Gender and Beauty in the Financial Analyst Profession: Evidence from the U.S. and China

Congcong Li
Duquesne University
lic3@duq.edu

An-Ping Lin
Singapore Management University
aplin@smu.edu.sg

Hai Lu *
University of Toronto
hai.lu@rotman.utoronto.ca

Kevin Veenstra
McMaster University
veenstk@mcmaster.ca

This version: August 25, 2019

* Corresponding author. We gratefully acknowledge helpful comments from Daniel Bens, Mark Bradshaw, Hye Sun Chang, Xia Chen, Richard Crowley, Peter Joos, Meng Li, Mark Ma, Siew Hong Teoh, Holly Yang, two anonymous reviewers and seminar participants at the 2017 ABFER Annual Conference, the 2017 Conference on the Convergence of Financial and Managerial Accounting Research, the 2017 FARS Midyear Meeting, and Singapore Management University. We also thank PhD students Chao Dou, Yuanyuan Liu and Meng Luck Tay for their excellent research assistance in data collection. We thank McMaster University, Singapore Management University, University of Toronto and Social Sciences and Humanities Research Council of Canada for their financial support.

Gender and Beauty in the Financial Analyst Profession: Evidence from the U.S. and China

ABSTRACT

We examine how gender and beauty affect the likelihood of being voted as an All-Star in the financial analyst profession in both the U.S. and China. We find that female analysts are more likely to be voted as All-Star analysts in the U.S., but good-looking female U.S. analysts are less likely to be voted as All-Stars. The conclusion is the opposite for Chinese financial analysts. We find that female analysts in China are less likely to be voted as All-Stars, but the likelihood increases with their facial attractiveness. These findings implicate a beauty penalty for female analysts in the U.S. and gender discrimination against female analysts in China. This career path evidence from a competitive financial industry suggests that gender and beauty biases may be rooted deeply in culture and legal environment and should not be treated homogenously.

JEL Classification: D83, G11, G24, J24, J44, M41

Keywords: Analysts, Gender, Beauty, Labor Market

Gender and Beauty in the Financial Analyst Profession: Evidence from the U.S. and China

1. Introduction

This study investigates gender discrimination and beauty bias in the financial industry in the U.S. and China. Specifically, we examine whether the likelihood of becoming a star analyst, one of the most important career outcomes for financial analysts, is affected by analysts' gender and beauty in both the U.S. and China. These two countries have very different legal policies and cultures. After the passage of the federal civil rights legislation of the 1960s in the U.S., overt forms of discrimination, such as stating that a particular gender or type of appearance is preferred in job postings, are illegal. In China, however, these explicit forms of discrimination and biases are common. Kuhn and Shen (2013) find that over one-third of Chinese companies seeking highly educated urban employees have at least one Internet job posting stating a preferred gender. In addition, many Chinese government departments and private companies make explicit requests for age, height and beauty in their Internet job postings (Human Rights Watch 2018).[1] Such cultural and legislative differences call into question whether the gender and beauty biases in the labor market are different in the U.S. and China.

The issue as to whether gender discrimination exists in the labor market remains debated among academics and policymakers. While the Civil Rights Act of 1964 and subsequent related legislation purged overt forms of gender discrimination in the U.S., gender discrimination continues in more hidden and subtle forms in American society (Kuhn and Shen 2013). Prior literature documents that American firms pay male employees higher wages and are more likely to promote male employees relative to their female counterparts with similar credentials or output

[1] https://www.hrw.org/report/2018/04/23/only-men-need-apply/gender-discrimination-job-advertisements-china.

(Altonji and Blank 1999; Goldin and Rouse 2000). One recent study finds that female financial advisers in America are more likely to lose their jobs and less likely to find new jobs relative to their male peers (Egan, Matvos, and Seru 2017). In the Chinese context, Gao, Lin and Ma (2016) document that firms headquartered in more gender-discriminatory areas in China hire fewer female executives who also face a higher likelihood of dismissal and receive lower compensation than their male counterparts. This finding is supported by a recent survey in which 86% of respondents believe that gender discrimination is prevalent in Asian companies.[2] Our first research question is whether gender discrimination exists in the financial analyst industry in the U.S. and China.

We further examine whether a beauty bias exists and whether beauty has a moderating effect on gender discrimination. Prior studies document a significant beauty premium in the labor market. A large body of literature in economics and psychology suggests that people are often rewarded for their attractiveness. For example, physically attractive workers earn more than other workers (Hamermesh and Biddle 1994; Frieze, Olson, and Russell 1991) and good-looking job applicants are more likely to get calls from employers (Ruffle and Shtudiner 2015). Hamermesh, Meng, and Zhang (2002) find that beauty increases females' compensation in Shanghai, China. Liu and Sierminska (2015) find that the beauty premium in China is the highest among the countries in their sample. Besides education, skills and experience, beauty can also be a selling point when Chinese companies try to market their current employees (Human Rights Watch 2018).

On the other hand, some recent studies suggest that beauty may not always benefit employees. One example is attractive female discrimination, which is prevalent in U.S. culture. In this form of discrimination, women prefer female job applicants with low attractiveness over female applicants with high attractiveness due to intra-sexual competition (Luxen and Van de

[2] https://centres.insead.edu/emerging-markets-institute/documents/GenderReport_000.pdf

Vijver 2006). As a result, attractive female workers are perceived to face a penalty in the labor market. Agthe, Sporrle and Maner (2010) document that attractive candidates tend to be rated lower than unattractive candidates by same-sex evaluators. Ruffle and Shtudiner (2015), in an Israeli context, find that physically attractive female job candidates receive a lower rate of callbacks than other female candidates. It is, therefore, an empirical question whether beauty biases mitigate gender discrimination differently in the U.S. and China.

In this paper, we focus on financial analysts for three reasons. First, prior studies suggest that financial analysts are a male-dominated profession (Kumar 2010; Fang and Huang 2017). Although many investment banks have adopted different types of programs to promote work-life balance in an effort to attract female analysts, the impact of gender discrimination and beauty bias on opportunities for career advancement remains unknown.

Second, financial analysts are among the most important information intermediaries in the capital markets. Their profession is fiercely competitive and highly compensated in both the U.S. and China. In the U.S. each year, *Institutional Investor* magazine surveys a large number of buy-side fund managers and asks them to vote for the All-Star sell-side analysts. The voting results from these surveys are powerful determinants of sell-side analyst compensation and career advancement opportunities. In fact, Groysberg, Healy and Maber (2011) find that All-Star analysts earn 61% higher compensation than other analysts. In China, similar to the *Institutional Investor* magazine in the U.S., *New Fortune Magazine* sends surveys to buy-side money managers and conducts an annual All-Star analyst ranking for Chinese financial analysts. The base salary for the top ranked Chinese analysts is around one million U.S. dollars, which is extremely high relative to the average income level in China.[3] In addition to the quantitative attributes, such as financial

[3] China Analysts' Cut-Throat Fight for $1 Million Paycheck, *Bloomberg News*, September 28, 2017. https://www.bloomberg.com/news/articles/2017-09-28/is-an-analyst-worth-75-000-or-1-million-how-china-decides

models and earnings estimates, the voting evaluation dimensions include qualitative attributes such as industry knowledge, accessibility, responsiveness and special service. While gender and beauty are not explicit parameters in the voting process, we conjecture that they will influence voting results in this competitive profession.

Third, prior research on female financial analysts suggests that female analysts are more likely to move to high-status brokerage firms and have better career advancement in the U.S. This finding is consistent with the assertion that female analysts have stronger skills than their male counterparts due to self-selection (Kumar 2010). However, no prior studies have examined the effects of the combination of gender and beauty on career path, particularly across the two largest capital markets with different laws and cultures.

We conduct empirical analyses separately on analysts in the U.S. and China. Our final sample consists of 1,121 U.S. analysts and 442 Chinese analysts participating in an All-Star analyst voting process. Our measure of analyst beauty is based on the ratings of photos by human subjects. The photos of the U.S. analysts are extracted from their LinkedIn pages, whereas the photos of the Chinese analysts are provided by *New Fortune Magazine* and are the actual photos used for All-Star analyst voting.[4] We manually identify the gender for each analyst from these photos. We employ the Amazon Mechanical Turk (MTurk) service to rate the photos.[5] For both the U.S. and the Chinese analyst samples, each analyst photo is rated, on average, by 10 MTurk workers.[6]

---

[4] All-Star analyst voting, started by *The Institutional Investor* in 1972, has gained a significant reputation among investors over the past four decades. *New Fortune Magazine* has been hosting All-Star analyst voting in China each October since 2003. The award ceremony is one of the most significant events in China's financial industry. High resolution photos must be provided by analysts in order to register for the voting. The award was paused for a year in 2018 but will resume in 2019.

[5] MTurk is an online service through which individual workers can perform standardized tasks with compensation (e.g., Duarte, Siegel, and Young 2012).

[6] In a robustness analysis, we use 610 undergraduate and MBA business students from a major research university in North America as photo raters. For the U.S. analyst sample, each photo is rated, on average, by 20 students. For the Chinese analyst sample, each photo is rated by 67 students on average. Our results are robust to using student raters.

We first examine whether gender affects the outcome of All-Star analyst voting. We find that female analysts are more likely to be voted as All-Stars in the U.S., even though these female analysts do not have better forecasting skills than the U.S. male analysts. This latter finding is consistent with Fang and Huang (2017). Interestingly, we find that female analysts are more likely to face discrimination in China. They are less likely to be voted as All-Stars compared with male analysts, despite the fact that Chinese female analysts show better-than-average skills and have lower forecast errors than Chinese male analysts.

Next, we examine whether a beauty bias exists in All-Star analyst voting. The results show that beauty, on average, does not affect the likelihood of being voted as an All-Star analyst in the U.S. In contrast, we find a beauty premium among Chinese analysts. Good-looking analysts are more likely to be voted as All-Stars in China. We find consistent results when analyst beauty is measured based on raw quantitative scores (scale from 1 to 100), quantitative scores mean-adjusted at the rater level, qualitative scores (below average, average, attractive, very attractive), and the residual scores from regressing the mean-adjusted quantitative scores on analysts' age.

We further examine whether beauty differentially affects the voting outcomes of male and female analysts. We find that attractive female U.S. analysts are discounted in the All-Star voting, with the "female" advantage muted for those female analysts with above average looks. In contrast, we find a beauty premium among Chinese female analysts. Attractive female Chinese analysts can overcome gender discrimination and have a similar likelihood of being voted as All-Stars compared with male analysts. Given that attractive female analysts do not perform differently from other female analysts, we conclude that beauty biases work in opposite directions in the U.S. and China. Our results persist after controlling for age and a host of widely documented analyst, brokerage and firm characteristics, including forecast error, forecast frequency, forecast horizon, experience, portfolio complexity, broker size, firm size, market-to-book ratio and return on assets.

