression relating an analyst's forecast timeliness to his *fWHR* (measured by $D(fWHR > median)$). The forecast timeliness is captured by analyst leader-follower-ratio introduced by Cooper et al. (2001), calculated as the cumulative lead-time (i.e., the cumulative number of days by which the two preceding forecasts issued by other analysts lead analyst i 's forecast for firm j) divided by the cumulative follow-time (i.e., the cumulative number of days by which the two subsequent forecasts issued by other analysts follow analyst i 's forecast for firm j). We scale this ratio into a range between 0 and 1 to construct the dependent variable (*LFR*), following Clement and Tse (2005). A higher value of *LFR* indicates a greater leader-follower-ratio (i.e., more timely forecasts). We lose 48 observations due to the calculation of *LFR* variable. Appendix of the manuscript provides a detailed description of the construction of all the variables. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Variables	<i>LFR</i>
<i>D(fWHR > median)</i>	0.034* (0.019)
<i>LagAccuracy</i>	0.011 (0.009)
<i>Dayselap</i>	-0.108*** (0.012)
<i>Frequency</i>	-0.024*** (0.009)
<i>Horizon</i>	0.118*** (0.018)
<i>Brokersize</i>	-0.013 (0.012)
<i>Experience</i>	0.017 (0.014)
<i>Companies</i>	-0.136*** (0.048)
<i>Industries</i>	0.064 (0.044)
<i>AnalystAge</i>	0.033 (0.025)
<i>Postgraduate</i>	-0.050** (0.024)
<i>Constant</i>	0.750*** (0.034)
<i>N</i>	19,276
<i>Adj R2</i>	0.047

Online Appendix 7 Analysts' *fWHR* and forecast optimism

This table reports results from a logistic regression relating an analyst's forecast optimism to his *fWHR* (measured by $D(fWHR > median)$). The dependent variable (*Optimism*) is an indicator variable coded as one if an analyst forecast is greater than the consensus forecast and zero otherwise. We lose 357 observations for *Optimism* because, to calculate the consensus forecast, we require at least two other analysts' forecasts issued before the focal analyst's forecast for each firm-year. Appendix of the manuscript provides a detailed description of the construction of all the variables. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Variables	<i>Optimism</i>	
	Coefficient	Odds Ratio
<i>D(fWHR > median)</i>	-0.059 (0.067)	0.943
<i>LagAccuracy</i>	-0.604*** (0.070)	0.547
<i>Dayselap</i>	-0.596*** (0.076)	0.551
<i>Frequency</i>	-0.187*** (0.061)	0.829
<i>Horizon</i>	0.428*** (0.105)	1.534
<i>Brokersize</i>	-0.098 (0.064)	0.907
<i>Experience</i>	-0.231*** (0.086)	0.794
<i>Companies</i>	0.373** (0.154)	1.452
<i>Industries</i>	-0.316* (0.176)	0.729
<i>AnalystAge</i>	0.137 (0.105)	1.147
<i>Postgraduate</i>	-0.109 (0.103)	0.897
<i>Constant</i>	-0.695*** (0.143)	0.499
<i>N</i>	18,967	
<i>Adj R2</i>	0.020	

Online Appendix 8 Analysts' *fWHR* and analyst overconfidence

This table reports results from an OLS regression relating an analyst's degree of overconfidence to his *fWHR* (measured by $D(fWHR > median)$). The dependent variable ($Overconfidence_{ijt}$) captures the bold but inaccurate forecasts according to Hilary and Menzly (2006). It is calculated as the absolute difference between analyst i 's forecast for firm j in year t and the consensus forecast based on forecasts issued in 90 days prior to analyst i 's forecast (scaled into range between 0 and 1 following Clement and Tse (2005)'s approach) if his forecast accuracy is less than the sample median, and zero otherwise. A higher value of $Overconfidence_{ijt}$ indicates a higher degree of overconfidence. Appendix of the manuscript provides a detailed description of the construction of all the variables. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Variables	<i>Overconfidence</i>
<i>D(fWHR > median)</i>	0.012 (0.016)
<i>LagAccuracy</i>	0.008 (0.018)
<i>Dayselap</i>	0.186*** (0.018)
<i>Frequency</i>	0.071*** (0.018)
<i>Horizon</i>	-0.186*** (0.032)
<i>Brokersize</i>	0.035*** (0.014)
<i>Experience</i>	0.068*** (0.019)
<i>Companies</i>	-0.104*** (0.035)
<i>Industries</i>	0.096** (0.040)
<i>AnalystAge</i>	-0.032 (0.024)
<i>Postgraduate</i>	-0.012 (0.035)
<i>Constant</i>	0.400*** (0.055)
<i>N</i>	19,324
<i>Adj R2</i>	0.017

