UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays in Applied Economics

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Economics

by

Runjing Lu

Committee in charge:

Professor Gordon B. Dahl, Co-Chair
Professor Joseph Engelberg, Co-Chair
Professor Prashant Bharadwaj
Professor Julie Berry Cullen
Professor Ruixue Jia
Professor William Mullins

2020

The dissertation of Runjing Lu is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Co-Chair

University of California San Diego

2020

DEDICATION

To my dear family.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisors Gordon Dahl and Joey Engelberg and committee members William Mullins, Julie Cullen, Prashant Bharadwaj, and Ruixue Jia. In each of their own ways, they have taught me how to think deeply about research questions, examine the evidence critically, and keep moving forward despite challenges. Their unwavering support and guidance have been indispensable to me in times of difficulties and will always inspire me to aim higher. To my coauthors and long-time friends Siyu Chen and Teng Li, I would like to thank them for their hard work, encouragement, and friendship. I am also grateful to my comrades in the Applied Picnic, Rebecca Fraenkel, Samuel Krumholz, Remy Levin, Kye Lippold, Jiajun Lu, Wayne Sandholtz, Chelsea Swete, and Wang Yang. I am fortunate to have such a smart and supportive group to share early-stage ideas and projects and to seek help and comfort during stressful times.

On a more personal level, I would like to thank many friends here in San Diego including but not limited to Sher Ali Butt, Qi Cheng, Yongjae Choi, Viraj Deshpande, Ying Feng, Peicong Hu, Yuesong Shi, Yuehui Wang, Yidan Ying, Min Zhang, and Weiqi Zhao. I am fortunate to have their company in a very meaningful stage of my life in one of the most beautiful places in the world. Most importantly, I am grateful to my parents and my sister for always encouraging me to pursue what I want and being there when I need them. Wherever you are is my home.

Chapter 1 is coauthored with Siyu Chen and is being prepared for publication. The dissertation author was the principal author on this chapter.

Chapter 2 is coauthored with Teng Li and has been submitted for publication. The dissertation author was the principal author on this chapter.

Chapter 3 is being prepared for publication. The dissertation author was the sole author on this chapter.

VITA

2014	B. A. in Economics and Mathematics (joint major) *summa cum laude*, Emory University, Atlanta
2015-2020	Teaching Assistant, University of California San Diego
2020	Ph. D. in Economics, University of California San Diego

FIELDS OF STUDY

Major Field: Financial Economics, Labor Economics

ABSTRACT OF THE DISSERTATION

Essays in Applied Economics

by

Runjing Lu

Doctor of Philosophy in Economics

University of California San Diego, 2020

Professor Gordon B. Dahl, Co-Chair
Professor Joseph Engelberg, Co-Chair

This dissertation contains three essays exploring the factors influencing investor behavior, worker productivity, and unintended consequences of government policies.

Chapter 1 examines the role of media in equity markets. We exploit exogenous attention shocks generated by the announcements of a financial analyst award — award winners are featured on the front page of a high-profile financial magazine while analysts just missing the award are not. We find that the award announcement immediately causes higher market reaction to *pre-existing stale* recommendations from analysts barely winning the award than those by analysts barely missing it. However, the reaction *fully* reverses in six weeks. Evidence supports the notion that the

overreaction is mainly driven by attention trading induced by media exposure of award winners rather than by ability signaling from winning the award. In terms of the longer-term consequences of the award, we find that brokerages assign more resources to awardees; the awardees issue more accurate and less biased earnings forecasts, but only for stocks unaffiliated with the brokerages.

Chapter 2 studies the effects of a non-pecuniary symbolic award on winners, losers, and their peers, using a regression discontinuity design. We identify newly recruited insurance salespeople who barely won a quarterly “Best Rookie” award and those who barely missed it in a large insurance company. Our main finding is that barely winners earn less life insurance commission than barely losers in the quarter following the award designation. Interestingly, the performance difference is mainly driven by winners earning less rather than losers’ earning more. Several mechanisms, such as signaling, effort reallocation, licensing, and mean reversion are tested and ruled out. One mechanism, which we have empirical support for, is peer sabotage of winners triggered by the award designation. Finally, we examine spillover effects of the award and find no evidence that coworkers of winners and losers perform differently in any measurable aspects after the award.

Chapter 3 examines the effect of Recreational Marijuana Legalization (RML) in Colorado on the illegal marijuana possessions in its neighboring states. I use a difference-in-differences design with distance to Colorado border as treatment intensity. I find that RML in Colorado increases marijuana possession offenses among adult males in counties closer to Colorado border relative to those farther away. These findings add to the heated policy debate by pointing out externalities of RML to other states.

Chapter 1

Stale Information in the Spotlight: The Effects of Attention Shocks on Equity Markets

Abstract

Media exposure of *new* information has been shown to facilitate information incorporation into asset prices. But the causal evidence on how asset prices are affected when the media draws investor attention to *stale* information is still scarce. We exploit exogenous attention shocks generated by the announcements of a financial analyst award — award winners are featured on the front page of a high-profile financial magazine while analysts just missing the award are not. We find that the award announcement immediately causes higher market reaction to *pre-existing stale* recommendations from analysts barely winning the award than those by analysts barely missing it. However, the reaction *fully* reverses in six weeks. Evidence supports the notion that the overreaction is mainly driven by attention trading induced by public exposure of award winners rather than by ability signaling from winning the award. Suggestive evidence further shows that speculative trading based on leaked award information exacerbates the price fluctuation. To

understand the longer-term consequences of the announcements, we explore how brokerages and analysts respond in the year after the award. We find that brokerages assign more resources to awardees; the awardees issue more accurate and less biased earnings forecasts, but only for stocks unaffiliated with the brokerages. Our results highlight the temporary price-destabilizing effects of media when it draws investor attention to stale information and the long-lasting effects of public recognition on sell-side research.

1.1 Introduction

Information is fundamental to asset prices and efficient markets. The media plays a central role in its dissemination. Several papers provide causal evidence on how the media facilitates information incorporation into asset prices when it draws investor attention to *new* information that is newly released to the public.[1] However, not all media articles contain new information. If investors fail to realize the extent to which others have already traded on the publicly available stale information, investors may overreact when the media draws their attention to stale information. Yet, it has been difficult to causally identify how asset prices are affected when the media draws investor attention to *stale* information. One challenge is to find exogenous media exposure.[2]

We overcome this challenge by exploiting a series of exogenous attention shocks generated by the announcements of a financial analyst award. Analysts who barely win the award receive front-page coverage on a high-profile financial magazine while those barely missing it do not. We find that investors overreact to award winners' *pre-existing stale* stock recommendations right after the announcement, which is driven by attention trading rather than by ability signaling.

[1]Several recent papers examine the causal impact of media exposure of new information on retail trading (Engelberg and Parsons, 2011; Peress, 2014) and on stock prices (Fedyk, 2018; Lawrence, Ryans, Sun, and Laptev, 2018). They find little evidence of subsequent reversal in trading or in prices, suggesting that media exposure of new information enhances informational efficiency.

[2]See Lawrence et al. (2018) for a discussion on the identification issues.

Evidence further suggests that speculative trading based on leaked award information exacerbates the price fluctuation. To understand the longer-term consequences of the announcements, we also examine brokerages' resource allocation and analysts' performance after the award.

The award we study is *New Fortune* (NF) magazine's *Best Financial Analyst* award. NF magazine is a widely subscribed financial magazine known as the Chinese analogue to *Institutional Investor* magazine and is highly influential among financial market participants in China. NF's Best Financial Analyst award is the best-known award for sell-side financial analysts in China. Each year, based on votes from directors of research and fund managers at major investment institutions, the five analysts who obtain the highest votes in each of approximately 30 industries are elected as "star winners".[3] They are awarded trophies in a widely publicized award ceremony and receive front-page coverage of their names and votes on NF magazine in print and online. Analysts who receive the sixth and seventh highest votes in each industry are known as the finalists of the award (hereafter, star finalists). Although their names and votes can be found on the NF website, they are in a section given much lower prominence. Information on analysts who do not make it into the finalist groups goes unpublished.

Our empirical strategy makes use of a regression discontinuity (RD) design. We use NF magazine's proprietary data to identify the analysts with a number of votes just above and just below the award cutoffs. We then compare the return of stocks with *pre-existing stale* recommendations from these two groups of analysts *after* the award announcement. The key identifying assumption is that the award designations and the front-page coverage are assigned as good as random around the award cutoffs. In other words, close to the cutoffs, which analysts win the award and receive investor attention is not perfectly predictable.[4]

Our approach has two advantages. First, while most existing studies rely on investor

[3]NF magazine categorizes industries with fewer than 20 competing analysts as "small industries" and only elects the three analysts with the highest votes as "star winners" in these industries. In this paper, we focus on large industries because they are more important and receive greater investor attention.

[4]Supporting the identifying assumption, we find no difference between analysts above and below the award cutoffs along almost all observable dimensions, such as demographics, characteristics of brokerages, characteristics of recommended stocks, and market reactions to stock recommendations before the award announcement.

attention induced by endogenous media coverage, our attention shocks are based on the exogenous variation in the front-page coverage of analysts around the award cutoffs. Second, while past papers rely on reprints or recombination of stale news events whose resulting narratives or perspectives may introduce new information, our paper focuses on stale recommendations that are *exactly the same* as when they were first released.

Our first main finding is a clear discontinuity in the post-award return of stocks with pre-existing stale recommendations from analysts barely winning and barely missing the award. In the first two trading days after the award announcement, stocks with pre-existing recommendations from analysts above the winner cutoff see 66 basis points higher adjusted buy-and-hold cumulative abnormal return (BHAR) than those with recommendations from analysts below the cutoff. The difference is 1.3 times the average two-day BHAR of non-stale stock recommendations issued in the year before the award announcement. In addition, the former stocks also experience 30 percent higher cumulative abnormal turnover (CAT) than the latter. These findings are inconsistent with the semi-strong form efficient market hypothesis, under which information in stale recommendations should have been reflected in prices and the will not react to stale recommendations.

Next, we explore the mechanisms underlying the market reactions to stale recommendations. Our evidence is consistent with the notion that media exposure of star winners draws investor attention to winners' stale recommendations and changes investors' trading decisions (*attention trading*). First, the media mentions of and the search volume for the award both surge right after the award announcement. Second, the difference in the return of star winners' and non-winners' stocks is more pronounced among less-known analysts, less-known stocks, and stocks held by more retail investors. In contrast, the signal of analyst ability from winning the award (*ability signaling*) does not appear to drive our findings. First, investors do not react more to star winners' new stock recommendations issued shortly after the award announcement. Second, the difference in the return of star winners' and non-winners' stocks *fully* reverses within

20 trading days after the award.

Interestingly, investors also react differently to stocks with pre-existing stale recommendations from analysts above and below the vote cutoff for finalists. Compared to stocks previously recommended by non-finalists, the stocks by analysts who barely make it into the finalist groups experience *higher* average two-day BHAR in the week *before* the award announcement but *lower* two-day BHAR right *after* the announcement.[5] Evidence suggests that the list of finalists may be leaked and traded on by institutional investors before the award announcement. A portfolio buying stocks recommended by all finalists before the award announcement while selling stocks by *failed* finalist after the announcement earns a risk-adjusted daily return of 18 basis points over the 10 days around the announcement. The findings highlight that insider trading can exacerbate price fluctuation and market inefficiency.