In additional tests, we carefully consider potentially correlated omitted variables and endogeneity concerns. Specifically, for omitted variables, we consider: 1) industry knowledge and social connections which may be correlated with gender and beauty; and 2) visibility which may be correlated with beauty and contribute to all-star competition success. With respect to endogeneity, we consider: 1) the potential that some brokers with more females (more beautiful) analysts are better at promoting their female (good-looking) analysts in the All-Star competitions; 2) the possibility that female analysts tend to follow specific industries; and 3) the likelihood that some good-looking Chinese analysts, with high income and/or large family net worth, may have undergone plastic surgery. In all tests, our results for gender and beauty are robust. As such, we cautiously attribute our opposing results for gender and beauty to different cultures and legal environments in the U.S. versus China. We acknowledge that we cannot fully rule out the effects of other differences between the two countries.

Our study contributes to several streams of literature. First, it provides international evidence on gender discrimination in the financial industry. There are substantial differences in the likelihood of being voted as an All-Star analyst across genders in different cultures. Although Chinese female analysts have stronger forecast ability than their male counterparts, they receive less favorable career outcomes. We do not observe the same effect in the U.S. Our finding implies that female analysts face more severe gender discrimination in China.

Second, this paper expands our knowledge about the beauty premium. Most economics and psychology studies on beauty employ experimental settings which may not be generalizable to a highly competitive profession. Our study provides large-sample empirical evidence documenting that the beauty premium is not universal. We find a beauty premium for All-Star analyst voting in China, but our results show that financial analysts with better looks do not enjoy a beauty premium in the U.S.

Finally, our study advances our understanding of the role of beauty in gender discrimination across two different countries. We find that beauty helps mute gender discrimination among female analysts in China. In contrast, we find that attractive female analysts in the U.S. financial industry are subject to a "beauty penalty." The fact that this penalty does not exist among female Chinese analysts suggests that discrimination against attractive females is not universal. These results indicate that the interaction effect of gender discrimination and beauty bias is conditional, and may be affected by culture.

Section 2 discusses the related literature and develops the hypotheses. Section 3 describes the sample and the methodology. Section 4 presents the empirical results. Section 5 discusses the additional tests. Section 6 concludes.

2. Related Literature and Hypotheses

2.1 All-Star Analyst Voting

In the U.S. each year, *Institutional Investor* magazine surveys a large number of buy-side fund managers and asks them to vote for the All-Star sell-side analysts. The vote is only open to a proprietary database, consisting of the global top fund managers of pension and hedge funds. *Institutional Investor* does not accept nominations for All-Star analysts. Rather, it allows the fund managers to vote for any sell-side participants who publish investment research and distribute this research to clients during the period covered by the poll. The voting results are solely determined by a numerical score. As per *Institutional Investor*, the final score for each financial analyst is constructed by weighting each vote based on the voter's equity and/or fixed-income assets.[7] The voting results from the surveys are powerful determinants of sell-side analysts' payment and career advancement opportunities. Groysberg et al. (2011) find that U.S. All-Star analysts earn 61% higher compensation than their peers. Brown, Call, Clement, and Sharp (2015) suggest that All-Star analysts gain more access to management and enjoy stronger bargaining power when they are promoted or change jobs. Analysts are known to lobby fund managers heavily before the voting (Hong and Kubik 2003).[8]

Turning to the All-Star analyst voting in China, *New Fortune Magazine* has conducted an annual All-Star analyst ranking for Chinese financial analysts since 2003. Similar to *Institutional Investor* in the U.S., *New Fortune* sends ballots to buy-side money managers in China each October and solicits their votes on top analysts in each industry sector based on the quality of their research and service. The buy-side firms consist of mutual funds, pension funds, insurance companies, banks, private equity and foreign investment funds (QFII). The final ranking is based on the sum of weighted votes.[9] The entire process and outcome are audited by Deloitte, one of the world's "Big Four" accounting firms. Both anecdotal evidence and formal surveys accompanying the voting process suggest that the voting outcome serves as an important determinant of Chinese analysts' compensation and future career trajectories.

The All-Star analyst rankings in the U.S. and China thus provide an ideal setting to examine whether gender discrimination and beauty bias exist in the highly competitive financial industry. We can trust the integrity of the ranking process as its importance is not lost on buy-side analysts.

---

[7] More details on the voting requirements and process can be found on the website of *Institutional Investor* magazine. http://www.institutionalinvestor.com/

[8] Anecdotal evidence suggests that analysts visit most funds to lobby for voting each year, including small funds that do not get analyst visits often and only receive reports periodically.

[9] Three analysts with the highest total scores will be recognized as stars for those industries with up to 20 analysts. Five analysts will be recognized as stars for industries with more than 20 analysts.

They rely on the performance and skills of the sell-side analysts to analyze industries and firms. Their votes make a difference in determining which sell-side analysts are promoted. We can also confidently assert that the sell-side analyst industry is indeed highly competitive. Many analysts cover the same industry sectors and the turnover rate within the industry is extremely high.

2.2 Gender Discrimination in the Labor Market

Kuhn and Shen (2013) define gender discrimination as taking an action, such as paying a different wage or choosing to hire a person, based not on that person's individual merit but on his or her gender. Many prior studies have investigated whether and where gender discrimination exists in the U.S. and in China, but few have specifically focused on the highly competitive financial industry.

Prior to the passage of the federal civil rights legislation of the 1960s, overt forms of gender discrimination, such as stating that a particular gender is preferred in job postings, were common in the U.S. (Darity and Mason 1998; Kuhn and Shen 2013). The Civil Rights Act of 1964 and subsequent related legislation have purged these overt forms of gender discrimination. However, gender discrimination has persisted in more hidden and subtle forms in American society. For example, Altonji and Blank (1999) document that American companies pay male employees more than comparable female employees. Goldin and Rouse (2000) further find that male musicians are more likely to be promoted in symphony orchestras relative to their female counterparts with similar credentials or outputs. Once gender information is hidden, the probability that a female musician will be hired and advanced significantly increases. Similarly, Egan et al. (2017) show that female financial advisers are more likely to lose their jobs and less likely to find new jobs relative to their male counterparts in the U.S. Consistent with these findings, a recent report from management consultant firm Oliver Wyman (2016) suggests that females face a glass ceiling in the financial services industry and lists it as the top ranked concern for females in the industry. Former FDIC chair Sheila Bair (2016) writes that the glass ceiling in the finance industry is barely cracked for females.

Since the financial analyst industry is highly competitive and compensated in the U.S., it is reasonable to conjecture that tangible skills sets may dominate the influence of female discrimination. Some literature supports this notion. For example, Kumar (2010) suggests that U.S. female analysts are actually more likely to move to high-status brokerage firms and have better career advancement because female analysts have stronger skills than their male counterparts due to self-selection. Li, Sullivan, Xu and Gao (2013) suggest that gender does not negatively affect female analysts' star rankings and job mobility among brokerage firms. In contrast to these findings, Fang and Huang (2017) find that male analysts receive higher benefits from social capital than female analysts in the U.S. labor market. Overall, these studies indicate that female analysts, on balance, do not have worse career advancement opportunities than male analysts in the U.S.

In contrast to the U.S., overt forms of gender discrimination remain common in China. Kuhn and Shen (2013) find that over one-third of Chinese companies seeking highly educated urban employees have one Internet job posting stating a preferred gender. These overt forms of gender discrimination decline when job skill requirements increase. Interestingly, the researchers further find that the percentage of Internet job postings preferring males versus females is approximately the same. According to INSEAD (2018), 86% of respondents believe that gender discrimination is still prevalent in companies in Asia. Recently, Human Rights Watch analyzed over 36,000 Chinese job postings on company websites and social media platforms. Many of these job postings specify a requirement or preference for males, including job postings from Chinese government departments and agencies. More specifically, 99% of job postings from Railway

Public Security Bureaus, 41% of job postings from the National Bureau of Statistics, 40% of job postings from the Maritime Safety Administration and 32% of job postings from the Civil Aviation Administration specify male preferred or male-only. Many large companies in various industries specify gender preference in their job postings as well. For example, a 2016 job posting for a film program manager at Baidu states: "strong logical reasoning ability, effective execution skills… men and manly women only." [10] An internship at Citic Group, one of China's largest state-owned investment companies, describes its ideal candidate as "high productivity, hardworking and swift, able to bear relatively high workload and pressure, male students preferred." Another job posting for a chief financial officer at an energy firm in Beijing states: "Man preferred. From any industry. Bachelor's degree or above. Age between 40 and 50." This job posting did not give any reason or explanation for the gender preference.

Besides these overt forms of gender discrimination, hidden forms of gender discrimination exist in China as well. Gao et al. (2016) document that firms headquartered in more gender-discriminatory areas hire fewer female executives in China and that these female executives receive lower compensation and face a higher likelihood of dismissal than their male counterparts. These overt or hidden forms of gender discrimination reflect the stereotype about females: that they are less intellectually, physically and psychologically capable than males, or that females are not fully committed to their jobs because some will eventually leave their positions to have a family.

Based on the preceding discussion, we predict that gender does not affect an analyst's likelihood of being voted as an All-Star in the U.S. but affects the likelihood in China. Despite the different predictions, we state our first hypothesis in null form:

[10] "Manly women" is slang in China used to refer to women showing "manly" traits, such as being strong and independent (Human Rights Watch 2018).

H1: Gender does not affect the likelihood of being voted as an All-Star analyst in the U.S. or China.

2.3 Beauty Premium in the Labor Market

As with gender preference, U.S. employers may not state their preferences regarding physical appearance in job postings since the passage of the federal civil rights legislation in the 1960s. However, a significant body of literature in economics and psychology finds a beauty premium in the U.S. labor market. Hamermesh and Biddle (1994) published a seminal work which examines the effect of beauty in the labor market. In their paper, a "beauty premium" is defined as the extra compensation earned by physically attractive workers compared with workers with below-average looks. They find that good-looking workers earn 10% to 15% more than plain-looking workers. Biddle and Hamermesh (1998) further focus on law school graduates as a specific profession and find that good-looking attorneys earn more than their less attractive peers. Deryugina and Shurchkov (2015) find that the beauty premium is more pronounced in bargaining tasks.

Prior research also suggests that beauty can predict election results. Todorov, Mandisodza, Goren and Hall (2005) show that politician appearance predicts the winner in 71.6% of U.S. Senate elections. Berggren, Jordahl and Poutvaara (2010) ask 10,011 survey respondents to rate the physical attractiveness of 1,929 Finnish political candidates. Their finding suggests that a one standard deviation increase in physical attractiveness is associated with a 20% increase in the number of votes for non-incumbent candidates. A few recent studies also document a beauty premium in corporate settings. For example, Halford and Hsu (2014) find that physically attractive CEOs are correlated with higher returns upon their job announcements and better acquirer returns on acquisition announcements. Graham, Harvey and Puri (2017) show that CEOs with a "look of competence" enjoy higher compensation.

In China, stating that a particular type of appearance is preferred in job postings is common. For example, a company in Chongqing states in their job posts that "base salary is 1,000 RMB if the candidate is taller than 160cm; base salary is 1,500 RMB if the candidate is taller than 165cm."[11] In their 2018 report, Human Rights Watch revealed that some job postings require female applicants to have certain physical attributes with respect to height, weight, voice and facial appearance, which are irrelevant to job duties. For example, a job posting in 2015 for court assistants by the Daxing District Court in Beijing states: "associate degree or above… have proper looks… clear enunciation…. and under 35-year-old, female." A job posting for a receptionist internship position at the recruiting company, Zhilian Zhaopin, states: "nice appearance and temperament, sweet and beautiful voice."