Online Appendix 9 Analysts' *fWHR* and forecast walkdown

This table reports results from an OLS regression relating an analyst's forecast walkdown to his *fWHR* (measured by $D(fWHR > median)$). To construct the dependent variable (*Walkdown_{ijt}*), we first calculate the ending forecast minus the beginning forecast scaled by the absolute value of the ending forecast for analyst *i* following firm *j* in year *t*. Then we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst *i* following firm *j* in year *t*. A higher value of *Walkdown_{ijt}* indicates a higher degree of forecast walkdown. We lose 2,647 observations because of calculation of *Walkdown*. Appendix of the manuscript provides a detailed description of the construction of all the variables. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Variables	<i>Walkdown</i>
<i>D(fWHR > median)</i>	0.006 (0.008)
<i>LagAccuracy</i>	-0.050*** (0.011)
<i>Dayselap</i>	0.064*** (0.011)
<i>Frequency</i>	0.061*** (0.009)
<i>Horizon</i>	-0.341*** (0.016)
<i>Brokersize</i>	-0.026*** (0.008)
<i>Experience</i>	0.204*** (0.011)
<i>Companies</i>	0.082*** (0.025)
<i>Industries</i>	-0.040* (0.024)
<i>AnalystAge</i>	0.001 (0.013)
<i>Postgraduate</i>	0.020* (0.010)
<i>Constant</i>	0.381*** (0.018)
<i>N</i>	16,677
<i>Adj R2</i>	0.092

Online Appendix 10 Analysts' *fWHR* and forecast accuracy—untransformed continuous measure of *fWHR*

This table uses the untransformed continuous *fWHR* variable as our independent variable of interest and reports results for OLS regressions relating an analyst's forecast accuracy to his *fWHR*. The dependent variable is the scaled forecast accuracy measure based on Clement and Tse (2005). Appendix of the manuscript provides a detailed description of the construction of all the variables. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Variables	<i>Accuracy</i> (1)	<i>Accuracy</i> (2)
<i>fWHR</i>	0.120*** (0.045)	0.044*** (0.015)
<i>LagAccuracy</i>		0.432*** (0.009)
<i>Dayselap</i>		0.006 (0.006)
<i>Frequency</i>		-0.022*** (0.005)
<i>Horizon</i>		-0.446*** (0.011)
<i>Brokersize</i>		-0.002 (0.005)
<i>Experience</i>		0.014** (0.006)
<i>Companies</i>		0.025** (0.012)
<i>Industries</i>		-0.021 (0.013)
<i>AnalystAge</i>		-0.006 (0.008)
<i>Postgraduate</i>		0.008 (0.007)
<i>Constant</i>	0.422*** (0.099)	0.455*** (0.036)
<i>N</i>	19,324	19,324
<i>Adj R2</i>	0.003	0.538

Online Appendix 11 Analysts' *fWHR* and forecast accuracy—alternative measure of *fWHR*

This table reports results for an OLS regression relating an analyst's forecast accuracy to his *fWHR* (measured by $D(fWHR > median)$). The dependent variable is the scaled forecast accuracy measure based on Clement and Tse (2005). We alternatively measure *fWHR* by dividing the average of the two research assistants' measures of bizygomatic width by the average of their measures of height of upper face. Appendix of the manuscript provides a detailed description of the construction of all the variables. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Variables	Accuracy
<i>D(fWHR > median)</i>	0.014*** (0.004)
<i>LagAccuracy</i>	0.435*** (0.009)
<i>Dayselap</i>	0.002 (0.006)
<i>Frequency</i>	-0.023*** (0.005)
<i>Horizon</i>	-0.454*** (0.011)
<i>Brokersize</i>	-0.002 (0.004)
<i>Experience</i>	0.014*** (0.006)
<i>Companies</i>	0.027** (0.012)
<i>Industries</i>	-0.015 (0.013)
<i>AnalystAge</i>	-0.010 (0.008)
<i>Postgraduate</i>	0.006 (0.006)
<i>Constant</i>	0.545*** (0.012)
<i>N</i>	19,324
<i>Adj R2</i>	0.544