Our second main finding is that brokerages and analysts change their behavior in the year following the award. We find that analysts above the finalist cutoffs are assigned to larger teams and to teammates with better forecast accuracy than those below, consistent with the notion that brokerages allocate additional resources to them. These analysts also issue more profitable stock recommendations and more accurate and less biased earnings forecasts than those below. Interestingly, the improved performance is only among stocks which are *not* underwritten by the brokerages. The findings imply that public recognition can discipline sell-side research but conflicts of interest attenuate the disciplinary effect.

This paper makes several contributions. Firstly, we contribute to the burgeoning literature evaluating the impact of media and investor attention in financial markets. Several recent papers provide causal evidence on how the media affects retail trading (Engelberg and Parsons, 2011; Peress, 2014) and stock prices (Fedyk, 2018; Lawrence, Ryans, Sun, and Laptev, 2018). Most of them focus on media coverage of *new* information and find little evidence of subsequent

[5]Star finalists are privately notified of their finalist status in the week before the award announcement (notification week), although they do not know their exact ranking or vote. In the analysis at the finalist cutoff, we only look at stock recommendations issued before the notification week.

reversal in trading or prices, suggesting that the media can enhance informational efficiency by drawing investor attention to new information. Meanwhile, papers such as Huberman and Regev (2001), Tetlock (2011), Gilbert, Kogan, and Lochstoer (2012) and Fedyk and Hodson (2019) document a linkage between media coverage of *stale* information and market *over*-reaction.[6] A causal interpretation of these findings is hindered by the endogeneity in the media coverage.[7] Moreover, as operationalized in these studies, the "stale" information is not entirely stale, because it tends to be a recombination of stale pieces whose resulting narratives or perspectives may contain new information. In contrast, we exploit exogenous media exposure of the authors of stale recommendations. We are thus able to provide causal evidence that the media can make prices less efficient in the short run when it draws investor attention to stale information.[8] Our paper helps to paint a more comprehensive and nuanced picture of the role of media in the market. In addition, our findings provide empirical evidence for models featuring investors who do not distinguish stale from new information (e.g., Tetlock, 2011). Investors may fail to realize the staleness of pre-existing recommendations even when the recommendations are unambiguously

[6]Huberman and Regev (2001) show that a *New York Times* article, which repeats previously published information, generates a price increase and subsequent partial reversal for the covered firm. Tetlock (2011) documents an overreaction to coverage of stale news events and a negative relationship between absolute abnormal returns and the staleness of news. Gilbert et al. (2012) find that market reactions following the publication of a macroeconomic index based on already public data reverse within one day. Fedyk and Hodson (2019) show that the relationship between absolute abnormal returns and coverage of stale news events is mainly driven by recombination of several stale news pieces.

[7]There could be unobservables simultaneously driving media coverage of stale news events and market reaction. For example, when Tesla's car batteries catch fire, media outlets may re-report Tesla's disappointing earnings announcement last quarter. Even if there is a market reaction after the coverage of the stale earnings announcement, it is unclear whether the reaction is to the battery fire or to the coverage of the stale information. It is worth pointing out that Gilbert et al. (2012) solve the endogeneity issue by focusing on pre-scheduled release of a macroeconomic index, but their time series analysis cannot control for contemporaneous common shocks and is hindered by a small sample size, as pointed out by Tetlock (2011).

[8]For outcomes other than stock prices, Kaniel and Parham (2017) find that fund flows respond to mutual funds' appearance on the WSJ Category King list which is based on past fund performance; Phillips, Pukthuanthong and Rau (2014) show a linkage between fund flows and stale fund returns arising from horizon effects in holding period returns. Our paper differs from theirs in the timing and the main mechanisms of the market reaction. First, Kaniel and Parham (2017) document a gradual increase in fund flows starting around two weeks after the ranking announcement, while we find an immediate market reaction in just days after the award announcement. Second, fund complexes' advertisement and investors' updated beliefs regarding fund performance are the main mechanisms for the above two papers, while attention trading due to heightened media exposure of the analysts is a more plausible channel in our setting.

old.

Secondly, our paper extends the literature on the impact of awards on stock prices. Prior papers focus on product quality awards and CEO awards and explain market reactions following the award announcement by citing the role of awards in signaling quality (Hendricks and Singhal, 1996) and incentivizing performance (Ammann, Horsch, Oesch, 2016; Malmendier and Tate, 2009). By contrast, we highlight the under-studied role of awards in increasing visibility of awardees, which can attract investor attention and affect asset prices.

Thirdly, we contribute to the literature on the post-award performance of star analysts.[9] Using a RD design, we confirm past findings that star analysts perform better in earnings forecasts and are more likely to be promoted than their unranked peers (Stickel, 1992; Leone and Wu, 2007; Fang and Yasuda, 2009; Wu and Zang, 2009; Xu, Chan, Jiang, and Yi, 2013). Moreover, we provide the first piece of causal evidence that brokerages reallocate resources in response to award designations. This resource reallocation can help to explain the persistent difference in the post-award performance between star and non-star analysts documented in the literature. Our findings also highlight the impact of teammates on worker productivity and broadly relate to the literature on peer effects in the workplace (e.g., Mas and Moretti, 2009). Finally, we find that star analysts' performance improvement differs between affiliated and non-affiliated stocks. This adds to the literature on how conflicts of interests affect financial analysts' research (Ljungqvist et al., 2007; Corwin, Larocque and Stegemoller, 2017).

Taken together, this paper highlights the temporary price-destabilizing effects of media when it draws investor attention to stale information and the long-lasting effects of public recognition on sell-side research. The remainder of the paper is as follows. Section 2 describes the institutional background and the data. Section 3 outlines the empirical strategy, presents market reactions after the award, and explores the underlying mechanisms. Section 4 examines the changes among brokerages and analysts after the award. Section 5 concludes.

[9]See Bradshaw (2011) for a review on studies about star financial analysts.

1.2 Background and Data

1.2.1 Background

China's stock market. Our testing ground is China's stock market. China opened its Shanghai and Shenzhen stock exchanges in 1990. Following more than two decades of rapid growth, the country's stock market reached a total equity value of six trillion U.S. dollars and became the world's second largest market in 2014. Given its large size, active trading, and market features representative of many emerging markets, China's stock market has become an important subject for mainstream research in finance (Carpenter and Whitelaw, 2017). Several features of China's stock market also make it an interesting setting for our study. First, the market is dominated by retail investors who are less informed and more prone to behavioral biases. Retail investors accounted for over 98.2 percent of total investor accounts and over 42 percent of all stock holding value in 2014 (China Securities Depository and Clearing, 2014; WIND, 2014). This percentage is much larger than that in many developed countries like the U.S. (Andrade, Bian, and Burch, 2013). Second, short sale was restricted, and future and option contracts for individual stocks were not allowed during our sample period. Such restrictions hinder the ability of sophisticated investors to arbitrage away mispricing in stocks. Therefore, China's stock market is farther away from an efficient market than its U.S. counterpart, and investor behavior should play a more important role in asset pricing in such a setting.

The "Best Financial Analyst" award. Financial analysts are important information agents in financial markets (Cohen, Frazzini, and Malloy, 2010). Their opinions are among the most widely solicited, anticipated, and dissected news items in financial markets. In addition, their data are available in large quantity and relatively standardized formats. Therefore, the industry of financial analysts is a useful setting to address our research question.

The award we exploit is *New Fortune* (NF) magazine's "Best Financial Analyst" award. NF magazine is a widely subscribed monthly financial magazine known as the Chinese version of *Institutional Investor* magazine and is highly influential among financial market participants in China. The goal of the award is to identify the best sell-side financial analysts who provide original and insightful opinions about China's stock market. Since its beginning in 2003, the award has been the largest and the most important award among sell-side financial analysts in China. In 2014 alone, about 2,700 investment professionals across 870 institutions managing fund sizes of around 1.3 trillion U.S. dollars participated as voters for the award. Over 1,500 analysts at 47 brokerage houses participated as award candidates. This accounts for 50 percent of registered financial analysts and 40 percent of brokerage houses in China. The award is also crucial for financial analysts' promotion and income. Anecdotally, star winners can get an annual income boost for over one million U.S. dollars.

The timeline of the award is summarized in Figure 1.1. Each year in August, investment professionals register as voters for the award. In mid- to late-October, eligible voters receive ballots from NF, which includes a list of brokerage-endorsed analysts and their bios in each of approximately 30 NF industries.[10] The voters rank the top five analysts in each NF industry according to the overall quantity and quality of the analysts' reports in the past year.[11] NF then weighs the vote to each analyst based on the analyst's ranking in the vote and the fund size managed by the voter.[12] The resulting weighted sum of votes is known as *scores*. Within each NF industry, the analysts with top five scores are elected as the Best Financial Analysts (*star winners*), and the analysts with the sixth and seventh scores are known as the finalists of the award (*star finalists*). On a Friday or Saturday afternoon at the end of each year, NF will hold a widely-publicized award ceremony to present trophies to star winners in front of all

[10]NF magazine categorizes industries with fewer than 20 competing analysts as "small industries"and elects fewer "star winners" in these industries. In this paper, we focus on large industries because they are more important and receive greater investor attention. In addition, we also exclude Economics & Strategy category industry, because analysts in this industry usually do not recommend individual stocks.

[11]There are no explicit voting guidelines or rubrics. According to surveys conducted by NF, institutional investors tend to evaluate an analyst by whether or not the analyst has a solid understanding of the market, provides original and insightful opinions based on quantitative analysis, and maintains timely communication with the institutional investors.

[12]The voting scheme changes slightly from year to year, and we use year fixed effects to control for these.

major institutional investors. At the same time, the names and votes of the star winners are also published on the front page of NF website and the magazine. Although the names and votes of star finalists are also publicly available, they appear in a much less prominent section on the NF website. The votes of the analysts who do not make it to the finalist groups goes unpublished.[13]

It is important to note that in the week before the award announcement, namely the notification week, NF sends finalists invites to the award announcement. Notified finalists do not know their votes or rankings and are required to keep the information private until the public award announcement. An extensive internet search performed by the authors confirm that no lists of NF finalists are posted on line before the date of the award announcement between 2005 and 2014. Therefore, the general public do not know who the finalists are before the award announcement takes place.

1.2.2 Data Sources

We obtain our data from three sources. The first is the proprietary vote and ranking data from NF. The data comprises a list of top 15 analysts (or teams) in each NF industry each year between 2005 and 2014 and includes analysts' names, gender, highest academic degree, and work history up till the point they last participated in the ranking.[14] We supplement this dataset with information in analysts' resumes posted on platforms such as the Security Association of China (http://www.sac.net.cn/xxgs/cyryxxgs), homepages of brokerage houses, personal websites of analysts, and financial industry job sites (*Golden Compass* http://stock.sohu.com/s2011/jlp and *Ifeng Finance* http://star.finance.ifeng.com). The data from NF is crucial, because we cannot identify the analysts who fail to make it to the finalist groups without the data, which is necessary

[13] Although investors can identify these analysts' names from a publicly available list of all award candidates, they do not know the ranking or votes of these analysts.

[14] If a team rather than an individual appears on the NF ranking list, we assign the same ranking and score to all analysts in the team. If an analyst participates in more than one NF industry in a year, we keep the analyst's most-covered NF industry, i.e., the industry where she covers the largest number of companies, following Boni and Womack (2006) and Emery and Li (2009).

to implement the RD design.