Other postings use the physical attributes of companies' current female employees to attract male applicants. For example, some of China's giant technology companies, such as Tencent, Baidu and Alibaba, have repeatedly published job posts boasting that there are "beautiful female colleagues" or "goddesses" working for their firms (Human Rights Watch Report 2018).[12] Besides education, skills and experience, beauty can also be a selling point when companies try to market their employees. Companies and media often use phrases such as "beauty CEO" or "beauty analyst" to describe their employees.[13] Consistent with anecdotal evidence, Liu and Sierminska (2015) find that the beauty premium in China is the highest among the countries in their sample.

[11] http://news.ifeng.com/society/2/detail_2007_11/06/853753_0.shtml
[12] https://www.hrw.org/sites/default/files/report_pdf/china0418_web.pdf;
https://www.hrw.org/news/2018/04/26/chinese-tech-companies-delete-gender-discriminatory-job-ads;
https://www.hrw.org/news/2018/11/08/china-female-civil-servants-face-discrimination-harassment
[13] http://finance.sina.com.cn/money/fund/jjzl/2016-04-14/doc-ifxriqqx2391894.shtml

Prior studies provide several potential explanations for the beauty premium. First, the beauty premium may be driven by discrimination from employers and/or customers. For example, physically attractive workers are wrongly considered more capable. In addition, customers or other related parties can also overestimate the abilities of physically attractive workers and thus increase beauty's productivity as a result. Second, beauty may actually be an indirect measure of unobservable skill or ability. For example, Mobius and Rosenblat (2006) find that physically attractive workers have better oral skills, which contribute to the beauty premium. Third, self-selection can drive the beauty premium. Good-looking workers are more likely to self-select into high-paying professions.

Based on the preceding discussion, we predict that beauty increases an analyst's likelihood of being voted as an All-Star both in the U.S. and in China. We state our second hypothesis in null form:

H2: Beauty does not affect the likelihood of being voted as an All-Star analyst in the U.S. or China.

2.4 Interaction Effect of Gender and Beauty in the Labor Market

Beauty may not always benefit employees. Luxen and Van de Vijver (2006) find that women (both human resource professionals and students) prefer female job applicants with low attractiveness over female applicants with high attractiveness. Agthe et al. (2010) document that attractive candidates tend to be rated lower by same-sex evaluators than unattractive candidates. Even though the financial analyst profession is male-dominated and such discrimination by same-sex evaluators may not apply to female analysts, prior studies in the U.S. suggest that beautiful females are more likely to be egotistical, snobbish and have unsympathetic personalities. Andreoni and Petrie (2008) find experimental evidence that beauty has different effects on male and female subjects.

In a Chinese context, Kuhn and Shen (2013) find that many Chinese companies have explicit requests for age, height and beauty in their Internet job postings and these requests for attractiveness are highly correlated with explicit requests for female applicants. As another example, Hamermesh et al. (2002) find that beauty increases females' compensation in Shanghai, China. We thus generate our third null hypothesis:

H3: There is no differential impact of beauty on the relationship between gender and the likelihood of being voted as an All-Star analyst in the U.S. or China.

3. Sample Selection and Beauty Measure

3.1 Sample Selection

We use a sample of U.S. analysts and a sample of Chinese analysts to test our hypotheses. Panel A of Table 1 shows the sample selection procedure for the U.S. analyst sample. We first retrieve the analysts' names from the Thomson Reuters Institutional Brokers' Estimate System (I/B/E/S) recommendation file, collect their gender information and photos from their LinkedIn profiles and then retain the analysts with high-quality photos for the rating process. The resulting analyst photos are reasonably standardized, with 1) 94% of the photos featuring head and upper body, 2) 93% of the photos being color photos and 3) 90% of the analysts wearing a suit/shirt/dress.[14] Specifically, among the 4,377 names of U.S. sell-side analysts available from I/B/E/S for the year 2014, we are able to identify 1,427 analysts with LinkedIn profiles and high-quality photos. These 1,427 analysts map to 3,695 analyst-years for the period 2013 to 2015. The

[14] In a robustness check, we repeat our empirical analysis with the analyst photos that meet all three conditions. The inferences remain unchanged.

final U.S. analyst sample consists of 2,709 analyst-years, representing 1,121 distinct analysts. The smaller sample size is due to a combination of factors, including 1) I/B/E/S data restrictions, 2) Compustat data restrictions and 3) the requirement for at least one All-Star analyst in our sample for a given industry and year.

Panel B of Table 1 shows the similar sample selection procedure for the Chinese analysts. We start with 922 candidates who are registered in the Chinese star analyst voting for 2015. *New Fortune* magazine exclusively provides us with the names and photos of these candidates, with the photos being the actual photos used in the All-Star voting. After excluding candidates that hold a sales manager title or work in the macro-economy, strategy, financial engineering or fixed income sectors, we are left with 585 sell-side analysts. There are 1,274 analyst-years associated with these analysts, based on the information available in the GTA China Stock Market & Accounting Research (CSMAR) database, for the period 2013 to 2015. After excluding observations with group photos and meeting data restrictions similar to those applied to the U.S. analyst sample, we arrive at the final China analyst sample, consisting of 960 analyst-years representing 442 unique analysts for the period 2013 to 2015.

3.2 Measuring Beauty

Our measure of analyst beauty is based on the analysts' facial attractiveness as perceived by human raters. Each analyst photo is rated on two complementary dimensions: (1) quantitative: a scale from 1 to 100; and (2) qualitative: below average, average, attractive and very attractive.

3.2.1 Methodology

The ratings of analyst beauty are obtained from Amazon Mechanical Turk (MTurk), a crowdsourcing Internet marketplace that enables businesses and individuals to coordinate the use of human intelligence to perform tasks. In this marketplace, workers can browse jobs and complete them for a monetary payment set by the employer. As noted above, each analyst photo is rated on both a quantitative and a qualitative dimension.

Each analyst photo is rated, on average, by 10 MTurk workers. The actual number of ratings per photo varies slightly because a random number generator is used to select photos for each rater. We measure analyst beauty as the average of the independent quantitative scores received for the analyst, after excluding raters of inconsistent rating quality and dropping the highest and lowest rating for each analyst.[15] The use of a composite rating is consistent with prior work, which shows that the estimated coefficients on beauty are smaller when based on the evaluations of a single rater rather than a composite measure. Composite measures are more reliable because they are based on aggregations of correlated responses.

One potential issue with the raw quantitative beauty measure is that each rater may have different benchmarks for beauty, which would add noise to the measure. To address this concern, in the main analyses, we use the quantitative scores mean-adjusted at the individual rater level to proxy for analyst beauty. Specifically, we subtract the mean quantitative score given by a rater from each quantitative score received from the same rater. Next, we recalculate the average of such mean-adjusted quantitative scores for each analyst.[16] In un-tabulated sensitivity analyses, we complement the mean-adjusted quantitative beauty measure with the alternative beauty measures based on the raw quantitative scores and the raw qualitative scores. The latter alternative measure is calculated as the average independent qualitative rating received for each analyst (i.e., we code

[15] To control the quality of rating results, we only include raters' scores in our final sample if their ratings are of consistent quality. We proxy for consistent quality in two ways: (1) the correlation between quantitative and qualitative ratings for a given rater is at least 0.60; and (2) the standard deviation of quantitative scores for all photographs coded by an individual is at least 6 (quantitative scores range from 0 to 100). While these cutoffs are admittedly somewhat arbitrary, they seem reasonable based on our review of the raw data.
[16] The mean-adjusted quantitative scores remove the potential effects of a rater's demographic characteristics. See next subsection for details.

"below average" as 1, "average" as 2, "attractive" as 3, and "very attractive" as 4). This alternative beauty measure also deals with the concern that raters may give different quantitative scores to analysts. In addition to the MTurk ratings, we obtain students' ratings as a robustness check. The student subjects consist of 610 undergraduate and MBA business students from a major research university in North America.[17]

3.2.2 Summary Demographics of Raters

The average age of the MTurk raters is 36 with 57% being male. Regarding ethnicity, 64% of the raters are white/Caucasian, 9% are African-American and 20% are Asian. To examine the effect of these demographic characteristics on the ratings of analyst beauty, we regress the raw quantitative scores of analyst beauty on raters' age, gender and ethnicity. The results show that the raters' age is positively associated with the raw quantitative scores of analyst beauty (p-value < 0.1), but their gender and ethnicity have no significant influence. These results are consistent with previous research that suggests little cross-cultural variation in people's perceptions of which facial characteristics are attractive (e.g., Langlois, Kalakanis, Rubenstein, Larson, Hallam, and Smoot 2000; Perrett, May, and Yoshikawa 1994). Importantly, when we regress the mean-adjusted quantitative scores of analyst beauty on the raters' age, gender and ethnicity, none of these demographic characteristics has a significant effect. As such, we believe that the mean-adjusted quantitative beauty measure is sufficiently unbiased. In addition, as raters are highly unlikely to know the identities of the individuals they are rating, we are not concerned that familiarity will bias the results.

[17] These students were sourced from an introductory managerial accounting course, an advanced financial accounting course and an accounting theory course. Ethics clearance was obtained from the Research Ethics Board of the university. For the U.S. analyst sample, each photo is rated by 20 students on average. For the China analyst sample, each photo is rated by an average of 67 students.

4. Hypothesis Tests

4.1 Test of H1: Does Gender Affect the Likelihood of Being Voted as an All-Star Analyst?

4.1.1 Empirical Specification

Our first research question is whether gender discrimination exists in the All-Star analyst voting process in the U.S. and China. Before testing the first hypothesis, we explore whether the male and female analysts perform differently on earnings forecasts. Although it is not our primary interest, the test follows the theme of the analyst literature in accounting and the results may help us distinguish between: 1) the analyst ability explanation (due to self-selection of female analysts) and 2) the gender discrimination explanation if a significant gender effect is later observed in the All-Star analyst voting.

We use average relative earnings forecast error to proxy for an analyst's performance and regress this proxy on gender, controlling for the analyst's age, earnings forecasting activities, research portfolio complexity, experience, brokerage firm resources and reputation and the characteristics of the firms followed. Specifically, we estimate the following OLS model:

$$\begin{aligned} AFE_{i,t} = {} & \beta_0 + \beta_1 \cdot Female_i + \beta_2 \cdot Age_{i,t} + \beta_3 \cdot Horizon_{i,t} + \beta_4 \cdot Freq_{i,t} + \beta_5 \cdot BSize_{i,t} \\ & + \beta_6 \cdot NFirm_{i,t} + \beta_7 \cdot NInd_{i,t} + \beta_8 \cdot GExp_{i,t} + \beta_9 \cdot FExp_{i,t} + \beta_{10} \cdot Size_{i,t} \\ & + \beta_{11} \cdot MTB_{i,t} + \beta_{12} \cdot ROA_{i,t} + Industry\text{-}Year\ Fixed\ Effects + \varepsilon_{i,t}\,. \end{aligned} \tag{1}$$

AFE is the analyst's average relative forecast error for the firms followed in year t, where the relative forecast error is the analyst's most recent forecast error (i.e., |analyst earnings forecast − actual earnings|) relative to the most recent forecast errors of all analysts following the firm in year t and ranges from 0 to 100 (Clement and Tse 2003).[18] *Female* is an indicator variable set to one for female analysts and zero for male analysts.