Online Appendix 12 Analysts' *fWHR* and forecast accuracy—correlated random effect

This table reports results for an OLS regression relating an analyst's forecast accuracy to his *fWHR* (measured by $D(fWHR > median)$) using the correlated random effects approach (Wooldridge 2012). The dependent variable is the scaled forecast accuracy measure based on Clement and Tse (2005). The variables starting with "M" are the mean of the control variables for each analyst. Appendix of the manuscript provides a detailed description of the construction of all the variables. All continuous variables are winsorized at 1% and 99%. The values in the parentheses are standard errors clustered by analyst in regressions. The symbols *, ** and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

Variables	Accuracy
<i>D(fWHR > median)</i>	0.011*** (0.004)
<i>Lagaccuracy</i>	0.424*** (0.009)
<i>MLagaccuracy</i>	0.147*** (0.028)
<i>Dayselap</i>	-0.001 (0.006)
<i>MDayselap</i>	0.017 (0.025)
<i>Frequency</i>	-0.023*** (0.005)
<i>MFrequency</i>	-0.003 (0.020)
<i>Horizon</i>	-0.458*** (0.011)
<i>MHorizon</i>	0.119*** (0.036)
<i>Brokersize</i>	-0.003 (0.005)
<i>MBrokersize</i>	-0.006 (0.012)
<i>Experience</i>	0.014** (0.006)
<i>MExperience</i>	0.005 (0.016)
<i>Companies</i>	0.020 (0.016)
<i>MCompanies</i>	0.017 (0.024)
<i>Industries</i>	0.016 (0.016)
<i>MIndustries</i>	-0.044* (0.024)
<i>AnalystAge</i>	0.016

Online Appendix 12 *continued*

	(0.016)
<i>MAnalystAge</i>	-0.029*
	(0.018)
<i>Postgraduate</i>	0.005
	(0.006)
<i>Constant</i>	0.430***
	(0.029)
<i>N</i>	19,324
<i>Adj R2</i>	0.546

Online Appendix 13 Analysts' *fWHR* and following decisions

This table reports the univariate tests on differences in the means of firm characteristics between firms followed by high-*fWHR* analysts and those followed by low-*fWHR* analysts. High-*fWHR* (low-*fWHR*) sample comprises analysts with above (below) median *fWHR* in the analyst-level sample. Appendix of the manuscript and Online Appendix 15 provide a detailed description of the construction of all the variables. *BHAR*, *BHAR_mktadj*, *Sales Growth*, *Earnings Growth*, *RetVolatility*, and *Leverage* have missing observations because of the missing values in the database. Two-tailed probability values are reported for the two-sample t-test. The values in the parentheses are standard errors. The symbols *, **, and *** denote significance at the 10%, 5% and 1% levels for two-tailed tests, respectively.