The second dataset is the China Stock Market Accounting Research (CSMAR) database. CSMAR is a comprehensive database on China's stock market, containing important financial information and publicly available analyst reports on all companies listed on the Shanghai and the Shenzhen Stock Exchanges. The dataset is included in Wharton Research Data Service and widely used in research on China's stock market. We obtain daily and monthly stock returns with reinvestment of cash dividends, quarterly market capitalization, quarterly institutional holdings, and annual book values for all A share companies listed in China between 2005 and 2014. We also get all stock recommendations and earnings forecasts on these companies issued between 2005 and 2014. Each stock recommendation has a report ID, stock ID, publication date, rating, rating expiration date, and the information on the authors and their brokerage houses. The rating is standardized to a five-point scale: strong sell=1, sell=2, neutral=3, buy=4, and strong buy=5. Each earnings forecast has a report ID, stock ID, publication date, earnings forecast, forecast end date, and the information on the authors and their brokerage houses. Lastly, we obtain the underwriters for initial public offering, right issue, and rationed shares occurring between 1989 and 2015.

The third dataset is the Choice Financial Terminal (Choice). Choice compiles information from mandatory filings of financial institutions in China. We obtain the annual revenue and trading commissions of brokerage houses and the annual stock holdings of mutual funds between 2005 and 2014.

1.2.3 Variable and Sample Construction

To answer our main research question, we need empirical proxies for market reactions. Following Loh and Stulz (2011), we measure market reactions to a stock recommendation with the adjusted two-day buy-and-hold abnormal return (BHAR):

$$BHAR_{st} = \prod_{\tau=t}^{t+1}(1+R_{s\tau}) - \prod_{\tau=t}^{t+1}(1+R_{s\tau}^{DGTW})$$

Day t is the day of the measurement. $R_{s\tau}$ is the return of stock s on day τ. $R_{s\tau}^{DGTW}$ is the return on a benchmark portfolio with similar size, book-to-market, and momentum characteristics as stock s on day τ (Daniel, Grinblatt, Titman, and Wermers, 1997). We assign a minus sign to BHAR if the rating in the stock recommendation is neutral, sell, or strong sell. As robustness check, we use the expected return estimated from a Fama-French five-factor model as the benchmark return for a stock (Fama and French, 2015). The detailed construction is in Appendix A.1.

For the regressions sample, we will include both recommendation revisions and reiterations. The reasons are as follows. First, reiterations may include new information despite having the same rating and may be viewed as a confirmation of past opinions (Dontoh, Ronen, and Sarath, 2003). Indeed, Jegadeesh and Kim (2006) find that investors still react to reiterations, although to a lesser extent than revisions. Second, the sample size for recommendations is too small if we exclude reiterations which account for 94 percent of recommendations during the sample period. Therefore, we err on the conservative side and include both reiterations and revisions in our sample.

The first step of sample construction is to link analysts across datasets and time. Since neither NF nor CSMAR assign unique ID to analysts, we follow Cohen et al. (2010) and use name and work history to identify analysts. If two analysts share the same name and work in the same brokerage house at the same time, we assume them to be the same person. Out of the 1,633 analysts in the NF ranking, we are able to identify 1,600 analysts and locate 1,588 analysts' stock recommendations or earnings forecasts in CSMAR.

We assemble two samples to analyze the immediate market reactions to stale recommendations after the award announcement. The first sample is for the award designations of star winners. We identify the last stock recommended by each NF analyst within 1-30 days before the award announcement. We focus on stocks recommended *before* the award announcement to alleviate the concern that analysts may change stock recommendations after they know the award results. We focus on the *last* stock recommended by each analyst, because attention-limited investors are likely to pay more attention to recent recommendations.[15] If an analyst covers more than one stock on their last day of recommendation before the award announcement, we keep all stocks covered on that day. If one stock is categorized as the last stock for more than one analyst, we keep all the observations unless the analysts are from the same NF team.[16] 2-day BHAR is measured on the first trading day after the award announcement. The base sample includes 1,157 analysts issuing 1,927 stock recommendations covering 717 stocks. We then calculate the IK bandwidth using the 2-day BHAR as the outcome and the centered vote share from the winner cutoff as the running variable (Imbens and Kalyanaraman, 2012).[17] In the regression, we only include stocks by analysts within the IK bandwidth to ensure comparability between analysts above and below the winner cutoff. The main RD sample at the winner cutoff consists of 1,003 analysts issuing 1,535 recommendations covering 644 stocks. The summary statistics are reported in Table 1.1 column (1).

The second sample is to analyze the market reactions to stale recommendations from analysts who make it into the finalist groups. Since analysts above the finalist cutoff are notified before the award announcement, analysts above and below the cutoff may change recommenda-

[15] As robustness check, we study the last three, five, or all stocks recommended by each analyst within 1-30 days before the award announcement. We also look at the last stock recommended by each analyst within 31-60 days, ..., 121-150 days before the award announcement. The estimates in these samples are generally smaller than those in the main sample, implying that investors indeed pay more attention to more recent recommendations. The results can be found in Section 1.3.5.

[16] 20 percent of stocks in the sample are categorized as the last stock for more than one analyst. We err on the conservative side and include all observations. As long as the stock bundles of analysts above and below the award cutoffs do not perfectly overlap, our estimates will be a lower bound of the local average treatment effect. As robustness check, we drop stocks recommended by more than one analyst or keep the observation with the best NF ranking. The estimates in these samples are slightly larger than those in the main sample. The results can be found in Section 1.3.5.

[17] The centered vote share from the winner cutoff is the distance between an analyst' score and the score at the winner cutoff for a NF industry in a year, normalized by the total scores of top 15 analysts for that NF industry and that year.

tion strategies discontinuously ahead of time. We thus identify the last stock recommended by each analyst within 1-30 days before the notification week rather than the award announcement. Not knowing the exact date of private notification, we use the average of 2-day BHAR on the Monday through Thursday in the notification week to proxy market reactions to stale recommendations following private notification. The base sample includes 1,249 analysts issuing 2,322 recommendations covering 767 stocks. In the regression, we only include the stocks by analysts within the IK bandwidth which is calculated using the BHAR on the first two trading day after the award announcement as the outcome and the centered vote share from the finalist cutoff as the running variable. The main RD sample at the finalist cutoff consists of 714 analysts issuing 1,088 recommendations covering 518 stocks. The summary statistics are reported in Table 1.1 column (4).

To study how brokerages and analysts change behavior in the year after the award (*post-award period*), we construct two more samples. The first sample consists of stock recommendations issued by NF analysts in the post-award period. We drop recommendations if the authors have different NF rankings or are in different NF industries, which account for 9.31% of the full sample. We then remove recommendations issued during one day before and one day after the suspension period, which accounts for 2.64% of the remaining sample, to prevent our results from being driven by abnormal reaction to suspension. 2-day BHAR is measured either on the recommendation publication date or the closest subsequent trading day. The base sample consists of 1,412 analysts issuing 71,520 stock recommendations covering 1,696 stocks. In the regression, we only include recommendations issued by NF analysts within the IK bandwidth, calculated using the 2-day BHAR as the outcome and the centered vote share as the running variable. The post-award recommendation sample at the winner cutoff consists of 804 analysts issuing 50,493 stock recommendations covering 1,596 stocks, and the sample at the finalist cutoff consists of 1,025 analysts issuing 42,805 stock recommendations covering 1,585 stocks. The summary statistics are reported in Table 1.1 columns (2) and (5).

The second sample consists of the last earnings forecast issued by each NF analyst for each stock in the post-award period. We drop earnings forecasts if the authors have different NF rankings or are in different NF industries, which account for 7.48% of the full sample. The remaining sample includes 1,440 NF analysts issuing 46,730 earnings forecasts covering 1,772 stocks. In the regression, we only include the forecasts by NF analysts within the IK bandwidths calculated using the forecast error as the outcome and the centered vote share as the running variable. The post-award forecast sample at the winner cutoff consists of 561 analysts issuing 11,860 earnings forecasts covering 1,319 stocks, and the sample at the finalist cutoff consists of 594 analysts issuing 12,051 earnings forecasts covering 1,352 stocks. The summary statistics are reported in Table 1.1 columns (3) and (6).

There are several noteworthy points. Firstly, across all regression samples, over 70% of financial analysts are male, around 90% of them have master or above degree, and their average work experience is about 3 years (Table 1.1). These statistics are in line with earlier studies on China's stock market (Hu, Lin, and Li, 2008). Secondly, the average stock ratings are around 4.34, and over 93% of the stock ratings are buy or strong buy. Compared to the stock ratings in the U.S. documented in Barber, Lehavy, McNichols and Trueman (2006), the stock rating in China is on average higher, suggesting that financial analysts are more positively biased in China. Finally, not all NF analysts in the raw sample appear in the base samples. There are three reasons: we do not have enough information to identify them in CSMAR, their reports are not recorded in CSMAR, or the stocks they cover do not have valid outcomes. This sample selection should not invalidate the RD design as long as the probability of analysts entering the base sample is uncorrelated with their award status. Therefore, we examine the probability of a NF analyst in the raw sample being included in the base sample in Table A.1. The probability changes smoothly across the award cutoffs for all base samples. Moreover, analysts in the raw sample are similar in baseline characteristics as those in the base samples, as shown in Table A.2.

1.3 Immediate Market Reaction

1.3.1 Empirical Strategies

We now explain the empirical strategies to examine the market reaction to stale recommendations after the award announcement. The main empirical strategy is an RD design.[18] We follow a standard specification for an RD design and include additional setting-specific control variables. The specification is:

$$\begin{aligned} Y_{si} = \beta_1 Win_{i(jy)} + \beta_2 f(X_{i(jy)} - C_{jy}) + \beta_3 Win_{i(jy)} f(X_{i(jy)} - C_{jy}) + \gamma_1 DayToAnnounce_{sit} \\ + \alpha_y + \alpha_b + \alpha_j + \alpha_{j_s} + \varepsilon_{si} \quad (1.1) \end{aligned}$$

where Y_{si} is the outcome of interest for stock s recommended by analyst i, e.g., market reaction on the first two days after the award announcement. Stock s is in industry j_s and receives recommendation on date t. Analyst i works in brokerage b and participates in NF industry j in year y. For the sample at the winner (finalist) cutoff, $Win_{i(jy)}$ equals one if analyst i ranks top five (seven) in NF industry j in year y and zero otherwise. $X_{i(jy)}$ is analyst i's score normalized by the sum of all analysts' scores in NF industry j in year y (*vote share*). C_{jy} is the vote share at the winner (finalist) cutoff.[19] We include an interaction term between $X_{i(jy)} - C_{jy}$ and $Win_{i(jy)}$ to allow different slopes on different sides of the award cutoff. We include the number of days from the recommendation date to the date of information event ($DayToAnnounce_{sit}$) to control for the

[18]We do not use a difference-in-differences (DID) design, because winners and average analysts are non-comparable in their baseline characteristics and performance, and the parallel pre-trend assumption for DID design does not hold in our setting. In addition, we do not use an RD-DID design, because the pre-existing difference in market reaction to the stocks recommended by analysts above and below the award cutoffs is economically and statistically small (see Table A.3 and Figure 1.2). Combining an RD design with a DID design in our setting will difference out a *zero* baseline difference and will not quantitatively change the coefficient of interest.