For the U.S. analyst sample, we measure an analyst's age based on the year the analyst graduated from college or university. Specifically, we proxy for analyst's age as: [22 + (current year – graduation year)]. We are able to collect the education backgrounds from LinkedIn for 80% (i.e., 892 out of 1,121) of the analysts. We then estimate the linear relationship between age, general experience as an analyst (*GExp*), *Female* and the interaction term of *GExp* and *Female* for these analysts in an effort to estimate the ages of analysts with missing education information. For the China analyst sample, *New Fortune Magazine* provides us with the age information for 76% (i.e., 337 out of 442) of the analysts. We collect the education backgrounds of the remaining analysts from the Securities Association of China website.[19] We estimate an analyst's age as [22/24/27 + (current year – graduation year)], depending on whether the analyst's highest degree is a Bachelor/Master/Doctoral degree.

Regarding the other control variables, earnings forecast horizon (*Horizon*) is the analyst's average relative forecast horizon for the firms followed in year t, where the relative forecast horizon is the analyst's forecast horizon (i.e., the number of days between the analyst's earnings forecast and the firm's actual earnings announcement) relative to the forecast horizons of all analysts following the firm in year t and ranges from 0 to 100. Earnings forecast frequency (*Freq*) is the analyst's average relative forecast frequency for the firms followed in year t, where the relative forecast frequency is the analyst's forecast frequency relative to the forecast frequencies

[18] We multiply all relative measures by 100 to make the coefficient estimates more interpretable.
[19] http://exam.sac.net.cn/pages/registration/sac-publicity-report.html (in Chinese).

of all analysts following the firm in year t and ranges from 0 to 100. The analyst's research portfolio complexity is measured by the number of firms followed (*NFirm*) and the number of industries followed (*NInd*) in year t. The analyst's experience is measured by his or her general experience as an analyst (*GExp*) as well as the average number of years that the analyst has followed the firms in the research portfolio (*FExp*) (Clement 1999). In addition, we control for brokerage firm resources and reputation by including brokerage firm size (*BSize*) in the model, measured by the number of analysts employed by the brokerage firm in year t. We further control for the characteristics of an average firm in the analyst's research portfolio in year t. These characteristics include: 1) average firm size (*Size*), measured as the mean of the natural logarithm of market value of the firms followed, 2) average market-to-book ratio of the firms followed (*MTB*) and 3) average return on assets of the firms followed (*ROA*) by the analyst in year t. These firm characteristics reflect analysts' coverage selection and to some extent the potential market impact of their equity research. Finally, we control for industry-year fixed effects to facilitate a within-industry comparison.

To test H1, we regress an analyst's All-Star award status on gender, controlling for the analyst's age, performance, earnings forecasting activities, research portfolio complexity and experience, brokerage firm resources and reputation, and the characteristics of the firms followed. Specifically, we estimate the following probit model:

$$\begin{aligned} Star\_Award_{i,t} = {} & \beta_0 + \beta_1 \cdot Female_i + \beta_2 \cdot Age_{i,t} + \beta_3 \cdot CAR_{i,t} + \beta_4 \cdot AFE_{i,t} + \beta_5 \cdot Horizon_{i,t} \\ & + \beta_6 \cdot Freq_{i,t} + \beta_7 \cdot BSize_{i,t} + \beta_8 \cdot NFirm_{i,t} + \beta_9 \cdot NInd_{i,t} + \beta_{10} \cdot GExp_{i,t} \\ & + \beta_{11} \cdot FExp_{i,t} + \beta_{12} \cdot Size_{i,t} + \beta_{13} \cdot MTB_{i,t} + \beta_{14} \cdot ROA_{i,t} \\ & + Industry\text{-}Year\ Fixed\ Effects + \varepsilon_{i,t}, \end{aligned} \tag{2}$$

where *Star_Award* denotes All-Star analyst award status, an indicator variable set to one if a U.S. (Chinese) analyst is ranked in the top three or runner-up by *Institutional Investor* (*New Fortune Magazine*), and zero otherwise. We additionally control for the price impact of the analyst's stock recommendations (*CAR*), measured as the average three-day abnormal cumulative stock returns surrounding the analyst's stock recommendations for the firms covered in year t.[20] Finally, we include industry-year fixed effects to alleviate the concern that analysts with certain characteristics tend to be assigned to cover specific industries. *Female* is our variable of interest. All other control variables are as previously defined.

4.1.2 Results

Table 2 presents descriptive statistics for the variables used in the empirical analysis. In the U.S. (China) analyst sample, 9% (37%) of the analysts are awarded All-Star status and 9% (24%) of the analysts are female.[21] Figure 1 shows that the percentages of female analysts for both samples are relatively stable over time. Regarding the beauty measures, the average raw quantitative beauty measure is 49.05 (49.61). On average, analysts cover 13 (20) firms and 4 (5) industries and have 8.5 (2.2) years of experience as financial analysts. The continuous variables are winsorized at the top and bottom 1%. The standard errors are clustered by brokerage firm and by year.

[20] In our robustness check, we calculate *CAR* separately for recommendation upgrades/downgrades and reiterations and include both in the model for the U.S. sample. We are not able to conduct this separation for the China sample as 84.3% of Chinese analysts either: 1) issue reiteration recommendations only or 2) only issue recommendation once a year.

[21] The percentage of All-Star analysts is different between the two samples mainly because we use all analysts in I/B/E/S to construct the U.S. analyst sample and the analyst candidates in the voting ballot (rather than all analysts in CSMAR) to construct the China analyst sample.

Table 3 reports the results from estimating model (1). In the U.S. analyst sample, we find that the coefficient estimate on *Female* is statistically insignificant, suggesting that the U.S. female analysts do not perform better than their male counterparts. This result is consistent with Fang and Huang (2017). In contrast, in the China analyst sample, we find that the coefficient estimate on *Female* is significant and negative (p-value < 0.05), suggesting that the Chinese female analysts in our sample tend to perform better than their male counterparts. In economic terms, the female analysts in China are 2% more accurate than their male counterparts, which is approximately 6% of the sample mean.

Table 4 reports the results from estimating model (2). In the U.S. analyst sample, we find a significant and positive coefficient estimate on *Female* (p-value < 0.1). This result suggests that, after controlling for age and analyst performance, the U.S. female analysts in our sample are more likely than their male counterparts to be voted as All-Stars. The marginal effect of *Female* in column 1 indicates that the probability of being voted as a star analyst is higher by 1.0%, approximately 11% of the sample mean (Table 2). In contrast, in the China analyst sample, we find a significant and negative coefficient estimate on *Female* (p-value < 0.01), suggesting that the Chinese female analysts are discriminated against in the All-Star analyst voting. The marginal effect of *Female* in column 2 indicates that the probability of being voted as a star analyst is lower by 3.3%, which is approximately 9% of the sample mean. Collectively, these results reject H1 and imply that female analysts face significant discrimination in the All-Star analyst voting in China whereas the effect is opposite in the U.S.

Turning to the control variables, we document some common determinants of All-Star analyst award status in both the U.S. and China. Analysts are more likely to be voted as All-Stars when they work for larger brokerage firms, follow more firms but fewer industries and follow larger firms. Nevertheless, some differences also exist. Older analysts are more likely to be voted as All-Star in the U.S. but not in China. In addition, the voters in the U.S. seem to place more weight on analysts' performance than the voters in China do.

Overall, we find that female analysts are discriminated against in opposing directions in the two countries. In particular, the voters in the U.S. discriminate in favor of female analysts, whereas the voters in China discriminate against female analysts.

4.2 Test of H2: Does Beauty Affect the Likelihood of Being Voted as an All-Star Analyst?

4.2.1 Summary Statistics for Beauty Measures

Panel A of Table 5 reports the summary statistics of the raw quantitative beauty scores based on either MTurk or student ratings. We find that both the mean and median of the MTurk ratings are approximately 50 for both the U.S. and China analyst samples. In contrast, we find that students' ratings, on average, are downward-biased, with mean raw scores of approximately 44 (41) for the U.S. (China) analyst sample. We consider the MTurk ratings to be more diversified and representative for our empirical analyses, so we only report results based on MTurk ratings. Importantly, while students are systemically more conservative than MTurk workers in their ratings, the relative rankings for analysts are remarkably consistent across both MTurk raters and student raters.[22]

[22] In Section 3.2.2, we document that the raters' age is positively associated with the raw quantitative scores of analyst beauty, which may explain why students' ratings are more conservative. The correlations between the MTurk and student ratings are +0.70 for the U.S sample, and +0.86 for the China sample.

Panel B of Table 5 reports the summary statistics of the raw quantitative beauty scores, age and general experience by gender and by All-Star analyst award status. When classifying analysts by gender, we find that female analysts, on average, are perceived to be better looking than male analysts. The average beauty scores for male and female U.S. (China) analysts are 47.94 and 59.93 (45.53 and 62.33), respectively, and this difference is statistically significant (p-values < 0.01). Next, when classifying analysts by All-Star analyst award status, we find that there is a significant difference between the raw quantitative beauty scores of star and non-star Chinese analysts (p-value < 0.1), but not for U.S. analysts. In the U.S., female analysts tend to be younger and less experienced than male analysts, and star analysts tend to be older and more experienced than non-star analysts. We do not observe such differences, however, in China.

Since the focus of the paper is gender and beauty, we further explore whether female analysts would choose specific industries to follow. Panel C of Table 5 reports the industries with the highest percentages of female analyst coverage. The results suggest that brokerage firms indeed tend to assign female analysts to cover certain industries. Based on the industries defined by *Institutional Investor* during our sample period, the industries with highest female analyst coverage in the U.S. analyst sample are: (1) food, beverage, and household and personal care products, (2) retailing/department stores and specialty soft lines and (3) retailing/food and drug chains. In contrast, based on the industries defined by *New Fortune* magazine for the China sample, the industries with highest female analyst coverage are: (1) textile garment and apparel, (2) services (hotel, restaurants, leisure) and (3) papermaking and paper products. Industries with significant overlap for female analysts across the two countries include clothing and service.