	<i>fWHR</i>	<i>Mean</i>	<i>Median</i>	<i>SD</i>	<i>N</i>	<i>T-test</i>
<i>MV</i>	<i>High</i>	16.108	15.923	1.133	9,289	0.076***
	<i>Low</i>	16.032	15.857	1.076	10,035	(4.760)
<i>Instown</i>	<i>High</i>	9.325	6.980	10.096	9,289	-0.013
	<i>Low</i>	9.338	7.020	10.171	10,035	(-0.089)
<i>NumAnalyst</i>	<i>High</i>	15.437	15.000	8.584	9,289	-0.077
	<i>Low</i>	15.514	15.000	8.439	10,035	(-0.629)
<i>ROA</i>	<i>High</i>	0.074	0.067	0.051	9,289	0.001
	<i>Low</i>	0.073	0.066	0.051	10,035	(1.493)
<i>ROA_indadj</i>	<i>High</i>	0.020	0.021	0.103	9,289	0.001
	<i>Low</i>	0.019	0.023	0.109	10,035	(0.600)
<i>ROE</i>	<i>High</i>	0.131	0.121	0.079	9,289	0.002
	<i>Low</i>	0.129	0.119	0.081	10,035	(1.611)
<i>ROE_indadj</i>	<i>High</i>	0.045	0.039	0.232	9,289	0.002
	<i>Low</i>	0.043	0.041	0.229	10,035	(0.725)
<i>BHAR</i>	<i>High</i>	1.306	1.068	0.793	9,229	0.025**
	<i>Low</i>	1.281	1.088	0.764	9,971	(2.231)
<i>BHAR_mktadj</i>	<i>High</i>	0.191	0.068	0.556	9,225	0.014*
	<i>Low</i>	0.177	0.060	0.529	9,967	(1.727)
<i>Sales Growth</i>	<i>High</i>	0.268	0.212	0.331	8,780	0.003
	<i>Low</i>	0.265	0.212	0.340	9,438	(0.552)
<i>Earnings Growth</i>	<i>High</i>	0.298	0.218	1.162	8,780	0.008
	<i>Low</i>	0.290	0.211	1.194	9,439	(0.449)
<i>RetVolatility</i>	<i>High</i>	0.030	0.027	0.025	9,228	-0.000
	<i>Low</i>	0.030	0.027	0.020	9,970	(-0.013)
<i>Leverage</i>	<i>High</i>	0.424	0.429	0.210	9,287	0.001
	<i>Low</i>	0.423	0.430	0.207	10,035	(0.464)
<i>Cash</i>	<i>High</i>	0.234	0.182	0.171	9,289	-0.002
	<i>Low</i>	0.236	0.186	0.172	10,035	(-0.741)
<i>Accrual</i>	<i>High</i>	0.026	0.007	0.175	9,289	0.002
	<i>Low</i>	0.024	0.005	0.170	10,035	(0.813)

Online Appendix 14 Distribution of *fWHR* reported in other papers

Paper	N	Ethnicity or Nationality	Bound of facial height	Gender	Participants	<i>fWHR</i> Summary Statistics		
						Mean	Std. dev	Range
Alrajih and Ward (2014)	93	UK	eyelids	M	CEOs of the UK FTSE-100 Index	2.04	0.16	–
	93	UK	eyelids	M	Individuals selected from the internet (matched to the CEOs in terms of appearance features)	1.87	0.15	–
Anderl et al. (2016)	96	Caucasian	eyelids	M	Undergraduate students from various disciplines at <i>Goethe University Frankfurt</i> (Germany)	–	–	1.6-2.2
	41	Caucasian	eyelids	M	Prisoners of a German young offenders institution (<i>Justizvollzugsanstalt Ebrach</i>)	–	–	1.6-2.4
Blankespoor et al. (2017)	224	USA	eyelids	M/F	U.S. CEOs presenting the IPO roadshow	2.08	0.16	–
Bird et al. (2016), Study 1	80	86% Caucasian, 5% Asian, 4% bi-racial, 5% others	brows	M	Participants from <i>Nipissing University (NU)</i> Research Participation Pool and from ads in <i>NU</i> community (ages 18-33)	1.83	0.13	1.56–2.10
Bird et al. (2016), Study 2	114	53% Caucasian, 19% bi-racial, 18% Asian, 10% Black	brows	M	Participants from <i>Wayne State University (WSU)</i> Research Participation Pool (ages 17-56)	1.71	0.12	1.44–2.11
Bird et al. (2016), Study 3	153	38% Caucasian, 20% Black, 18% Asian, 10% Middle Eastern, 14% others	brows	M	Undergraduate students between the ages of 18 and 34	1.77	0.15	1.48–2.23
Bird et al. (2016), Study 4	159	44% Caucasian, 22% Native American, 14% Asian, 8% bi-racial, 6% Black, 6% Latin American	brows	M	MBA students in the 2008 cohort at the <i>University of Chicago Graduate School of Business</i> (ages 24-35)	1.83	0.14	1.53–2.21
Bird et al. (2016), Study 5	95	46% Caucasian, 21% Black, 16% bi-racial, 17% others	brows	M	Undergraduate students between the ages of 18 and 30	1.76	0.16	1.38–2.09
Bird et al. (2016), Study 6	73	75% Caucasian, 9% Asian, 9% Middle Eastern, 7% others	brows	M	Randomly selected participants between the ages of 18 and 40	1.61	0.13	1.34–1.87
Bird et al. (2016), Study 7	119	78% Caucasian, 13% Aboriginal, 9% others	brows	M	Randomly selected participants between the ages of 18 and 35	1.73	0.15	1.41–2.26
Carré and McCormick (2008)	37	82% Caucasian, 18% others	brows	M	Undergraduate students	1.86	0.13	–
	51	82% Caucasian, 18% others	brows	F	Undergraduate students	1.80	0.10	–
Carré et al. (2013)	28	Caucasian	brows	M	Students for <i>Duke Neurogenetics Study</i> (ages 19-21)	1.81	0.12	–
	36	Caucasian	brows	F	Students for <i>Duke Neurogenetics Study</i> (ages 19-21)	1.75	0.12	–