[19]We use vote share rather than rank or scores as the running variable for the following reasons. Firstly, vote share is continuous while rank is discrete. Secondly, vote share is better than rank and scores at accounting for the difference in the level of competition across industries. Scores do not distribute evenly across rank or across industries, so rank fifth versus sixth might be close in scores and comparable in competitive industries but far away and less comparable in non-competitive industries. In Section 1.3.5, we conduct robustness check using rank and scores as running variables. Estimates are within one standard deviation from the one estimated using vote share.

staleness of the recommendation. We also include year, brokerage, NF industry and stock industry fixed effects to control for common shocks at various levels. Standard errors are clustered by NF industry-and-year.[20] The regression is estimated using local linear regression with triangular weights and IK bandwidths (Lee and Lemieux, 2010).[21]

β_1 is the coefficient of interest, which measures the difference in the market reaction to stale recommendations issued by analysts with a number of votes barely above and below the award cutoffs right after the award announcement. There are two threats to identification: (1) perfect manipulation over one's score; (2) discontinuous change in recommendation strategy at the winner cutoff before the award announcement. In the following section, we will provide evidence suggesting that neither threat exists.

1.3.2 Validity of RD Design

In this section, we examine the two core identifying assumptions of the RD design: (1) the award designations and the front-page coverage are assigned as good as random around the award cutoffs; (2) there is no discontinuous recommendation strategy change at the winner cutoff before the award announcement.

One threat to the quasi-randomness of award designations is the perfect manipulation of analysts over their scores to be above an award cutoff. However, perfect manipulation is unlikely in this setting. On average 950 analysts and 1,300 voters participated in the NF ranking each year. The large number of participants and voters make it very difficult to manipulate scores to be exactly above an award cutoff. To formally examine this, we plot the density of the vote

[20] According to Abadie, Athey, Imbens, and Wooldridge (2017), clustering should be used at the level where the probability of treatment assignment systematically differs. In our setting, the treatment assignment occurs in each NF industry each year, and the probability of winning differs systematically depending on the number of analyst candidates in each NF industry each year. Therefore, we cluster the standard errors by year-and-NF industry. We replicate the main results under various clusters, such as brokerage house, NF industry section, stock industry sector. The statistical significance remains similar. Results are in Section 1.3.5.

[21] We replicate the main results using various bandwidths, including those selected following Calonico, Cattaneo and Titiunik (2014). The coefficients on β_1 under various bandwidths are plotted in Figure A.1. We also replicate the main results using local quadratic regression and the results refer to Section 1.3.5.

share centered at the award cutoffs in Figure A.2.[22] The density change smoothly across the corresponding cutoffs. Following McCrary (2008), we run McCrary tests on the centered vote share and cannot reject the null hypothesis that the density is continuous at the winner cutoff (*p* value=0.210) or the finalist cutoff (*p* value=0.383).

To lend further support to the quasi-randomness of award designations, we compare the baseline performance and demographics of analysts above and below the award cutoffs. Figure 1.2 plots the 2-day BHAR of a stock on the recommendation publication date on the y-axis and the issuing analyst's margin of centered vote share on the x-axis. The return changes smoothly across the award cutoffs. Regression results reported in Table 1.2 column (1) tell the same story. In other words, investors do not react differently to stock recommendations from analysts above and below the award cutoffs *before* the award announcement. Moreover, Table A.3 also shows that there is no discontinuity in the average 2-day BHAR in the notification week. Besides market reactions to recommendations, the quality of earnings forecasts and NF winning history are also important measures of analysts' performance. We thus examine analysts' forecast error and their probability of being star winners or finalists in the year before each award announcement. Table 1.2 columns (2) and (3) show that neither outcome exhibit discontinuities at the award cutoffs. We further examine analysts' characteristics, such as gender, education and work experience in Table 1.2 columns (4) through (6), and the characteristics of stocks recommended by them, such as market capitalization and momentum, in Table 1.3. All variables

[22]The displayed density excludes the observations with centered vote share equaling zero, i.e., the analysts whose vote share is at the corresponding cutoff ranks. We do so because the density of centered vote share has a *mechanical* spike at zero due to the way we center the vote share, and McCrary test has lower power under a density with mechanical spike. To illustrate the issue, we run McCrary test on a series of simulated data. For each simulation, we generate 15 scores for each industry and each year from an industry-specific uniform distribution of scores based on the minimum and maximum scores in each industry in the actual data. We then use the simulated scores to construct centered vote share and run McCrary test on this centered vote share. The simulation and McCrary test are repeated 1,000 times. The centered vote share *including* the mechanical spike at zero passes McCrary only 0.1 percent of the times, while the centered vote share *excluding* the mechanical spike passes McCrary test over 85 percent of the times. This exercise implies that the mechanical spike at zero renders McCrary test low power, and excluding the spike improves the power of the test. Therefore, we only run McCrary test on the centered vote share excluding the mechanical zero in this paper. But observations right at the award cutoffs are included in all other figures and regressions.

change smoothly across the award cutoffs. Overall, analysts are similar in baseline performance, demographics, and stock characteristics, regardless of their award status.[23]

Since brokerage houses are sometimes in a better position than individual analysts to manipulate votes, we also check the characteristics of brokerage houses which employ the analysts above and below the award cutoffs. First, if certain brokerage houses always buy votes to push their analysts above the award cutoffs, we expect a drop in the diversity of brokerage houses and a jump in the probability of winning in previous years when we move from below an award cutoff to above the cutoff. Table A.4 columns (1) and (2) show that this is not the case. Neither the number of unique brokerage houses in each NF industry across years nor the probability of having at least one analyst being star winner or finalist in the last year changes significantly across the award cutoffs. In Table A.4 columns (3) through (8), we examine other characteristics of the brokerage houses, such as total assets, net profit, number of analysts, number of stock recommendations, whether the brokerage house is publicly listed, and whether it is held by mutual fund. All variables change smoothly across the award cutoffs.

One final issue is that we include stock recommendations issued in the notification week for the main RD sample at the winner cutoff. One may worry that finalists who are notified in this week may change how they recommend stocks, which then causes the discontinuity in market reactions after the award announcement. This is not a problem at the winner cutoff. Note that finalists do not know their winner status in the notification week. Therefore, changes in recommendation strategy, if any, have to be continuous at the winner cutoff and hence cannot cause the discontinuity in market reactions. To corroborate this argument, we examine analysts' recommendation strategy before and during the notification week. Table A.5 columns (1) and (2) show that analysts above the winner cutoff are no more likely to issue stock recommendations or initiate new stock coverage in the notification week compared to those below the cutoff. Table A.5 columns (3) through (7) compare the characteristics of stocks recommended by analysts above

[23]The construction of the above variables is in Appendix A.1.

and below the winner cutoff before and during the notification week. *Post* is one if the stock is recommended in the notification week, and zero if between the notification week and the 60 days before the award announcement. The insignificant coefficients on Win and on $Win \times Post$ suggest that analysts above and below the winner cutoff recommend similar stocks before the notification week and during the notification week.[24]

1.3.3 Effect at the Winner Cutoff

In this section, we discuss how investors respond differently to *stale* stock recommendations from star winners and non-winners right after the award announcement. Figure 1.3 plots the BHAR of stocks with pre-existing recommendations from star winners and non-winners on the first two trading days after the award announcement. There is a significant jump in the return of stocks from just below to just above the winner cutoff. In other words, the market reacts more to stocks previously recommended by star winners than to those by non-winners after the award announcement. Table 1.4 presents regression formalization of the above figures. Our preferred specification in column (3) shows that the stocks with stale recommendations from analysts barely winning the award experience 66 basis points higher BHAR in the first two trading days after the award announcement than those with recommendations from analysts barely missing the award. This difference amounts to 1.3 times the average two-day BHAR of non-stale stock recommendations in the year before the award. Estimates fluctuate less than half of the standard deviation when we vary the controls from column (1) to column (3).[25]

These findings suggest at the award announcement prompts some investors to react more to stale recommendations from star winners than those from non-winners. The reaction is

[24] As robustness check, we present the main results excluding the stocks recommended in the notification week. The estimate is within half the standard deviation from the one estimated using the main RD sample. The results are reported in Section 1.3.5.

[25] The main measurement of market reaction in this paper is BHAR calculated using a benchmark return following Daniel et al. (2020). As robustness check, we also construct the benchmark return based on Fama-French five-factor model (Fama and French, 2015). This alternative BHAR gives quantitatively similar results, as shown in Table A.6.

inconsistent with the semi-strong form efficient market hypothesis, under which information in stale recommendations should have been reflected in prices and the market should not react to stale recommendations. So what could be driving the reaction? There are at least two potential explanations.

Attention trading. It is possible that the heightened media exposure of star winners attract investor attention and prompt them to search for winners' stale recommendations. To understand this channel, we first examine the existence of its necessary condition — media exposure and public attention. We collect the search volume and the daily number of news articles mentioning ”New Fortune Best analyst award” on Baidu, the largest search engine and most easily accessible information source for retail investors in China. As shown in Figure 1.4, both media mentions and search volume surge following the award announcement. The media indeed extensively covers the award, and the public actively search for it.

If attention trading is the driving force, we expect the reaction to the award announcement to be more pronounced among less known stocks and the stocks recommended by less known analysts. This is because the stocks and analysts with lower ex-ante public exposure should gain more from the ex-post increase in exposure by winning the award. We measure how well known an analyst is by the number of media articles mentioning their name in the year before the award and by their NF winning history. We then split the stocks recommended by all analysts by whether the analyst has above- or below-median media exposure (11 mentions) and split the stocks recommended by star winners by whether the winner is a first-time or a repeated winner. Table 1.5 columns (1) through (4) show that the discontinuity in market reaction is indeed driven by stocks recommended by less known analysts, i.e., analysts with less media exposure and first-time winners. In addition, we proxy how well known a stock is by the number of recommendations covering it in the 30-day period before the award. We then split the stocks by the median number of analyst coverage (six recommendations). Table 1.5 columns (7) and (8) show that the discontinuity in market reaction is again driven by stocks that are less known to begin with. Finally, retail investors are more subject to limited attention (Barber and Odean, 2008). Therefore, we expect the reaction to be stronger among stocks with less institutional investor holding. We thus split the stocks by whether the share of institutional holding in the mid-year before the award is above or below the sample median (6.19 percent). Table 1.5 columns (5) and (6) indicate that the discontinuity in market reaction is driven by stocks with below-median institutional holding. Note that the correlation between each two of the four splitting variables is between -0.04 and 0.153. Therefore, the above tests are different enough and, when taken together, provide solid evidence that attention trading is at work.

Ability signaling. It is also possible that some investors view winning the award as a signal of analyst ability. For example, investors may update beliefs about the precision of price signal in winners' stale recommendations and change their investment decisions accordingly. If ability signaling is the main mechanism, we expect investors to continue reacting more to winners' *new* stock recommendations issued shortly after the award announcement. This is because investors' beliefs about analysts' ability are unlikely to rapidly change again given that new information on analysts' ability or stock fundamentals has not yet entered the market. We thus compare the 2-day BHAR of stock recommendations newly issued by star winners and non-winners in 6-35, 36-65, 66-95 days after the award announcement.[26] Table 1.6 panel A shows that the market does *not* respond significantly more to the new stock recommendations from star winners than those from non-winners. The result remains quantitatively similar even when we restrict winners to be *repeated* winners in panel B, whose signal of ability should be stronger than first-time winners. Importantly, Table A.7 shows that there is no discontinuity in the probability of issuing new recommendations between analysts above and below the winner cutoff, so differential selection in the post-award period cannot explain the lack of difference in market

[26] We exclude recommendations issued in the first five days after the award announcement because Figure 1.4 suggests that media mentions of the award continue till around five days after the award. Also, we include additional day-of-week fixed effects for this exercise, because the market reaction can be measured on any weekday depending on the publication date of the new recommendations. These fixed effects are unnecessary in the main RD regression, because the market reaction is always measured on Monday.

reaction. These findings imply that ability signaling does not persist once attention disappears, which strongly suggests that signaling is not the main mechanism.