4.2.2 Empirical Specification

The second research question is whether a beauty bias exists in the All-Star analyst voting process in the U.S. and in China. As in Section 4.1., we follow prior analyst literature and use an analyst's average relative earnings forecast error to proxy for the analyst's performance and examine its association with beauty. Specifically, we estimate the following OLS model:

$$\begin{aligned} AFE_{i,t} = {} & \beta_0 + \beta_1 \cdot Beauty_i + \beta_2 \cdot Age_{i,t} + \beta_3 \cdot Horizon_{i,t} + \beta_4 \cdot Freq_{i,t} + \beta_5 \cdot BSize_{i,t} \\ & + \beta_6 \cdot NFirm_{i,t} + \beta_7 \cdot NInd_{i,t} + \beta_8 \cdot GExp_{i,t} + \beta_9 \cdot FExp_{i,t} + \beta_{10} \cdot Size_{i,t} \\ & + \beta_{11} \cdot MTB_{i,t} + \beta_{12} \cdot ROA_{i,t} + Industry\text{-}Year\ Fixed\ Effects + \varepsilon_{i,t}\,, \end{aligned} \tag{3}$$

where *Beauty* is the analyst's facial attractiveness, measured as the average of the mean-adjusted quantitative scores received for the analyst, as explained in Section 3. When examining the incremental effect of *Beauty*, it is especially important to control for *Age* because there exists a negative correlation between *Age* and *Beauty*: the correlations are -0.25 and -0.24 for the U.S. male and female analysts, respectively, and -0.31 and -0.29 for the Chinese male and female analysts. All other variables are as previously defined.

To test H2, we regress an analyst's All-Star award status on beauty, again controlling for the analyst's age, performance, earnings forecasting activities, research portfolio complexity and experience, brokerage firm resources and reputation and the characteristics of the firms followed. Specifically, we estimate the following probit model:

$$\begin{aligned} Star\_Award_{i,t} = {} & \beta_0 + \beta_1 \cdot Beauty_i + \beta_2 \cdot Age_{i,t} + \beta_3 \cdot CAR_{i,t} + \beta_4 \cdot AFE_{i,t} + \beta_5 \cdot Horizon_{i,t} \\ & + \beta_6 \cdot Freq_{i,t} + \beta_7 \cdot BSize_{i,t} + \beta_8 \cdot NFirm_{i,t} + \beta_9 \cdot NInd_{i,t} + \beta_{10} \cdot GExp_{i,t} \\ & + \beta_{11} \cdot FExp_{i,t} + \beta_{12} \cdot Size_{i,t} + \beta_{13} \cdot MTB_{i,t} + \beta_{14} \cdot ROA_{i,t} \\ & + Industry\text{-}Year\ Fixed\ Effects + \varepsilon_{i,t}\,, \end{aligned} \tag{4}$$

where *Beauty* is our variable of interest. All other control variables are as previously defined. H2 predicts β_1 to be insignificant.

4.2.3 Results

Table 6 reports the results from estimating model (3). In both the U.S. and China analyst samples, we find an insignificant coefficient estimate on *Beauty*, suggesting that good-looking analysts on average do not perform better in earnings forecasts.

Table 7 reports the results from estimating model (4). In the U.S. analyst sample, we find that the coefficient estimate on *Beauty* is statistically insignificant, suggesting that good-looking analysts do not benefit from their facial attractiveness in the All-Star analyst voting. However, in the China analyst sample, we find a significant and positive coefficient estimate on *Beauty* (p-value < 0.01), suggesting that good-looking analysts are more likely to be voted All-Star analysts.[23] In economic terms, a one standard deviation increase in *Beauty* is associated with a 3.2% increase in the probability of being voted as a star analyst, which is approximately 9% of the sample mean. Collectively, these results reject H2.

Although the prior literature suggests that physical attractiveness is correlated with many positive attributes such as ambitiousness, industriousness, confidence, popularity and intelligence (Eagly, Ashmore, Makhijani, and Longo 1991; Feingold 1992), we do not observe a beauty premium in the U.S. analyst sample. The reason for this lack of a beauty premium is likely that the background and record of each analyst's relative performance are readily available to voters, so voters may place more weight on direct indicators of analysts' abilities (e.g., research output) than on indirect indicators (e.g., facial attractiveness). In contrast, the analysts in the China sample are

---

[23] All of these results hold when we replace *Beauty* with the raw quantitative or the qualitative beauty measure.

relatively young and inexperienced, with an average general experience of 2.2 years, so the voters may still place some weight on indirect indicators of analysts' abilities, such as analyst beauty.

Overall, our results suggest that, on average, the beauty premium in All-Star analyst voting exists in China but not in the U.S. As such, we conclude that the beauty premium is not a universal phenomenon.

4.3. Test of H3: Does an Interaction Effect Between Gender and Beauty Exist in the All-Star Analyst Voting?

4.3.1 Empirical Specification

We further examine whether there is an interaction effect between beauty and gender on the likelihood of receiving an All-Star analyst award, given the competitive and professional nature of the financial analyst industry. Before testing H3, we first augment model (1) by including the interaction effect between *Female* and *Beauty* and the interaction effect between *Male* and *Beauty.* We include *Beauty* separately for female and male analysts because the mean beauty scores of the two groups are significantly different.[24] Specifically, we estimate the following OLS model:

$$\begin{aligned} AFE_{i,t} = \; & \beta_0 + \beta_1 \cdot Female_i + \beta_2 \cdot Female_i \times Beauty_i + \beta_3 \cdot Male_i \times Beauty_i + \beta_4 \cdot Age_{i,t} \\ & + \beta_5 \cdot Horizon_{i,t} + \beta_6 \cdot Freq_{i,t} + \beta_7 \cdot BSize_{i,t} + \beta_8 \cdot NFirm_{i,t} + \beta_9 \cdot NInd_{i,t} \\ & + \beta_{10} \cdot GExp_{i,t} + \beta_{11} \cdot FExp_{i,t} + \beta_{12} \cdot Size_{i,t} + \beta_{13} \cdot MTB_{i,t} + \beta_{14} \cdot ROA_{i,t} \\ & + Industry\text{-}Year\ Fixed\ Effects + \varepsilon_{i,t}\,. \end{aligned} \tag{5}$$

[24] The inferences are similar if we directly interact *Female* and *Beauty,* or if we estimate model (3) and (4) separately for male and female analysts.

To test H3, we augment model (2) by including the interaction effect between *Female* and *Beauty* and the interaction effect between *Male* and *Beauty.* Specifically, we estimate the following probit model:

$$\begin{aligned} Star\_Award_{i,t} = \ & \beta_0 + \beta_1 \cdot Female_i + \beta_2 \cdot Female_i \times Beauty_i + \beta_3 \cdot Male_i \times Beauty_i \\ & + \beta_4 \cdot Age_{i,t} + \beta_5 \cdot CAR_{i,t} + \beta_6 \cdot AFE_{i,t} + \beta_7 \cdot Horizon_{i,t} + \beta_8 \cdot Freq_{i,t} \\ & + \beta_9 \cdot BSize_{i,t} + \beta_{10} \cdot NFirm_{i,t} + \beta_{11} \cdot NInd_{i,t} + \beta_{12} \cdot GExp_{i,t} \\ & + \beta_{13} \cdot FExp_{i,t} + \beta_{14} \cdot Size_{i,t} + \beta_{15} \cdot MTB_{i,t} + \beta_{16} \cdot ROA_{i,t} \\ & + Industry\text{-}Year\ Fixed\ Effects + \varepsilon_{i,t}, \end{aligned} \tag{6}$$

All variables are as previously defined.

4.3.2. Results

Table 8 reports the results from the estimation of model (5). In the U.S. analyst sample, we continue to find an insignificant coefficient estimate on *Female*, and we find no significant interaction between gender and beauty on U.S. analysts' performance. In the China analyst sample, we continue to find a significant and negative coefficient estimate on *Female* (p-value < 0.1), again suggesting that Chinese female analysts tend to perform better than their male counterparts. Meanwhile, we find that there is no significant interaction between gender and beauty on Chinese analysts' performance.

Table 9 reports the results from the estimation of model (6). In the U.S. analyst sample, we continue to find a significant and positive coefficient estimate on *Female* (p-value < 0.05), suggesting that U.S. female analysts on average are treated more favorably in the All-Star analyst voting. We also find a significant and negative interaction effect between *Female* and *Beauty* (p-value < 0.01), suggesting that good-looking female analysts not only receive no premium from their facial attractiveness but are also penalized in the All-Star analyst voting. In economic terms, *Female* is associated with a 1.5% increase in the probability of being voted as a star analyst. Based on the distribution of *Beauty*, the effect of *Female* is muted for a female analyst with the facial attractiveness in the top decile among all female analysts in the U.S.

As a sharp contrast, in the China analyst sample, we find a significant and negative coefficient estimate on *Female* (p-value < 0.01) and a significant and positive coefficient estimate on the interaction term of *Female* and *Beauty* (p-value < 0.01). This evidence suggests that while there is gender discrimination against Chinese female analysts in the All-Star analyst voting, good-looking female analysts are subject to less discrimination. In economic terms, *Female* is associated with a 13.2% decrease in the probability of being voted as a star analyst. Based on the distribution of *Beauty*, a female analyst with the facial attractiveness in the top quartile among all female analysts in China can completely mitigate the gender discrimination.[25]

In sum, these results reject H3 and suggest that gender and beauty discrimination function differently in two different countries.

5. Possible Explanations and Additional Tests

5.1. Brokerage Firms' Promotion Efforts

In an un-tabulated analysis, we observe variations in the percentage of female analysts and the average beauty ratings of analysts across brokerage firms in both the U.S. and China samples. As such, our results could be biased by the endogenous matching between analysts and brokerage firms. In other words, the female premium observed in the U.S. could be driven by brokerage firms

[25] In this regression, we control for *Age*. In an un-tabulated analysis, we regress *Beauty* on *Age* and use the residual to proxy for analysts' facial attractiveness. The results are robust.

who are better at promoting female analysts, and the female beauty premium observed in China could be driven by brokerage firms who are better at promoting good-looking analysts.

To address this concern, we construct two measures of brokerage firms' promotion efforts. The first measure, *Female%*, is the percentage of female analysts for a given broker and year. The second measure, *Avg Beauty*, is the mean of *Beauty* of all analysts for a given broker and year. The rationale is that the promotion efforts for female or good-looking analysts, as well as the net benefits of doing so, should be a function of how many of such analysts a brokerage firm has. Therefore, we augment model (6) by additionally controlling for *Female%* and *Avg Beauty*. In un-tabulated results, we find an insignificant effect of *Female%* and a significant and positive effect of *Avg Beauty* on *Star_Award* for both samples (p-value < 0.05 or better). While the latter result is consistent with brokerage firms' promotion efforts, the results of *Female* and its interaction effect with *Beauty* are robust to the findings reported in Table 9.[26]

5.2 Industry Expertise

Analysts' industry knowledge has been ranked by buy-side clients as one of the most important analyst attributes, ahead of earnings forecasts and stock selection (Groysberg et al. 2011; Brown et al. 2015). Therefore, while we control for analysts' earnings forecast accuracy and price impact of stock recommendations in the tests of All-Star status, we seek to further control for analysts' industry expertise to address the possibility that female analysts in the U.S. and China may possess different amounts of industry knowledge, thus leading to differential voting outcomes. We proxy for analysts' industry expertise by: (1) *Industry_Exp*, defined as the number of years an

[26] All un-tabulated results discussed in Section 5 will be tabulated and available upon request and through an online appendix if the paper is accepted.

analyst has followed her main industry, (2) *Industry_Cover%*, defined as the percentage of firms in an analyst's main industry that are covered by the analyst in the year, (3) *Industry_EPS%*, defined as the percentage of earnings forecasts in an analyst's main industry that are issued by the analyst in the year and (4) *Industry_Rec%*, defined as the percentage of stock recommendations in an analyst's main industry that are issued by the analyst in the year.[27] We then augment model (6) by additionally controlling for these measures of industry expertise, one at a time. In the un-tabulated results, we find that all four measures have a significant and positive effect on *Star_Award* for both samples (p-value < 0.1 or better). Importantly, the results of *Female* and its interaction effect with *Beauty* remain inferentially similar to those reported in Table 9.