Paper	N	Ethnicity or Nationality	Bound of facial height	Gender	Participants	fWHR Summary Statistics		
						Mean	Std. dev	Range
Geniole et al. (2014) , Study 1	25	Caucasian	brows	M	Randomly selected participants	1.87	0.16	–
Geniole et al. (2014) , Study 2	54	Caucasian	brows	M	Randomly selected participants	1.85	0.19	–
Geniole et al. (2014) , Study 3	22	Caucasian	brows	M	Randomly selected participants	1.88	0.15	–
Geniole et al. (2014) , Study 4	65	Caucasian	brows	M	Randomly selected participants	1.91	0.16	–
Goetz et al. (2013)	106	Caucasian	brows	M	Students at <i>Wayne State University</i>	1.72	0.12	–
	110	Caucasian	brows	F	Students at <i>Wayne State University</i>	1.70	0.15	–
Haselhuhn and Wong (2012), Study 1	51	USA	N/A	M	MBA students at <i>University of Wisconsin-Milwaukee</i>	1.78	0.12	1.48–1.98
	45	USA	N/A	F	MBA students at <i>University of Wisconsin-Milwaukee</i>	1.74	0.12	1.45–2.00
Haselhuhn and Wong (2012), Study 2	50	USA	N/A	M	Undergraduate students at <i>University of Wisconsin-Milwaukee</i> (avg. age 22)	1.79	0.15	1.44-2.08
	53	USA	N/A	F	Undergraduate students at <i>University of Wisconsin-Milwaukee</i> (avg. age 22)	1.73	0.21	1.21-2.31
Huh et al. (2014)	47	Asian (Republic of Korea)	brows	M	Korean celebrity endorsers	1.78	0.13	–
	50	Asian (Republic of Korea)	brows	F	Korean celebrity endorsers	1.78	0.11	–
Jia et al. (2014)		USA	eyelids	M	S&P 1500 CEOs in 2009	2.01	0.15	–
Kramer (2017)	56	East Asian (China)	eyelids	M	Chinese students at <i>Bangor University</i> (ages 19-44)	2.14	0.14	–
	79	East Asian (China)	eyelids	F	Chinese students at <i>Bangor University</i> (ages 19-44)	2.11	0.11	–
Kramer et al. (2012), Study 1	138	Caucasian (Germany)	eyelids	M	Individual volunteers	1.85	0.11	–
	277	Caucasian (Germany)	eyelids	F	Individual volunteers	1.87	0.11	–
Kramer et al. (2012), Study 2	66	Caucasian (UK)	eyelids	M	Students at <i>Bangor University</i> (ages 18-30)	2.01	0.16	–
	89	Caucasian (UK)	eyelids	F	Students at <i>Bangor University</i> (ages 18-30)	2.03	0.14	–
Kramer et al. (2012), Study 3	75	Caucasian (UK)	eyelids	M	Students at <i>Bangor University</i> (ages 18-29)	2.07	0.16	–
	105	Caucasian (UK)	eyelids	F	Students at <i>Bangor University</i> (ages 18-29)	2.07	0.15	–
Lebuda and Karwowski (2016)	44	N/A	eyelids	M	Laureates of the Nobel Prize	1.79	0.12	–
	368	N/A	eyelids	M	Nominees of the Noble Prize	1.79	0.15	–
Lefevre et al. (2014)	54	95% Caucasian (UK)	eyelids	M	Students at the <i>University of Bristol</i> (ages 18-30)	2.08	0.17	–
	49	95% Caucasian (UK)	eyelids	F	Students at the <i>University of Bristol</i> (ages 18-30)	2.01	0.13	–