Overall, our empirical evidence supports that attention trading is the main mechanism and ability signaling plays a limited role. Attention trading can exist *without* ability signaling, if investors search the winners not for the analysts' ability in recommending stocks but out of their curiosity generated by the analysts' frequent media appearance. The attention paid to the analysts' stale recommendations can then be a by-product of the search for the analysts, as the search engine often returns analysts' most recent recommendations as the top results. In other words, "Best *Looking* Financial Analyst" award is expected to generate similar immediate market reaction in our context, even though the award is unrelated to analyst ability.

Importantly, attention to stale recommendations can be beneficial for informational efficiency, if the information contained in the stale recommendations has *not* been fully incorporated into prices. However, the attention could lead to overreaction followed by reversal, if the information has already been reflected in the prices but some investors fail to realize it. To distinguish the two cases, we construct a time series of BHAR of stocks in the main RD sample starting at 0 from the 10^{th} trading day before the announcement and accumulating all the way to the 30^{th} trading day after the announcement. Figure 1.5 plots the average BHAR of stocks recommended by star winners and by non-winners in the main RD sample. The BHAR of the two groups converges around 20 trading days after the award announcement. To formalize the figure, we repeat regression 1.1 using BHAR during the 5-day, 10-day, ..., and 30-day period starting from the award announcement in Table A.8. Panel A shows that the coefficients on Win dummy decrease in both magnitude and statistical significance as the duration increases. In panel B, we restrict stocks to be those *without* any earnings announcements during this period, and the reversal pattern still exists. Therefore, overreaction being corrected by new fundamental information entering the market is not a leading cause for the reversal. Overall, the findings imply that the information contained in the stale recommendations is already in the prices before the award while some investors fail to realize this and overreact.

Will investors benefit from buying stocks with stale recommendations from star winners? First, investors who buy winners' stocks but do not sell them quickly enough are likely to face a loss in the short run due to the reversal. In addition, investors will not make a profit even if they keep holding the stocks. The risk-adjusted monthly return of a portfolio based on the stocks in star winners' stale recommendations from the month of the recommendation to the following 12 months is merely -1.7 basis points (p value=0.8729).[27]

There are three final points worth mentioning. Firstly, Figure A.3 shows that the magnitude of the immediate market reaction to winners' stale recommendations decreases from the start of the sample period (2005-2008) to the end of it (2012-2014), although the significance level of the estimate is smaller at the beginning due to the smaller sample size. One interpretation of this pattern is that investors learn about the overreaction overtime and lower their reaction.

Secondly, we do not find increase in the search volume on the tickers of the stocks recommended by star winners. This finding is similar to Lawrence et al. (2017) which also finds no increase in information acquisition by the Yahoo Finance users who experience the promotion of earnings announcements for certain stocks. The lack of further information acquisition may indicate that investors make purchase decisions based on minimal additional research or conduct research on platforms unobservable to us.

Finally, Figure 1.5 shows that the return of stocks recommended by both star winners and non-winners rise in the week *before* the award announcement, but the return of stocks by non-winners decreases after the announcement. One explanation is that the list of finalists is leaked in the notification week, and informed investors trade in anticipation of the subsequent market overreaction. For instance, informed investors can buy stocks with stale recommendations

[27]The portfolio consists of the latest stocks recommended by star winners in the 30 days before the award announcement. Stocks enter the portfolio in the month of the recommendation and remain in the portfolio for the following 12 months. The return of individual stocks is aggregated to the portfolio return using equal-weighted method. The portfolio alpha is calculated using the Fama-French five-factor model (Fama and French, 2015); the factors and the risk-free rate are from CSMAR.

from all finalists in the notification week but sell stocks from *failed* finalists after the award announcement. The drop in price among stocks from failed finalists after the award may then cause a panic among uninformed investors who follow suit to sell the stocks, further driving down the price. In appendix A.2, we construct a portfolio based on this strategy which is feasible for investors who obtain the list of finalists ahead of time. The portfolio earns a risk-adjusted daily return of 18 basis point during the 10 days around the announcement (Table 1.7). To understand who leaks and who trades on the information ahead of time, we now switch to the cutoff of finalist.

1.3.4 Effect at the Finalist Cutoff

In this section, we examine the difference in the return of stocks with stale recommendations from analysts above and below the cutoff of *finalist*. Figure 1.6 depicts the average 2-day BHAR on Monday through Thursday in the notification week in panel A and the 2-day BHAR on the first day after the award announcement in panel B.[28] The figure shows that the stocks recommended by analysts barely making it to the finalist groups experience significantly higher market reaction in the notification week than those who do not make it. However, the discontinuity turns negative right after the award announcement. Regressions in Table 1.8 formalize the findings.[29]

It is evident that information about finalists is leaked before the award announcement, but how so? One possibility is via notified finalists and their brokerages. Past literature has shown that analysts tip their institutional clients prior to releasing stock recommendations (Irvine, Lipson, and Puckett, 2012), and that analysts in brokerages who rely more on trading commission are more susceptible to biases (Cowen, Groysberg, and Healy, 2006). Therefore, brokerages relying more on the business of institutional clients (e.g., mutual funds) may have higher incentive to

[28]NF does not record the exact date of notification. We thus err on the conservative side and use the average market reaction in the notification week as a proxy.

[29]In Section 1.3.5, we show that BHAR calculated using the benchmark return from a Fama-French five-factor model gives quantitatively similar results.

leak the information. We thus expect the stocks recommended by analysts in these brokerages to experience higher price fluctuations before and after the award announcement. We proxy for a brokerage's reliance on institutional clients by the proportion of its operating income from mutual fund trading commission in the mid-year before the award. We then split the sample by whether the proportion is above or below the sample median (0.6 percent). Consistent with our conjecture, Table 1.9 columns (1) through (4) show that the price fluctuations in the notification week and on the first trading day after the announcement concentrate among the stocks recommended by analysts in brokerages that rely more on trading commission.

Institutional investors are often considered as a force for price stability and market efficiency. But here, it is possible that their speculative trading based on insider information exacerbates the price fluctuations around the award announcement. Nonetheless, investors' inability to tell stale from new information is still the root cause for the overall difference in the market reaction between star winners' stocks and non-winners' stocks.

1.3.5 Robustness Checks

In this section, we alter the running variable, dependent variables, inference methods, degree of polynomial, bandwidth, and sample selection to test the robustness of our main findings. We first re-estimate the main results using rank and raw scores as running variables in Table A.9. The new estimates are within one standard deviation from the ones estimated using vote share as running variable. We next replicate the main results under various clusters including brokerage house, NF industry section, and stock industry sector in Table A.10. The statistical significance remains similar. We also re-estimate the regression with local quadratic regression in Table A.11. The new estimates are also within one standard deviation from those estimated with local linear regression. We further estimate the main results under varying bandwidths, including the bandwidths selected following Calonico et al. (2014). The estimates are stable across various bandwidths (Figure A.1). In addition, we change the stock selection criteria from the last stock recommended by each analyst within 1-30 days before the award announcement (or notification week) to the last three, five, or all stocks recommended during the same period in Table A.12. Lastly, we change the day range of stock selection from within 1-30 days before the award announcement (or notification week) to 61-90, 91-120, and 121-150 days before the award announcement (or notification week) in Table A.13. The estimates are smaller than the ones in the main RD sample, suggesting that investors respond more to more recent recommendations.

It is possible that stocks receive new recommendations right around the award announcement, which can bias our estimates in an unknown direction. We thus exclude stocks that receive any new recommendations between the Saturday and the Monday around the announcement in Table A.14 column (3) and those that receive recommendations between the Saturday and the Monday around the notification week in Table A.15 column (3). The estimates are similar to the main results. In addition, firms may issue announcements around the award announcement, which could also bias our findings. Therefore, in Table A.14 column (4) and Table A.15 column (4), we drop stocks that issue announcements between the Friday in the notification week and the Tuesday after the award announcement. The estimates are largely unchanged. To alleviate the concern that analysts may change recommendation strategy in the notification week, we exclude the stocks recommended in the notification week in Table A.14 column (5). The estimate is within half the standard deviation from the one estimated in the main RD sample.

Finally, we examine how trading volume responds to the award announcement. We construct the daily abnormal turnover by subtracting the natural log of the daily turnover by the average of natural log of daily turnover in the previous 30 to 60 days.[30] Table A.16 column (1) shows that the cumulative abnormal trading volume in the two days following the award announcement is 30 percent higher among the stocks with stale recommendations from analysts above the winner cutoff than those from analysts below the cutoff. This finding implies that the award announcement prompts investors to trade more stocks recommended by the star winners.

[30]The construction is in Appendix A.1.

We also examine the trading volume at the cutoff of finalist. Although the coefficients on *Win* dummy in Table A.16 columns (2) and (3) have the same signs as expected, they are not statistically significant from zero.

1.4 Post-Award Effect

1.4.1 Market Reaction

So far, we have demonstrated a clear discontinuity in market reaction to stale recommendations from analysts barely winning and those barely missing the award following the award announcement. We also show that the market does not react differently to new stock recommendations issued by winning and losing analysts shortly after the award announcement. To understand market reaction in a longer period of time, we now examine stock recommendations issued in the whole year after the award announcement. We measure market reaction to stock recommendations with the adjusted two-day BHAR following Loh and Stulz (2011). The specification is similar to that in regression 1.1. Besides brokerage and NF industry fixed effects, we also control for month, day-of-week, and year-and-stock-industry fixed effects.[31] We cluster standard errors by year-and-NF industry.

Table 1.10 column (1) reports the regression results. In the whole year after the award announcement, the market does not react differently to stock recommendations from winning and losing analysts. As a check for the validity of the RD design, we examine market reaction to stock recommendations issued in the year *before* the award announcement. Table A.17 column (1) shows that no difference in market reaction exists before the announcement.

[31]We do not control for day-of-week or year-and-stock-industry fixed effects before, because the market reaction is always measured on a Monday, and the sample size is too small to include year-and-stock industry fixed effects.

1.4.2 Information Production

Besides market reaction to stock recommendations, the response of brokerages and analysts themselves is also worth understanding. Since providing earnings forecasts is an essential job task for financial analysts, we now examine their forecast performance in the year after the award. Following Fang and Yasuda (2009), we measure forecast performance using forecast error, forecast bias, and forecast boldness based on the *last* earnings forecast from each analyst for each stock in a calendar year.[32] Table 1.10 reports the regression results. Analysts above the winner cutoff issue earnings forecasts with similar error, bias and boldness as those below the cutoff after the award announcement, as indicated by the insignificant estimates on $Win(t)$ dummy in panel A columns (2) to (4). However, panel B columns (2) and (3) show that analysts barely making it to the finalist groups issue more accurate and less positively biased earnings forecasts than those not making it to the groups. To ensure the validity of the RD design, we examine the earnings forecasts issued in the year before the award announcement. Table A.17 columns (2) to (4) show that the earnings forecasts issued by analysts above and below either award cutoff are similar in all three dimensions before the award announcement.

A natural question that follows is why the analysts barely making it into the finalist groups become better at earnings forecasts than those barely not making it, who are ex-ante similar to the former? One possibility is that brokerage houses allocate more resources to analysts with award designations or who are likely to become star analysts in the future.[33] To measure how favorably brokerage houses assign resources to these analysts, we construct three empirical proxies: (1) teammate quality, (2) team size, and (3) lead author status. Teammate quality for an analyst's earnings forecast is defined as the average of the baseline forecast error among the analyst's coauthors in the same forecast. For a solo-author forecast, teammate quality is defined as the

[32] Detailed construction is in Appendix A.1.