5.3 Differential Social Networking Effects for Female and Male Analysts

The financial analyst profession is a male-dominated profession in both the U.S. and China. Consistent with this, Fang and Huang (2017) find that male analysts receive higher benefits from social capital than female analysts in the U.S. labor market. Therefore, it is possible that the discrimination on gender and beauty impacts the All-Star voting outcome through the social network. While it is ex-ante unclear whether the social networking effects for female and male analysts would apply differently to the U.S. versus China, we seek to answer this question empirically.[28]

[27] The *Industry_Rec%* measure also captures the effect of analysts' reiterations of existing stock recommendations, which tend to be associated with smaller market reactions and thus may not be fully captured by the *CAR* measure.

[28] This analysis also helps to control for analyst performance. Analysts with better social connections have more resources. Therefore, they are more likely to arrange client access to corporate managers and are also more likely to provide better service to clients. In this sense, the social connection between analysts and managers can be viewed as a performance measure to proxy for the ability to arrange client access to corporate managers and to provide special services such as corporate site visits. In a China setting, Gu et al. (2019) find that the social connection between analysts and the buy-side (i.e., fund managers) contributes to analyst ratings and analyst compensation.

Following Cohen, Frazzini and Malloy (2010) and Fang and Huang (2017), we identify the U.S. analysts who graduated from the top 100 universities in the U.S. and the top five universities in Canada. We then proceed to collect the education background of all of the executives and directors of the companies followed by these analysts. We match the schools of these analysts with the schools of the executives and directors. We then proceed to generate a school tie measure whereby the value of *SocialTies* is one if the analyst attended the same institution with any executive or director of the covered firms, and zero otherwise. Similarly, we collect the education background of Chinese analysts and all of the fund managers investing in the companies followed by these analysts, and generate a school tie measure whereby the value of *SocialTies* is one if the analyst attended the same institution with any fund manager investing in the covered firms, and zero otherwise (Gu, Li, Li, and Yang 2019). We then augment model (6) by additionally controlling for *SocialTies*. In the un-tabulated results, we find a significant and positive effect of *SocialTies* on *Star_Award* for both samples (p-value < 0.05 or better). The results of *Female* and its interaction effect with *Beauty* are inferentially similar to those reported in Table 9.

5.4 Does Gender and Beauty Discrimination Vary with the Extent of Female Analyst Presence?

Panel C of Table 5 reports that the extent of female analyst coverage varies across industries. As such, a natural question to ask is whether the gender and beauty discrimination toward female analysts is stronger or weaker in the more female-concentrated industries. To examine this issue, we divide the samples into two subsamples: 1) the top 3 industries with the highest percentage of female analysts and 2) all other industries. We then re-estimate model (6). The un-tabulated results show that, for the U.S. analyst sample, the gender and beauty biases toward females are not statistically significant in the top 3 industries largely covered by female analysts. For the China analyst sample, the gender bias is not statistically significant and the beauty bias is statistically weaker in the top 3 industries largely covered by female analysts. In other words, these results suggest that gender and beauty biases are less (more) pronounced in the less (more) male-dominated industries. However, these findings do not explain why gender and beauty discrimination function differently in the two countries.

5.5 Alternative Beauty Measure for the China Analyst Sample

Although previous research shows little cross-cultural variation in people's perceptions of which facial characteristics are considered attractive (e.g. Langlois et al., 2000; Perrett et al. 1994), we recognize that the ethnicity of MTurk raters is not representative of the ethnicity of voters for the Chinese All-Star analyst awards. This mismatch could potentially bias our *Beauty* measure and the empirical results. To address this concern, we recalculate the mean-adjusted quantitative beauty score for each Chinese analyst based on the ratings received from all Asian MTurk and student raters, and then re-estimate the regressions. The un-tabulated results from estimating model (4) show a significant and positive coefficient estimate on *Beauty* (p-value < 0.01). The un-tabulated results from estimating model (6) show a significant and negative coefficient estimate on *Female* (p-value < 0.01) and a significant and positive coefficient estimate on *Female*×*Beauty* (p-value < 0.1). Our main results are thus robust to the modified beauty measure for the China analyst sample.

5.6 Beauty Premium and Analyst Visibility

For the U.S. analyst sample, an alternative explanation for the lack of a beauty premium may be due to the visibility of financial analysts.[29] If fund managers never meet analysts in person or see their pictures, beauty should have little effect on voting outcomes. To test this alternative explanation, we investigate whether the beauty premium varies with the physical distance between analysts and fund managers or with media coverage of analysts. Specifically, regarding physical distance, we assume analysts working in New York City (*NYC*), where the major stock exchanges are located, have more opportunities for face-to-face interaction with voters. As for media coverage, we calculate the number of times an analyst appears in major news and business sources (*MCover*), such as press release wires, Reuters newswires and the *Wall Street Journal*. We then add *NYC* and *MCover* and their interactions with *Beauty* to Models (4) and (6) and re-estimate the regression.[30]

The un-tabulated results show that the coefficient estimates on *NYC* and *MCover* are positive and significant, but do not change the significance of *Beauty* and the interaction term *Female×Beauty*. In addition, the interaction terms *Beauty×NYC* and *Beauty×MCover* are positive but insignificant. These results suggest that the visibility of analysts affects the likelihood of All-Star awards but has little effect on the influence of the beauty premium in All-Star analyst voting. Taken together, our evidence does not support the conjecture that the visibility of analysts explains the absence of a beauty premium in All-Star analyst voting in the U.S.

[29] Eckel and Petrie (2011) find that people are willing to pay extra to see or have face-to-face interaction with peers, suggesting that facial cues have informational value.
[30] The locations of the U.S. analysts are extracted from their LinkedIn pages.

5.7 Plastic Surgery

In Asian countries such as Korea and China, females believe that they have a better chance in recruitment and promotion when they have plastic surgery. If this is the case, some Chinese attractive analysts may have undergone plastic surgery to do better in All-Star competitions. One concern is that the likelihood of having plastic surgery may be correlated with family wealth and background. To address this concern, we asked a Korean plastic surgery doctor to look at all the pictures of Chinese female analysts and pick the ones with a high probability of having undergone plastic surgery. The doctor identified 27 such analysts (24% of the Chinese female analyst sub-sample). The average quantitative raw (mean-adjusted) beauty scores for these 27 analysts, compared with the rest of the sample, are 65.52 and 61.83 (16.60 and 11.69), respectively. This difference suggests that plastic surgery improves the beauty score. In an un-tabulated analysis, we exclude these analysts from our sample and rerun all analyses. Our results are robust.[31]

6. Conclusion

This paper investigates whether gender discrimination and beauty bias exist and interact with each other in the All-Star analyst voting process. Exploring these two types of discrimination in the U.S. and China, two countries with different cultures and legal environments, may help us better understand the root of these biases. Collectively, we find that female analysts are more (less) likely to be voted as All-Stars in the U.S. (China). Beauty, on average, does not affect the likelihood of being voted as an All-Star analyst in the U.S., but increases such likelihood in China. Attractive female Chinese analysts can overcome gender discrimination and have a similar likelihood of

[31] In another un-tabulated analysis, our results hold after excluding female analysts with *Beauty* in the top decile, suggesting that our results are not driven by a few unusual females.

being voted as an All-Star compared with male analysts. Attractive female U.S. analysts, however, suffer a beauty penalty.

Our findings, therefore, suggest that the gender and beauty effect can manifest themselves distinctively within different countries. We attribute the different forms of discrimination to the differences in culture and legal environments, acknowledging the possibility that other differences between the U.S. and China could be driving our results. Future research could further explore this issue in different settings.

References

Agthe, M., Sporrle, M., Maner, J. (2010). Don't hate me because I'm beautiful: Anti-attractiveness bias and decision making. *Journal of Experimental Social Psychology,* 46, 1151–1154.

Altonji, J., Blank, R. (1999). Race and gender in the labor market. *Handbook of Labor Economics*, 3, 3143–3259.

Andreoni, J., Petrie, R. (2008). Beauty, gender, and stereotypes: Evidence from laboratory experiments. *Journal of Economic Psychology*, 29, 73–93.

Bair, S. (2016). BankThink--The glass ceiling in finance: Barely cracked. *American Banker.* https://www.americanbanker.com/opinion/the-glass-ceiling-in-finance-barely-cracked. Accessed 6 October 2018.

Berggren, N., Jordahl, H., Poutvaara, P. (2010). The looks of a winner: Beauty and electoral success. *Journal of Public Economics*, 94, 8–15.

Biddle, J., Hamermesh, D. (1998). Beauty, productivity, and discrimination: Lawyers' looks and lucre. *Journal of Labor Economics*, 16, 172–201.

Brown, L., Call, A., Clement, M., Sharp, N. (2015). Inside the "black box" of sell-side financial analysts. *Journal of Accounting Research*, 53, 1–47.

Clement, M. (1999). Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting & Economics*, 27, 285–303.

Clement, M., Tse, S. (2003). Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? *The Accounting Review*, 78, 227–249.

Cohen, L., Frazzini, A., Malloy, C. (2010). Sell-side school ties, *Journal of Finance*, 65, 1409–1437.

Darity, W. A., Mason, P. L. (1998). Evidence on discrimination in employment: Codes of color, codes of gender. *Journal of Economic Perspectives*, 12(2), 63–90.

Deryugina, T., Shurchkov, O. (2015). Now you see it, now you don't: The vanishing beauty premium. *Journal of Economic Behavior and Organization*, 116, 331–345.

Duarte, J., Siegel, S., Young, L. (2012). Trust and credit: The role of appearance in peer-to-peer lending. *Review of Financial Studies*, 25, 2455–2483.

Eagly, A., Ashmore, R., Makhijani, M., Longo, L. (1991). What is beautiful is good, but…: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin*, 110, 109–128.

Eckel, C., Petrie, R. (2011). Face value. *American Economic Review*, 101, 1497–1513.

Egan, M. L., Matvos, G., Seru, A. (2017). When Harry fired Sally: The double standard in punishing misconduct. *Harvard Business School Working Paper,* No. 19-047.

Fang, L., Huang, S. (2017). Gender and connections among Wall Street Analysts. *Review of Financial Studies,* 30, 3305–3335.

Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin*, 111, 304–341.

Frieze, I., Olson, J., Russell, J. (1991). Attractiveness and income for men and women in management. *Journal of Applied Social Psychology*, 21(13), 1039–1057.

Gao, H., Lin, Y., Ma, Y. (2016). Sex discrimination and female top managers: Evidence from China. *Journal of Business Ethics*, 138, 683–702.

Goldin, C., Rouse, C. (2000). Orchestrating impartiality: The impact of blind auditions on female musicians. *American Economic Review*, 90(4), 715–741.