Paper	N	Ethnicity or Nationality	Bound of facial height	Gender	Participants	fWHR Summary Statistics		
						Mean	Std. dev	Range
Lefevre et al. (2012), Sample 1	46	Caucasian (UK)	eyelids	M	Undergraduate students (ages 18-27)	2.12	0.18	–
	99	Caucasian (UK)	eyelids	F	Undergraduate students (ages 18-25)	2.17	0.16	–
Lefevre et al. (2012), Sample 2	137	Caucasian (UK)	eyelids	M	Scottish adults from the Lothian Birth Cohort 1921 (age 83)	2.06	0.17	–
	169	Caucasian (UK)	eyelids	F	Scottish adults from the Lothian Birth Cohort 1921 (age 83)	2.09	0.16	–
Lefevre et al. (2012), Sample 3	124	Caucasian (UK)	eyelids	M	Undergraduate students, with 3D photographs (ages 18-30)	1.84	0.13	–
	131	Caucasian (UK)	eyelids	F	Undergraduate students, with 3D photographs (ages 18-28)	1.88	0.11	–
Lefevre et al. (2012), Sample 4	108	Black (South Africa)	eyelids	M	Undergraduate students (ages 18-26)	2.20	0.24	–
	110	Black (South Africa)	eyelids	F	Undergraduate students (ages 18-29)	2.24	0.20	–
Lewis et al. (2012)	29	Caucasian (USA)	eyelids	M	Former US presidents	1.99	0.11	1.78-2.30
Mileva et al. (2014), Study 1	50	UK	brows	M	Undergraduate students (ages around 26)	1.96	0.14	–
	50	UK	brows	F	Undergraduate students (ages around 26)	1.96	0.14	–
Mileva et al. (2014), Study 2	31	UK	brows	M	Undergraduate students (ages around 22)	1.98	0.11	–
	29	UK	brows	F	Undergraduate students (ages around 22)	1.96	0.16	–
Mileva et al. (2014), Study 3	21	UK	brows	M	Undergraduate students (ages around 21)	1.96	0.14	–
	29	UK	brows	F	Undergraduate students (ages around 21)	1.95	0.16	–
Özener (2012)	230	Turkey	brows	M	Students from the Faculty of Health Sciences at three different universities in Ankara	1.89	0.12	–
	240	Turkey	brows	F	Students from the Faculty of Health Sciences at three different universities in Ankara	1.91	0.11	–
Sanchez-Pages et al. (2014)	147	Caucasian (Spain and UK)	brows	M	Students from Madrid and Edinburgh	2.09	0.16	–
	74	Caucasian (Spain and UK)	brows	F	Students from Madrid and Edinburgh	2.05	0.13	–
Stirrat and Perrett (2012)	32	Caucasian (UK)	brows	M	Students at the <i>University of St Andrews</i>	2.03	–	–

Paper	N	Ethnicity or Nationality	Bound of facial height	Gender	Participants	fWHR Summary Statistics		
						Mean	Std. dev	Range
Welker et al. (2015)	910	Multinational	brows	M	Football players from the 32 countries competing in the 2010 FIFA world cup finals and qualifying competitions	1.89	0.15	–
Welker, Stefan M M Goetz, et al. (2015)	165	38% Caucasian, 20% Black, 19% Asian, 23% others	brows	M	Participants enrolled at <i>Wayne State University</i>	1.77	0.15	–
Weston et al. (2007)	30	Native South Africa	nasion	M	MicroScribe G2 digitiser from dry skulls from the <i>Raymond Dart</i> collection, <i>University of Witwaterstrand</i>	1.92	0.10	–
	30	Native South Africa	nasion	F	MicroScribe G2 digitiser from dry skulls from the <i>Raymond Dart</i> collection, <i>University of Witwaterstrand</i>	1.84	0.10	–
Wong et al. (2011)	50	USA	brows	M	CEOs of Fortune 500 companies	1.96	0.15	1.61-2.25