[33] Brokerage houses have incentives to do so because star analysts are shown to attract a higher market share of trading volume and investment banking deal flows for their brokerage houses (Clarke, Khorana, Patel, and Rau, 2007; Niehaus and Zhang, 2010).

average of the baseline forecast error among all coauthors of the analyst in the year after the award. Team size for an earnings forecast is the number of analysts putting their names on the report. Lead author status is an indicator of whether an analyst's name is listed the first in an earnings forecast.[34] Since we are interested in the resource allocation in the whole year after the award, we include *all* earnings forecasts issued in the period. Table 1.10 columns (5) through (7) report the regression results. Overall, analysts making it to the finalist groups are assigned to larger and better teams than those not making it. One top of that, analysts barely winning the award are more likely to be lead authors than those barely missing the award. As a check for the validity of the RD design, we examine the three proxies in the year before the award announcement. Table A.17 columns (5) through (7) show that all proxies change smoothly across the award cutoffs.[35]

Finally, past papers have shown that conflicts of interest play a role in sell-side research. For example, brokerages reward analysts who promote stocks (Hong and Kubik, 2003), and analysts' earnings forecasts are influenced by their desire to win investment banking business (Chan, Karceski and Lakonishok, 2007). However, given the improvement in forecast performance documented above, is it possible that star finalists manage to overcome such conflicts of interest? To understand this possibility, we examine the heterogeneity in the improvement in forecast performance by the level of conflicts of interest. In Table 1.11, we split earnings forecasts by whether the covered firm has ever had any underwriting relationship in initial public offering, right issue or rationed shares with the analysts' brokerages before the forecasts. Interestingly, the improvement in forecast performance among star finalists only comes from forecasts for firms that are *not* affiliated with the brokerages. Among forecasts for firms affiliated with the brokerages, coefficients on *Win(t)* dummy for forecast error and forecast bias are positive, albeit insignificant. Taken together, our findings suggest that award designations could attract favourable resources

[34] In China, authors are listed in a descending order of hierarchy. The one listed first is usually the team leader. Lead author status is coded as missing for solo-author forecasts.

[35] All the above results are robust to dropping solo-author earnings forecasts, which account for 36.6% of the regression sample. Results are available upon request.

to the analysts, which improves their forecast performance. However, these analysts still have incentives to issue biased forecasts for certain firms to maintain investment banking business.

1.5 Conclusion

This paper studies the role of the media in financial markets. We exploit a series of exogenous attention shocks generated by the announcements of a financial analyst award on the front page of a high-profile financial magazine. We document a clear discontinuity in the return between the stocks with stale recommendations from analysts barely winning the award and appearing on the front page and the stocks with stale recommendations from analysts barely missing the award. However, the return fully reverses within 20 trading days. Evidence is consistent with the notion that investors are drawn to star winners' stale recommendations due to heightened media exposure. Investors' inability to tell stale from new information can be a root cause for the overreaction.

The documented investor behavior is generalizable to markets where information gathering is costly and investors have limited attention. However, whether the investor behavior can cause an observable price impact depends on the proportion of uninformed investors and the level of limit to arbitrage in the market. In markets which are dominated by retail investors and impose short sale restrictions, such as those in China and many developing countries, the investor behavior documented here is likely to have an observable price impact. In other words, the media can make price less efficient in the short run by drawing investor attention to stale information in these markets. Our findings paint a more comprehensive picture of the role of media in financial markets. Not only can the media facilitate information incorporation into asset prices, it can also introduce noises to the markets.

In addition, we find suggestive evidence that institutional investors' speculative trading based on leaked award information amplifies the price fluctuation around the award announcement. These findings highlight that insider trading can exacerbate price fluctuation and market inefficiency.

Finally, we find that brokerages assign awardees to larger and better teams, and these analysts issue more accurate and less biased earnings forecasts than others in the year after the award. Favourable resource allocation to analysts with award designations may contribute to the persistent difference in performance between star and non-star analysts documented in the literature. However, the better performance is only among earnings forecasts for firms not affiliated with the brokerages. Reputation and favourable resources obtained from winning the award may facilitate analysts' research, but analysts still have incentives to issue biased reports for certain firms to main investment banking business.

In conclusion, the media and its induced attention are a double-edged sword. This is especially the case when investors do not effectively distinguish stale from new information. As a potential remedy, the media is encouraged to give due credit to the initial information source to help investors gauge the staleness of the information.

Acknowledgements

Chapter 1 is coauthored with Siyu Chen and is being prepared for publication. The dissertation author was the principal author on this chapter.

1.6 Figures and Tables

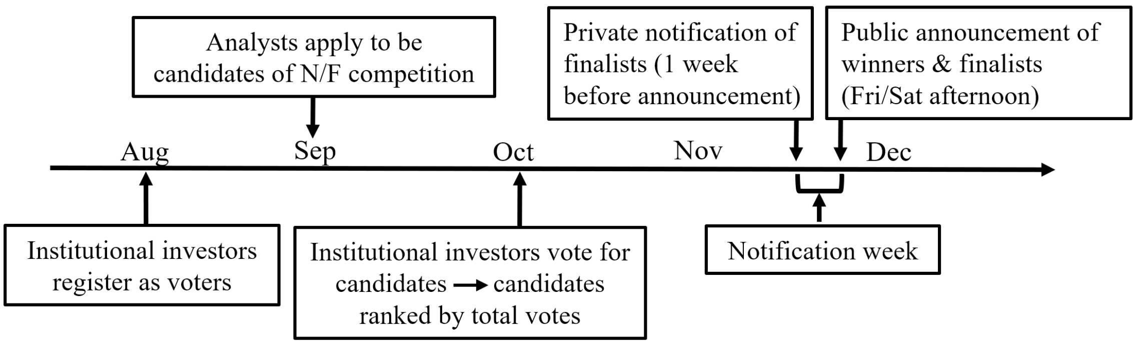
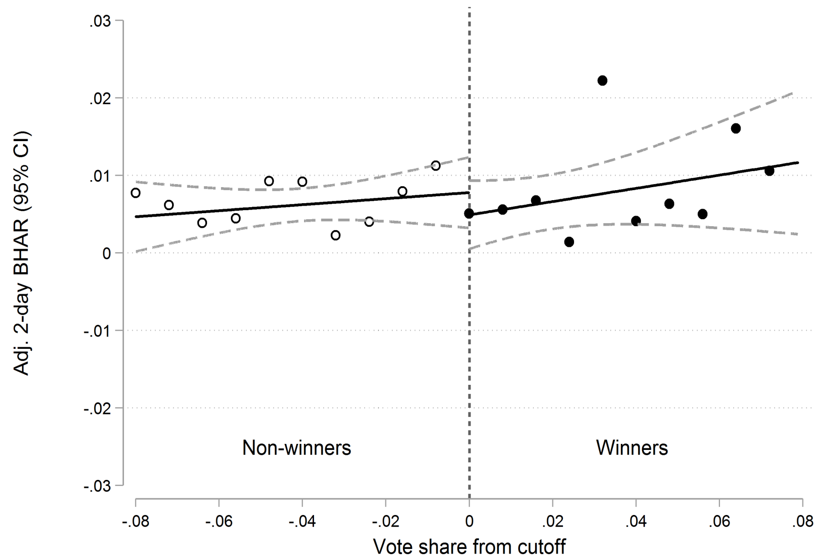
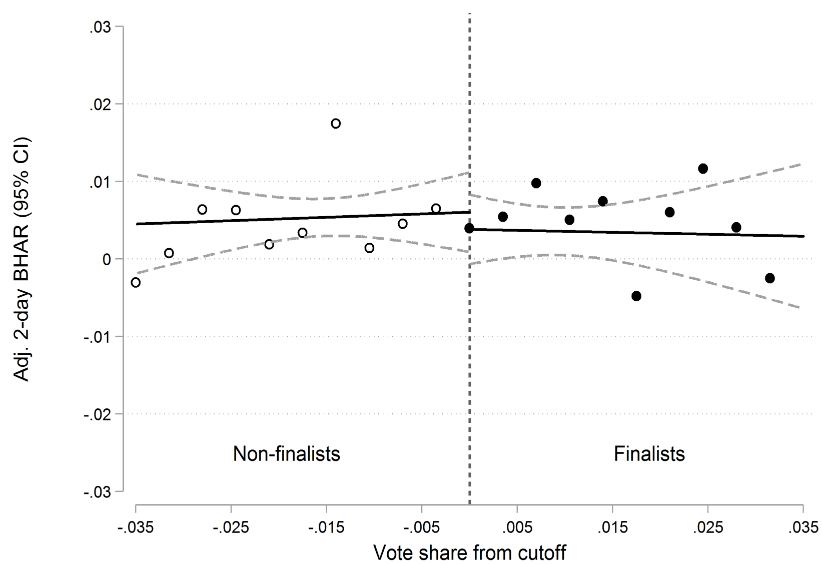


Figure 1.1: Timeline of New Fortune magazine's "Best Financial Analyst" ranking



A: Winner cutoff



B: Finalist cutoff

Figure 1.2: Placebo test – Market reaction on actual recommendation publication date

Notes: Each observation is the average two-day BHAR on the actual recommendation publication date in a 0.008 vote share bin (panel A) or a 0.0035 vote share bin (panel B) for the last stock recommended by each analyst within 1-30 days before the award announcement (panel A) or before the notification week (panel B). Dashed vertical line denotes the vote share at the cutoff of winner (panel A) and at the cutoff of finalist (panel B) for each industry each year and is normalized to 0. The solid lines are estimated using a local linear regression with triangular weights and individual-level data. The dashed lines denote the 95% confidence interval based on the standard errors clustered by NF industry and year.

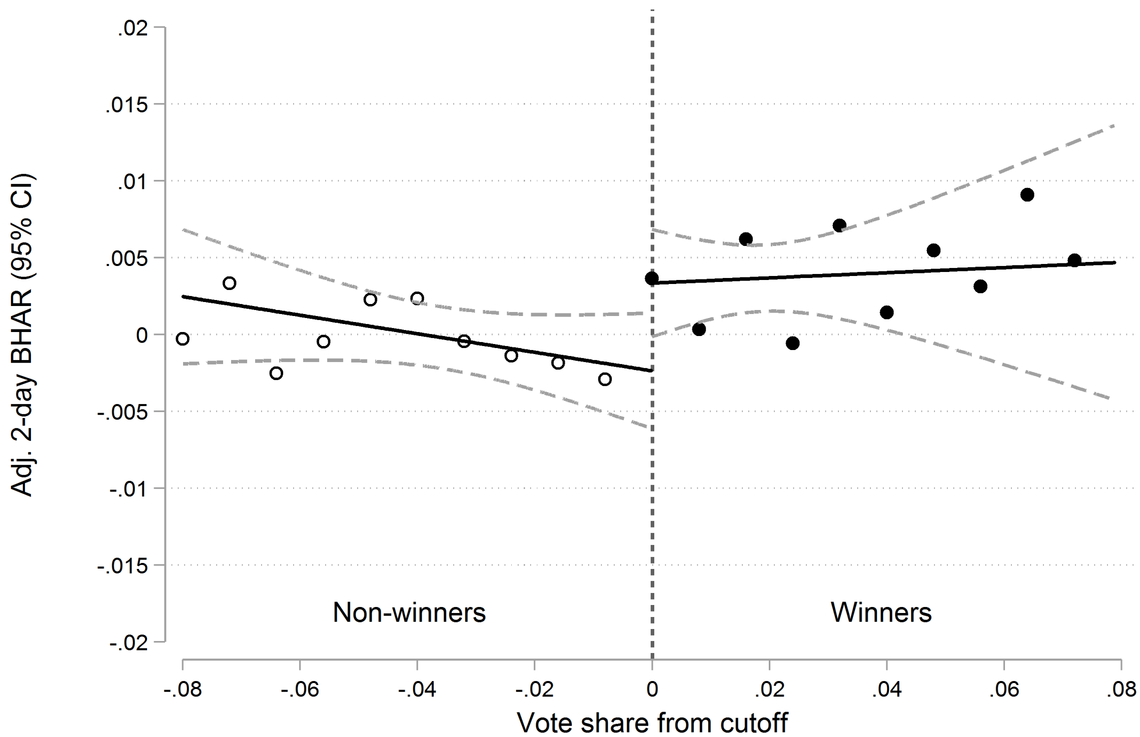


Figure 1.3: Main result – Market reaction after the award announcement
Winner cutoff

Notes: Each observation is the average two-day BHAR on the first trading day after the award announcement in a 0.008 vote share bin for the last stock recommended by each analyst within 1-30 days before the award announcement. Dashed vertical line denotes the vote share at the winner cutoff for each industry each year and is normalized to 0. The solid lines are estimated using a local linear regression with triangular weights and individual-level data. The dashed lines denote the 95% confidence interval based on the standard errors clustered by NF industry and year.