Graham, J., Harvey, C., Puri, M. (2017). A corporate beauty contest. *Management Science*, 63(9), 2773–3145.

Groysberg, B., Healy, P., Maber, D. (2011). What drives sell-side analyst compensation at high-status investment banks? *Journal of Accounting Research*, 49, 969–1000.

Gu, Z., Li, G., Li, Z., Yang, Y. (2019). Friends in need are friends indeed: The effect of social ties between financial analysts and mutual fund managers. *The Accounting Review*, 94, 153–181.

Halford, J., Hsu, H. (2014). Beauty is wealth: CEO appearance and shareholder value. Available at SSRN: https://ssrn.com/abstract=2357756. Accessed 6 October 2018.

Hamermesh, D., Biddle, J. (1994). Beauty and the labor market. *American Economic Review,* 84, 1174–94.

Hamermesh, D., Meng, X., Zhang, J. (2002). Dress for success–does primping pay? *Labour Economics*, 9(3), 361–373.

Hong, H., Kubik, J. (2003). Analyzing the analysts: Career concerns and biased earnings forecasts. *Journal of Finance,* 58, 313–351.

Human Rights Watch. (2018). Only Men Need Apply – Gender Discrimination in Job Advertisements in China. https://www.hrw.org/report/2018/04/23/only-men-need-apply/gender-discrimination-job-advertisements-china. Accessed 20 J 2019.

INSEAD. (2018). Gender equality in Asian corporate leadership – distant dream or achievable ideal? https://centres.insead.edu/emerging-markets-institute/documents/GenderReport_000.pdf. Accessed 1 August 2018.

Kuhn, P., Shen, K. (2013). Gender discrimination in job ads: Evidence from China. The *Quarterly Journal of Economics*, 128, 287–336.

Kumar, A. (2010). Self-selection and the forecasting abilities of female equity analysts. *Journal of Accounting Research*, 48, 393–435.

Langlois, J., Kalakanis, L., Rubenstein, A., Larson, A., Hallam, M., Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, 126, 390–423.

Li, X., Sullivan, R., Xu, D., Gao, G. (2013). Sell-side analysts and gender: A comparison of performance, behavior, and career outcomes. *Financial Analysts Journal*, 69(2), 83–94.

Liu, X., Sierminska, E. (2015). Evaluating the effect of beauty on labor market outcomes: A review of the literature. *LISER Working Paper Series.* Available at SSRN: https://ssrn.com/abstract=2613431 or http://dx.doi.org/10.2139/ssrn.2613431

Luxen, M., Van de Vijver, F. (2006). Facial attractiveness, sexual selection, and personnel selection: When evolved preferences matter. *Journal of Organizational Behavior*, 27: 241–255.

Mobius, M., Rosenblat, T. (2006). Why beauty matters. *American Economic Review*, 96, 222–235.

Oliver Wyman. (2016). *Women in Financial Services.* https://www.oliverwyman.com/content-/dam/oliverwyman/global/en/2016/june/WiFS/WomenInFinancialServices_2016.pdf. Accessed 6 October 2018.

Perrett, D., May, K., Yoshikawa, S. (1994). Facial shape and judgements of female attractiveness. *Nature,* 368, 239–242.

Ruffle, B., Shtudiner, Z. (2015). Are good-looking people more employable? *Management Science*, 61, 1760–1776.

Todorov, A., Mandisodza, A., Goren, A., Hall, C. (2005). Inferences of competence from faces predict election outcomes. *Science*, 308(5728), 1623–1626.

Figure 1
Percentage of Female Analysts

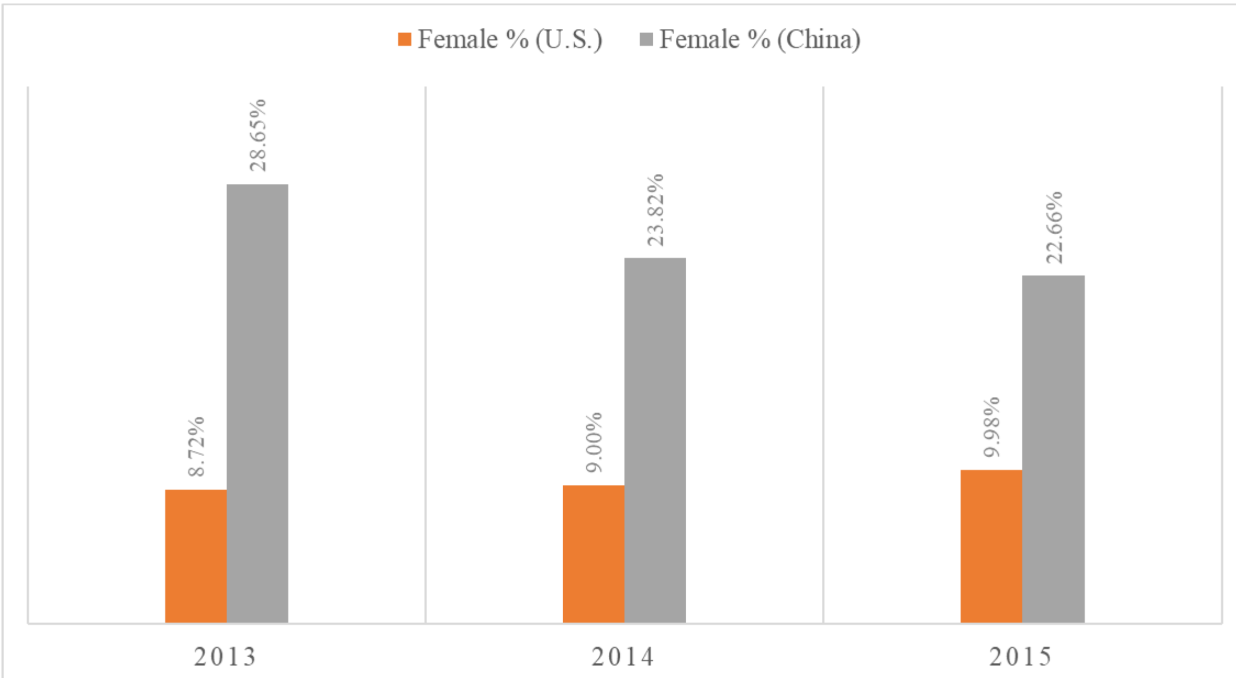

Table 1
Sample Selection

Panel A: U.S. analyst sample

Sample selection criteria	Number of Observations
Analysts with names in the I/B/E/S recommendation file, 2014	4,377
Exclude: without LinkedIn profile	(1,755)
Exclude: without quality photos for rating	(1,195)
	1,427
Analyst-years with quality photos and EPS forecasts, 2013 – 2015	3,695
Exclude: industries not classified by *Institutional Investor*	(614)
Exclude: without actual EPS to calculate EPS forecast errors	(13)
Exclude: without financial statement data to calculate control variables	(52)
Exclude: without any All-Star analyst in an industry-year	(307)
Final sample: number of analyst-years (analysts)	2,709 (1,121)

Panel B: China analyst sample

Sample selection criteria	Number of Observations
Candidates for All-Star voting, with names and photos provided by *New Fortune Magazine*	922
Exclude: sales managers	(170)
Exclude: candidates working in the Macroeconomy, Strategy, Financial Engineering, or Fixed Income sectors	(167)
	585
Analyst-years with quality photos and EPS forecasts, 2013 – 2015	1,274
Exclude: with group photos	(66)
Exclude: with actual EPS to calculate EPS forecast errors	(15)
Exclude: with financial statement data to calculate control variables	(27)
Exclude: with at least one All-Star analyst in an industry-year	(206)
Final sample: number of analyst-years (analysts)	960 (442)

This table presents the sample selection criteria. We start with a sample of 1,427 U.S. analysts (585 Chinese analysts) with quality photos for rating purposes. After merging with I/B/E/S and Compustat (CSMAR) databases, we derive the final sample consisting of 2,709 U.S. analyst-years (960 Chinese analyst-years) for the years 2013 to 2015.

Table 2
Summary Statistics for Key Variables

	U.S. Analyst Sample (N = 2,709)			China Analyst Sample (N = 960)		
Variable	Mean	Median	Stdev	Mean	Median	Stdev
Star_Award	0.09	0.00	0.28	0.37	0.00	0.48
Female	0.09	0.00	0.29	0.24	0.00	0.43
Beauty (Quantitative)	49.05	48.44	10.87	49.61	48.90	12.65
Beauty (Mean-Adj. Quantitative)	-2.35	-1.21	17.91	-2.73	-3.31	15.00
Age	40.39	40.00	7.46	31.41	30.00	4.54
CAR	-0.00	0.00	0.04	0.02	0.00	0.05
AFE	28.90	26.25	14.78	34.78	33.31	15.65
Horizon	75.22	78.34	16.40	57.40	58.46	20.83
Freq	43.58	44.50	20.32	30.94	29.08	18.70
BSize	60.92	42.50	56.07	61.27	53.67	27.28
NFirm	13.39	13.00	7.06	19.82	18.00	11.91
NInd	4.21	4.00	2.46	5.28	5.00	2.83
GExp	8.49	7.50	6.68	2.23	2.00	1.94
FExp	4.18	3.62	2.44	1.51	1.30	0.67
Size	8.91	8.97	1.44	16.49	16.43	0.66
MTB	5.76	3.91	7.17	2.09	1.86	1.06
ROA	0.00	0.03	0.12	0.06	0.06	0.03

This table presents descriptive statistics for the sample used in the empirical tests. *Star_Award* = All-Star analyst, an indicator variable set to one if the U.S. (Chinese) analyst is a star analyst in year t, and zero otherwise. *Female* = an indicator variable set to one if the analyst is female, and zero otherwise. *Beauty* = Facial attractiveness of the analyst, measured in raw quantitative and mean-adjusted quantitative terms. *Age* = Age of the analyst. *CAR* = The average three-day abnormal cumulative stock returns surrounding the analyst's stock recommendations for the firms covered in year t. *AFE* = Earnings forecast error, calculated as the mean of the analyst's relative forecast errors for the firms followed in year t, where relative forecast error is the analyst's most recent forecast error (i.e., |analyst earnings forecast − actual earnings|) relative to the most recent forecast errors of all analysts following the firm in year t and ranges from 0 to 100 (Clement and Tse 2003). *Horizon* = Earnings forecast horizon, calculated as the mean of the analyst's relative forecast horizon for the firms followed in year t, where relative forecast horizon is the analyst's forecast horizon (i.e., the number of days between the analyst's earnings forecast and the firm's actual earnings announcement) relative to the forecast horizons of all analysts following the firm in year t and ranges from 0 to 100. *Freq* = Earnings forecast frequency, calculated as the mean of the analyst's relative forecast frequency for the firms followed in year t, where relative forecast frequency is the analyst's forecast frequency relative to the forecast frequencies of all analysts following the firm in year t and ranges from 0 to 100. *BSize* = Brokerage firm size, calculated as the number of analysts employed by the sell-side firm in year t. *NFirm* = Number of firms followed by the analyst in year t. *NInd* = Number of industries followed by the analyst in year t. *GExp* = General experience, defined as the number of years between an analyst's first appearance in the I/B/E/S or CSMAR database and the end of year t. *FExp* = Firm-specific experience, defined as the average number of years that the analyst has followed the firms in his or her research portfolio in year t. *Size* = Average firm size, measured as the mean of the natural logarithm of the market value of the firms followed by the analyst in year t. *MTB* = Average market-to-book ratio of the firms followed by the analyst in year t. *ROA* = Average return on assets of the firms followed by the analyst in year t.