Online Appendix 15 Variables definition⁴

Variables	Definition	Data Source
<i>Bias_{it}</i>	The average value of <i>Bias_{jt}</i> for all the firms covered by analyst <i>i</i> in year <i>t</i> . Forecast bias is calculated as analyst <i>i</i> 's earnings forecast for firm <i>j</i> at time <i>t</i> minus firm <i>j</i> 's actual earnings. Then we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst <i>i</i> following firm <i>j</i> in year <i>t</i> to construct <i>Bias_{ijt}</i> .	CBAS
<i>First-mover_{it}</i>	The average value of <i>First-mover_{ijt}</i> for all the firms covered by analyst <i>i</i> in year <i>t</i> . <i>First-mover_{ijt}</i> is an indicator variable coded as one if analyst <i>i</i> is the first analyst who issues an earnings forecast for firm <i>j</i> in year <i>t</i> , and equals zero otherwise.	CBAS
<i>LagBHAR[0,+2]_{ijt}</i>	Stock recommendation profitability for analyst <i>i</i> 's stock recommendation of firm <i>j</i> in year <i>t-1</i> , calculated as <i>BHAR[0,+2]_{ijt-1}</i> .	CBAS CSMAR
<i>Recomms_{ijt}</i>	Stock recommendation, coded as 1 for a "Strong Buy", 2 for a "Buy", 3 for a "Neutral", 4 for a "Sell", and 5 for a "Strong Sell". Then we scale this variable to range from 0 to 1, employing the approach in Clement and Tse (2005) for analyst <i>i</i> following firm <i>j</i> in year <i>t</i> .	CBAS
<i>Instown_{jt}</i>	Institutional ownership of firm <i>j</i> in the year preceding <i>t</i> .	CSMAR
<i>ROA_indadj_{jt}</i>	<i>ROA_{jt}</i> minus the median of the industry peers' return on assets in the year preceding <i>t</i> . Industries are classified by the first digit of CSRC industry codes for non-manufacturing firms and first three digits for manufacturing firms.	CSMAR
<i>ROE_{jt}</i>	Return on equity defined as net income deflated by total equity of firm <i>j</i> in the year preceding <i>t</i> .	CSMAR
<i>ROE_indadj_{jt}</i>	<i>ROE_{jt}</i> minus the median of the industry peers' return on equity in the year preceding <i>t</i> . Industries are classified by the first digit of CSRC industry codes for non-manufacturing firms and first three digits for manufacturing firms.	CSMAR
<i>BHAR_{jt}</i>	Buy-and-hold return of firm <i>j</i> in the year preceding <i>t</i> .	CSMAR
<i>BHAR_mktadj_{jt}</i>	Buy-and-hold market-adjusted return of firm <i>j</i> in the year preceding <i>t</i> .	CSMAR
<i>Sales Growth_{jt}</i>	Growth in total sales of firm <i>j</i> in the year preceding <i>t</i> .	CSMAR
<i>Earnings Growth_{jt}</i>	Growth in net income of firm <i>j</i> in the year preceding <i>t</i> .	CSMAR
<i>RetVolatility_{jt}</i>	Standard deviation of daily return of firm <i>j</i> in the year preceding <i>t</i> .	CSMAR
<i>Cash_{jt}</i>	Ratio of cash holdings to the total assets of firm <i>j</i> in the year preceding <i>t</i> .	CSMAR
<i>Accrual_{jt}</i>	The absolute value of the difference between firm's earnings and cash flows scaled by previous year's sales of firm <i>j</i> in the year preceding <i>t</i> .	CSMAR

⁴ In this table, we only include variables that are not defined in the manuscript.