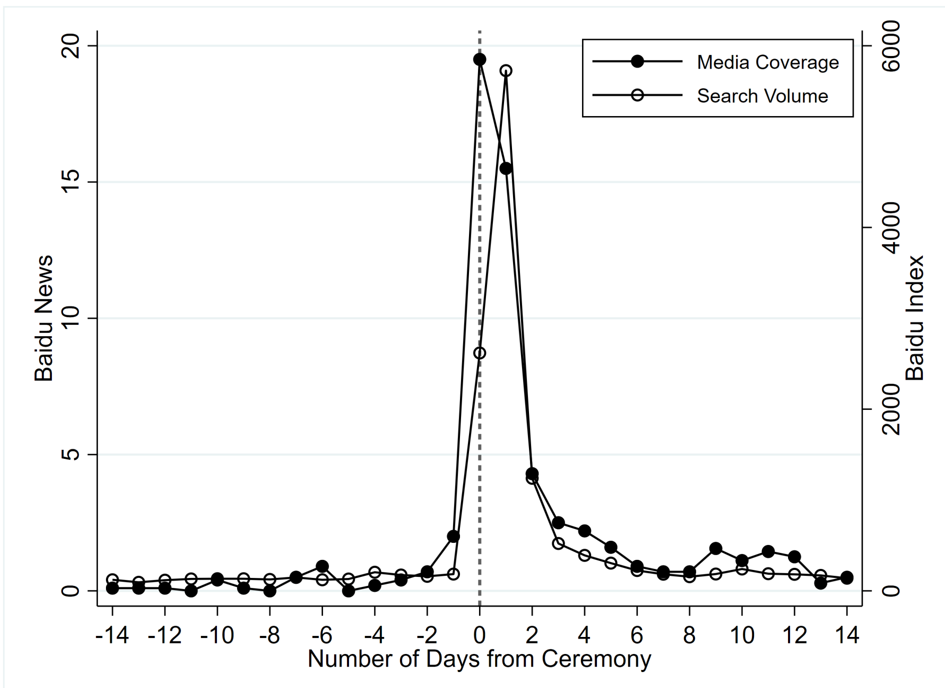


Figure 1.4: Media exposure and search volume of award announcement

Notes: The dashed vertical line donates the day of award announcement. The line with solid dots and the line with hollow dots depict the average number of news articles found on Baidu and the search volume on Baidu (”Baidu Index”) mentioning phrases ”New Fortune Best Financial Analyst”, respectively.

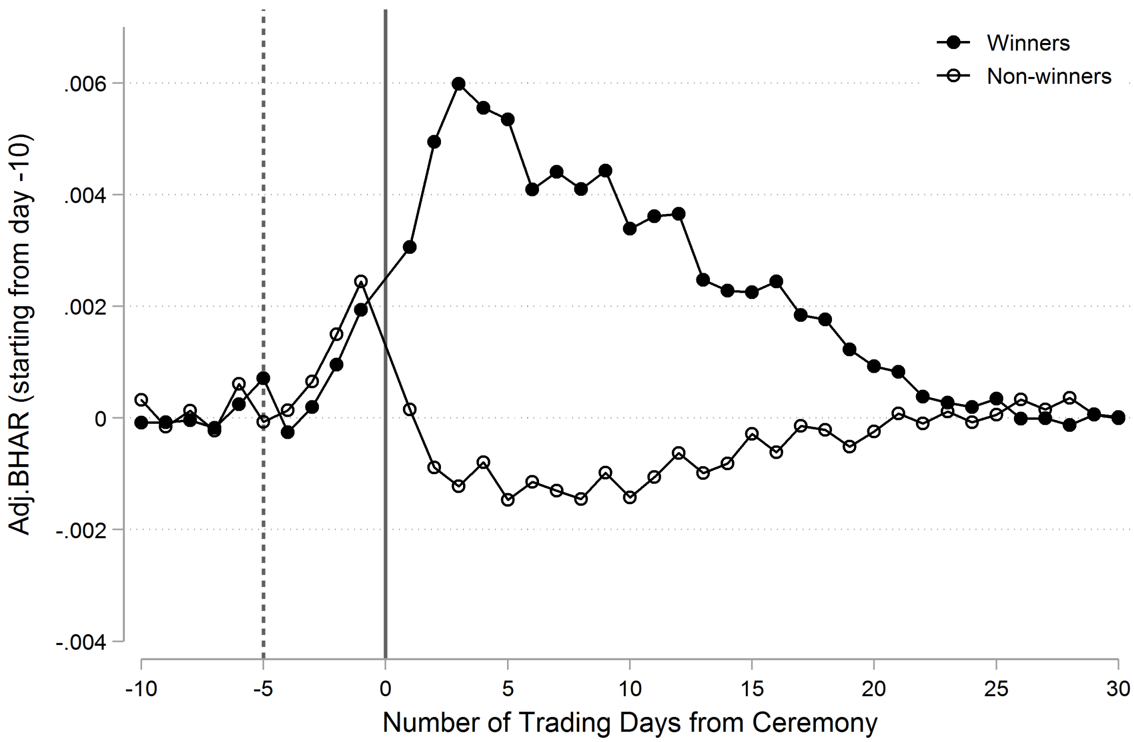
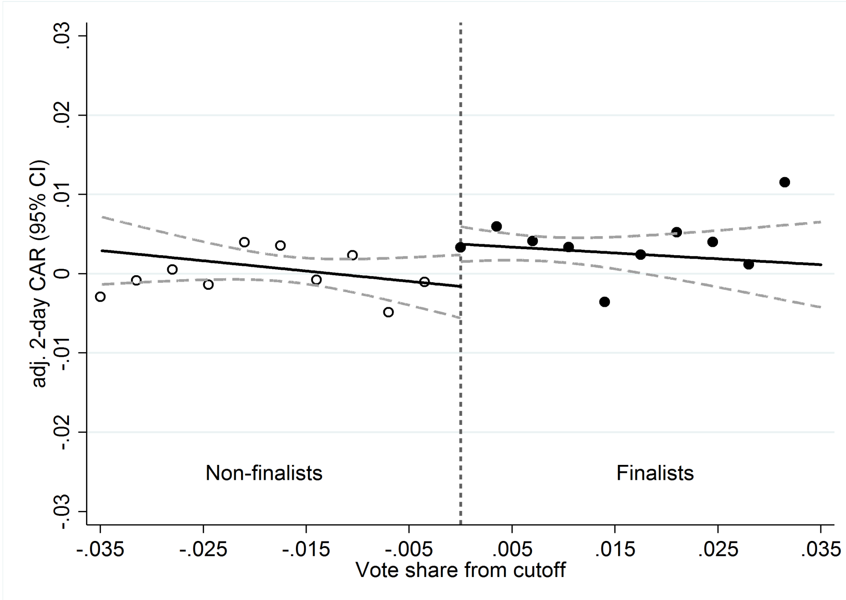
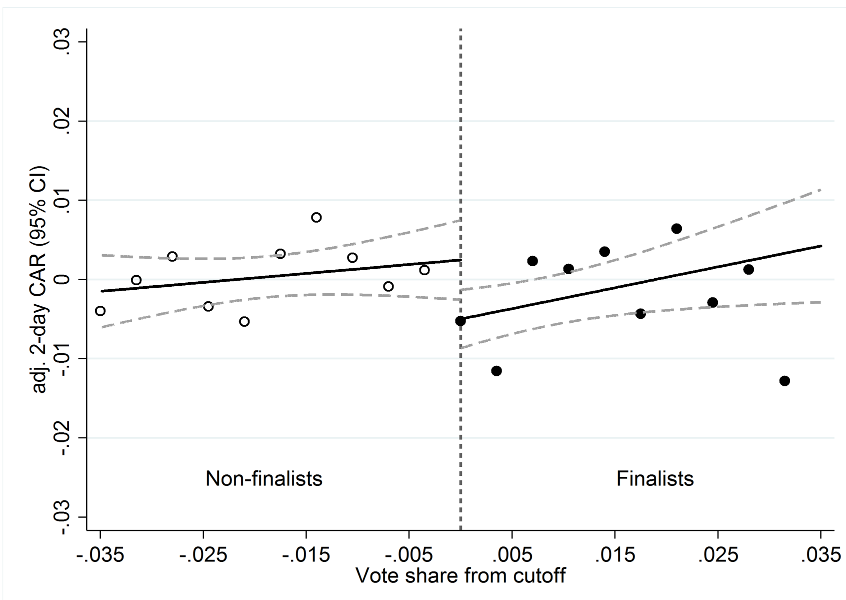


Figure 1.5: Buy-and-hold abnormal return around the award announcement
Winner cutoff

Notes: This figure plots the BHAR from the 10^{th} trading day before the award announcement to the 30^{th} trading day afterwards for stocks recommended by analysts in the main RD sample. The solid vertical line and the dashed vertical line mark the award announcement and the beginning of the notification week, respectively. The line with solid dots denote the average BHAR among the stocks recommended by analysts above the winner cutoff; the line with hollow dots denote the average BHAR among the stocks recommended by analysts below the cutoff.



A: In the notification week



B: After the award announcement

Figure 1.6: Main result – Market reaction in notification week and after award Finalist cutoff

Notes: Each observation is the average two-day BHAR from Monday to Thursday in the notification week (panel A) and on the first trading day after the award announcement (panel B) in 0.0035 vote share bin for the last stock recommended by each analyst within 1-30 days before the notification week. Dashed vertical line denotes the vote share at the finalist cutoff for each industry each year and is normalized to 0. The solid lines are estimated using a local linear regression with triangular weights and individual-level data. The dashed lines denote the 95% confidence interval based on the standard errors clustered by NF industry and year.