Table 3
Analyst Gender and Performance

	U.S. Analyst Sample	China Analyst Sample
	(1)	(2)
	AFE	*AFE*
Female	0.032	-2.156**
	(0.04)	(-2.07)
Age	-0.036	0.136
	(-0.69)	(1.53)
Horizon	0.393***	0.476***
	(17.51)	(14.29)
Freq	-0.021	0.067**
	(-0.87)	(2.49)
BSize	0.008	-0.014
	(1.11)	(-0.58)
NFirm	-0.232***	-0.116
	(-3.81)	(-1.59)
NInd	0.561***	0.525*
	(3.78)	(1.84)
GExp	-0.029	-0.275
	(-0.46)	(-0.60)
FExp	0.157	0.479
	(0.91)	(0.36)
Size	-1.826***	-2.034*
	(-3.62)	(-1.79)
MTB	-0.024	-1.285
	(-0.59)	(-0.96)
ROA	4.380	-67.710*
	(0.94)	(-1.87)
Industry-Year Fixed Effects	Included	Included
N	2,709	960
Adj. R-squared	0.397	0.432

This table presents the results from estimating the OLS regression of model (1). See Table 2 for the definition of all variables. *t*-statistics (in parenthesis) are calculated based on standard errors clustered by broker and by year. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 4
Test of H1: Does Gender Discrimination Exist in the Star Analyst Voting?

	U.S. Analyst Sample	China Analyst Sample
	(1)	(2)
	Star_Award	*Star_Award*
Female	0.301*	-0.093***
	(1.77)	(-3.42)
Age	0.025*	0.003
	(1.89)	(0.32)
CAR	2.119***	-0.136
	(5.08)	(-0.13)
AFE	-0.019***	0.007
	(-2.61)	(0.82)
Horizon	0.013**	0.002
	(2.13)	(0.18)
Freq	0.004	0.016***
	(1.15)	(4.33)
BSize	0.008***	0.011***
	(7.61)	(4.99)
NFirm	0.085***	0.047***
	(3.84)	(4.14)
NInd	-0.103**	-0.121*
	(-2.06)	(-1.86)
GExp	-0.026	-0.039
	(-1.01)	(-0.51)
FExp	0.125***	-0.089
	(4.33)	(-0.42)
Size	0.485***	0.453**
	(6.60)	(2.19)
MTB	0.002	0.068
	(0.35)	(0.49)
ROA	1.577*	-3.648
	(1.82)	(-0.43)
Industry-Year Fixed Effects	Included	Included
N	2,709	960
Pseudo R-squared	0.393	0.220

This table presents the results from estimating the probit regression of model (2). See Table 2 for variable definitions. *z*-statistics (in parenthesis) are calculated based on standard errors clustered by broker and by year. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 5
Summary Statistics for Beauty Measures and Female Analyst Coverage

Panel A reports the summary statistics of the raw quantitative beauty scores based on both MTurk and student ratings. It also reports the summary statistics by different ethnicity groups of raters. Panel B reports the summary statistics of the raw quantitative beauty measures, age, and general experience by analyst groupings. The significance of the mean difference at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively. Panel C reports the top 3 industries by the percentage of female analysts.

Panel A: beauty scores by ethnicity groups of raters

	U.S. Analyst Sample			China Analyst Sample		
	Mean	Median	Stdev	Mean	Median	Stdev
MTurk ratings						
All raters	49.56	50.00	20.98	50.17	50.00	21.27
By ethnicity:						
Non-Asian raters	48.03	50.00	20.60	50.49	50.00	21.09
Asian raters	55.80	57.00	21.37	47.82	50.00	22.38
Student ratings						
All raters	43.55	45.00	21.38	41.46	43.00	22.64
By ethnicity:						
Non-Asian raters	44.35	46.00	21.57	40.84	41.00	22.78
Asian raters	42.05	44.00	20.93	42.60	45.00	22.36

Panel B: average beauty score, age and general experience by analyst groupings

	U.S. Analyst Sample			China Analyst Sample		
	Beauty	*Age*	*GExp*	*Beauty*	*Age*	*GExp*
By gender:						
Male analysts	47.94	41	9	45.53	32	2
Female analysts	59.93	38	7	62.33	31	2
Difference: male - female	-11.99***			-16.80***		
By award status:						
Star Analysts	49.56	43	12	50.60	31	2
Non-Star Analysts	49.00	40	8	49.02	31	2
Difference: Star - non-Star	0.56			1.58*		

Panel C: top 3 industry sectors covered by female analysts

U.S. Analyst Sample		China Analyst Sample	
Industry Sector	Female%	Industry Sector	Female%
1. Food, Beverage, Household, Personal Care Products	24%	1. Textiles Garment and Apparel	80%
2. Retailing/Department Stores, Specialty Softlines	23%	2. Services (Hotel, Restaurants and Leisure)	60%
3. Retailing/Food, Drug Chains	19%	3. Papermaking and paper products	50%

Table 6
Analyst Beauty and Performance

	U.S. Analyst Sample	China Analyst Sample
	(1)	(2)
	AFE	*AFE*
Beauty	-0.022	-0.028
	(-1.31)	(-1.13)
Age	-0.050	0.125
	(-0.92)	(1.38)
Horizon	0.394***	0.479***
	(17.60)	(14.31)
Freq	-0.019	0.068**
	(-0.78)	(2.52)
BSize	0.008	-0.014
	(1.13)	(-0.55)
NFirm	-0.233***	-0.112
	(-3.82)	(-1.52)
NInd	0.557***	0.554**
	(3.74)	(2.01)
GExp	-0.028	-0.264
	(-0.45)	(-0.57)
FExp	0.157	0.247
	(0.90)	(0.19)
Size	-1.820***	-1.938*
	(-3.65)	(-1.66)
MTB	-0.024	-1.297
	(-0.59)	(-0.95)
ROA	4.423	-73.521**
	(0.94)	(-2.10)
Industry-Year Fixed Effects	Included	Included
N	2,709	960
Adj. R-squared	0.397	0.430

This table presents the results from estimating the OLS regression of model (3). See Table 2 for variable definitions. *t*-statistics (in parenthesis) are calculated based on standard errors clustered by broker and by year. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 7
Test of H2: Does Beauty Discrimination Exist in the Star Analyst Voting?

	U.S. Analyst Sample	China Analyst Sample
	(1)	(2)
	Star_Award	*Star_Award*
Beauty	-0.001	0.006***
	(-0.24)	(2.59)
Age	0.022*	0.010
	(1.70)	(0.87)
CAR	2.074***	-0.056
	(3.44)	(-0.05)
AFE	-0.019***	0.007
	(-2.63)	(0.83)
Horizon	0.013**	0.002
	(2.06)	(0.17)
Freq	0.005	0.016***
	(1.36)	(4.36)
BSize	0.008***	0.011***
	(7.84)	(4.85)
NFirm	0.084***	0.046***
	(3.78)	(4.23)
NInd	-0.105**	-0.118*
	(-2.12)	(-1.80)
GExp	-0.024	-0.034
	(-0.91)	(-0.44)
FExp	0.123***	-0.096
	(3.99)	(-0.45)
Size	0.486***	0.450**
	(6.55)	(2.10)
MTB	0.003	0.076
	(0.55)	(0.54)
ROA	1.553*	-4.413
	(1.71)	(-0.54)
Industry-Year Fixed Effects	Included	Included
N	2,709	960
Pseudo R-squared	0.391	0.222

This table presents the results from estimating the probit regression of model (4). See Table 2 for variable definitions. *z*-statistics (in parenthesis) are calculated based on standard errors clustered by broker and by year. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 8
The Interaction Effect of Analyst Gender and Beauty on Performance

	U.S. Analyst Sample	China Analyst Sample
	(1)	(2)
	AFE	*AFE*
Female	0.248	-2.469*
	(0.23)	(-1.92)
Female × Beauty	-0.017	0.023
	(-0.34)	(0.58)
Male × Beauty	-0.024	0.002
	(-1.41)	(0.05)
Age	-0.050	0.141
	(-0.91)	(1.59)
Horizon	0.394***	0.475***
	(17.63)	(13.94)
Freq	-0.019	0.067**
	(-0.79)	(2.43)
BSize	0.008	-0.015
	(1.09)	(-0.59)
NFirm	-0.232***	-0.116
	(-3.80)	(-1.58)
NInd	0.557***	0.523*
	(3.72)	(1.83)
GExp	-0.028	-0.275
	(-0.45)	(-0.60)
FExp	0.156	0.538
	(0.89)	(0.41)
Size	-1.817***	-2.044*
	(-3.62)	(-1.82)
MTB	-0.024	-1.271
	(-0.60)	(-0.94)
ROA	4.384	-67.570*
	(0.93)	(-1.86)
Industry-Year Fixed Effects	Included	Included
N	2,709	960
Adj. R-squared	0.397	0.431

This table presents the results from estimating the OLS regression of model (5). See Table 2 for variable definitions. *t*-statistics (in parenthesis) are calculated based on standard errors clustered by broker and by year. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.

Table 9
Test of H3: Does Beauty Affect Gender Discrimination in the Star Analyst Voting?

	U.S. Analyst Sample	China Analyst Sample
	(1)	(2)
	Star_Award	*Star_Award*
Female	0.418**	-0.384***
	(2.37)	(-3.87)
Female × Beauty	-0.014***	0.017***
	(-2.62)	(3.10)
Male × Beauty	0.000	0.009
	(0.07)	(1.63)
Age	0.024*	0.012
	(1.93)	(1.04)
CAR	2.202***	-0.151
	(5.89)	(-0.15)
AFE	-0.018***	0.007
	(-2.62)	(0.73)
Horizon	0.013**	0.002
	(2.15)	(0.18)
Freq	0.005	0.016***
	(1.35)	(4.08)
BSize	0.008***	0.011***
	(7.98)	(4.67)
NFirm	0.086***	0.046***
	(3.84)	(4.04)
NInd	-0.102**	-0.123*
	(-2.06)	(-1.90)
GExp	-0.028	-0.036
	(-1.11)	(-0.47)
FExp	0.127***	-0.064
	(4.53)	(-0.30)
Size	0.495***	0.433**
	(6.57)	(2.06)
MTB	0.003	0.084
	(0.61)	(0.63)
ROA	1.597*	-3.802
	(1.81)	(-0.45)
Industry-Year Fixed Effects	Included	Included
N	2,709	960
Pseudo R-squared	0.395	0.226

This table presents the results from estimating the probit regression of model (6). See Table 2 for variable definitions. *z*-statistics (in parenthesis) are calculated based on standard errors clustered by broker and by year. Significance at the 10%, 5%, and 1% levels is indicated by *, **, and ***, respectively.