Table 1.1: Summary statistics

	(1)	(2)	(3)	(4)	(5)	(6)
		Winner cutoff			Finalist cutoff	
VARIABLES	Main RD recommendation	Post-award recommendation	Post-award forecast	Main RD recommendation	Post-award recommendation	Post-award forecast
Baseline 2-d BHAR	0.00474	0.00519	-	0.00424	0.00455	-
	(0.03492)	(0.03545)		(0.03493)	(0.03516)	
Baseline forecast error	-	-	0.03115	-	-	0.03173
			(0.03619)			(0.03626)
Vote share	0.06966	0.08234	0.08626	0.05727	0.05882	0.06368
	(0.04525)	(0.02828)	(0.01941)	(0.02561)	(0.02996)	(0.01732)
Win(t)	0.36808	0.48263	0.59444	0.50551	0.50966	0.63638
	(0.48244)	(0.49971)	(0.49102)	(0.50020)	(0.49991)	(0.48106)
Stock rating	4.33616	4.34307	-	4.25184	4.32274	-
	(0.62984)	(0.60204)		(0.63550)	(0.61884)	
Observation	1,535	34,113	11,860	1,088	42,805	12,051
Analysts						
Male	0.73354	0.73044	0.73439	0.72222	0.70999	0.72429
Master+	0.89962	0.89364	0.89416	0.89423	0.89606	0.90613
Experience	3.33908	3.32029	3.32029	3.35804	3.30420	3.16939
Win(t-1)	0.28084	0.30164	0.30164	0.35623	0.39376	0.33159

Notes: This table reports the mean and standard error (in parentheses) for our regression samples. Columns (1) and (4) are at the analyst × stock × date level, columns (2) and (5) are at the stock recommendation level, and columns (3) and (6) are at the analyst × stock × year level. *Baseline 2-d BHAR* is the average adjusted two-day BHAR for stock recommendations issued in the year before the award announcement. *Baseline forecast error* is the average forecast error of stock forecasts issued in the year before the award announcement. *Vote share* is an analyst's score divided by the sum of scores among top 15 analysts (teams) in a NF industry in a year. *Win(t)* equals 1 if the analyst is above the winner cutoff (columns(1) to (3)) or finalist cutoff (columns (4) to (6)) in the year of an award announcement. *Stock rating* is the analyst's rating for a stock: strong sell=1, sell=2, neutral=3, buy=4, and strong buy=5. *Master+* is an indicator of the analyst having master or above degree. *Experience* is the number of years since the analyst made his/her first recommendation till the year of the measurement.

Table 1.2: Validity of RD – Analyst characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline performance			Demographics		
VARIABLES	2-d BHAR(t-1)	Forecast error(t-1)	Win(t-1)	Male	Master+	Experience
Panel A: Winner Cutoff						
Win(t)	-0.00021	-0.00133	-0.04864	-0.01449	0.01640	0.07486
	(0.00103)	(0.00136)	(0.05468)	(0.03822)	(0.03237)	(0.18638)
Outcome mean	0.00474	0.03364	0.28084	0.73354	0.89962	3.2967
Observations	1,702	1,606	1,702	1,610	1,614	1,702
R-squared	0.189	0.269	0.160	0.110	0.126	0.109
Number of clusters	146	142	146	146	144	146
Bandwidth	0.080	0.080	0.080	0.080	0.080	0.080
Panel B: Finalist Cutoff						
Win(t)	0.00061	0.00299	-0.07241	0.01601	0.01911	-0.33510
	(0.00123)	(0.00198)	(0.06244)	(0.05050)	(0.04167)	(0.23876)
Outcome mean	0.00424	0.03208	0.35623	0.72222	0.89423	3.35804
Observations	1,106	1,035	1,106	1,044	1,040	1,106
R-squared	0.229	0.295	0.188	0.143	0.211	0.121
Number of clusters	143	141	143	143	143	143
Bandwidth	0.035	0.035	0.035	0.035	0.035	0.035

Notes: The data are at the analyst $\times$ year level consisting of all analysts whose margin of vote share is within 0.08 from the winner cutoff (panel A) and 0.035 from the finalist cutoff (panel B). *Win(t)* equals 1 if the analyst is above the cutoff of winner (panel A) or finalist (panel B) in year t. $t-1$ refers to the year before the award announcement. Variables are defined in the same way as in Table 1. Sample size varies due to missing dependent variables. All regressions control for year, brokerage house, and NF industry fixed effects, and are estimated using a linear RD model with triangular weights. Standard errors in parentheses are clustered by NF industry $\times$ year. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.3: Validity of RD – Stock characteristics

VARIABLES	(1) Market cap	(2) Under-valued	(3) P/E ratio	(4) Beta	(5) Momentum
Panel A: Winner Cutoff					
Win(t)	-8.60010	-0.01561	18.80812	0.01528	0.00829
	(7.37266)	(0.03225)	(16.85633)	(0.02016)	(0.00995)
Outcome mean	55.92195	.41704	55.51162	1.09219	.01311
Observations	1,525	1,525	1,490	1,494	1,506
R-squared	0.519	0.500	0.167	0.570	0.601
Number of clusters	146	146	146	146	146
Bandwidth	0.080	0.080	0.080	0.080	0.080
Panel B: Finalist Cutoff					
Win(t)	3.28350	-0.02396	-9.22075	0.04156	0.00376
	(7.91492)	(0.05293)	(28.63591)	(0.03202)	(0.01532)
Outcome mean	64.853	.44196	50.14221	1.08721	.00211
Observations	1,077	1,077	1,049	1,058	1,068
R-squared	0.483	0.580	0.178	0.594	0.613
Number of clusters	143	143	142	143	143
Bandwidth	0.035	0.035	0.035	0.035	0.035

Notes: The data are at the analyst $\times$ stock $\times$ date level consisting of the last stock recommended within 1-30 days before the announcement by an analyst whose margin of vote share is within 0.08 from the winner cutoff (panel A) and 0.035 from the finalist cutoff (panel B). Sample size varies due to missing dependent variables. *Market cap* is stock's 3rd-quarter market capitalization measured in billion CNY. *Under-valued* equals 1 if stock's 3rd-quarter book-to-market ratio ≤ 1. *P/E ratio* is stock's 3rd-quarter price-to-earnings ratio. *Beta* is stock's risk factor estimated from CAPM regression using daily data in the past 250 trading days starting from October 31^{st}. *Momentum* is stock's return with cash dividend reinvestment in the month before announcement, i.e., October. *Win(t)* is 1 if analysts are above the winner cutoff (panel A) or above the finalist cutoff (panel B). All regressions control for year, brokerage, NF industry, and stock industry fixed effects and are estimated using a linear RD model with triangular weights. Standard errors in parentheses are clustered by NF industry in column (1) and by NF industry $\times$ year in columns (2) through (8). *** p<0.01, ** p<0.05, * p<0.1.

Table 1.4: Main result – Market reaction after the award announcement
Winner cutoff

VARIABLES	(1) 2-d BHAR	(2) 2-d BHAR	(3) 2-d BHAR
Win(t)	0.00591**	0.00504*	0.00659**
	(0.00290)	(0.00303)	(0.00293)
Baseline mean	0.00474	0.00474	0.00474
Observations	1,535	1,535	1,535
R-squared	0.012	0.082	0.183
Number of clusters	146	146	146
Year FE	Yes	Yes	Yes
Brokerage FE	No	Yes	Yes
NF Industry FE	No	Yes	Yes
Stock Industry FE	No	No	Yes
IK bandwidth	0.080	0.080	0.080

Notes: The data are at the analyst $\times$ stock $\times$ date level consisting of the last stock recommended within 1-30 days before the award announcement by an analyst whose margin of vote share is within 0.08 from the winner cutoff. The dependent variable is the adjusted two-day BHAR on the first trading day after the award announcement. Baseline mean refers to the average two-day BHAR for the stock issued by the analysts in the year before the award announcement. *Win(t)* equals 1 if the analyst is above the winner cutoff in the year of the award announcement. All regressions are estimated using a linear RD model and triangular weights. Standard errors in parentheses are clustered by NF industry $\times$ year. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.5: Mechanism - Attention trading
Winner cutoff

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Media exposure		Winning frequency		Analyst coverage		Institutional holding	
VARIABLES	$\leq$ median 2-d BHAR	$>$ median 2-d BHAR	1^{st} time 2-d BHAR	Repeated 2-d BHAR	$\leq$ median 2-d BHAR	$>$ median 2-d BHAR	$\leq$ median 2-d BHAR	$>$ median 2-d BHAR
Win(t)	0.01150***	0.00181	0.01030***	0.00515*	0.01016**	0.00608*	0.01025**	0.00463
	(0.00418)	(0.00528)	(0.00389)	(0.00307)	(0.00481)	(0.00352)	(0.00433)	(0.00408)
Baseline mean	0.00467	0.00474	0.00466	0.00473	0.00482	0.00449	0.00458	0.00477
Observations	812	723	1,119	1,386	815	720	768	767
R-squared	0.268	0.289	0.239	0.204	0.245	0.397	0.323	0.265
Number of clusters	142	133	146	144	140	142	140	138
Bandwidth	0.080	0.080	0.080	0.080	0.080	0.080	0.080	0.080

Notes: The data are at the analyst $\times$ stock $\times$ date level consisting of the last stock recommended within 1-30 days before the award announcement by an analyst whose margin of vote share is within 0.08 from the winner cutoff. In columns (1) and (2), stocks are split by whether the number of news articles mentioning the issuing analysts in the 360-day period before the award announcement is below or above the sample median (11 mentions). In columns (3) and (4), stocks recommended by analysts above the winner cutoff are split by whether the analyst is a first-time or a repeated winner. In columns (5) and (6), stocks are split by whether the number of recommendations covering the stock in the 30-day period before the award announcement is below or above the sample median (6 recommendations). In columns (7) and (8), stocks are split by whether the percentage of institutional holding for the stock in the mid-year before the award announcement is below or above the sample median (6.19 percent). The dependent variable is the adjusted two-day BHAR on the first trading after the award announcement. *Win(t)* equals 1 if the analyst is above the winner cutoff in the year of the award announcement. Specifications mirror the one in Table 4 column (3). All regressions are estimated using a linear RD model and triangular weights. Standard errors in parentheses are clustered by NF industry $\times$ year. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.6: Mechanism - Ability signaling
Winner cutoff

VARIABLES	(1) 6-35 days 2-d BHAR	(2) 36-65 days 2-d BHAR	(3) 66-95 days 2-d BHAR
Panel A: Winners v.s. non-winners			
Win(t)	-0.00048	-0.00231	0.00209
	(0.00372)	(0.00319)	(0.00315)
Baseline mean 0.00474	0.00474	0.00474	
Observations	1,666	2,276	2,377
R-squared	0.155	0.117	0.132
Number of clusters	141	142	140
Bandwidth	0.080	0.080	0.080
Panel B: Repeated winners v.s. non-winners			
Win(t)	-0.00044	-0.00179	0.00205
	(0.00439)	(0.00320)	(0.00317)
Baseline mean	0.00473	0.00473	0.00473
Observations	1,479	2,253	2,371
R-squared	0.162	0.118	0.132
Number of clusters	138	141	140
Bandwidth	0.080	0.080	0.08

Notes: The data are at the recommendation level consisting of all recommendations newly issued in 6-35, 36-65, and 66-95 days after the award announcement. Panel A includes the recommendations from all winners and non-winners whose margin of vote share is within 0.08 from the winner cutoff. Panel B restrict the winners to be repeated winners. The dependent variable is the adjusted two-day BHAR on the actual recommendation publication date. *Win(t)* equals 1 if the analyst is above the winner cutoff in the year of the award announcement. Specifications mirror the one in Table 4 column (3) plus day-of-week fixed effects. All regressions are estimated using a linear RD model and triangular weights. Standard errors in parentheses are clustered by NF industry $\times$ year. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.7: Portfolio return for informed investors

VARIABLES	(1) Equal-weighted	(2) Equal-investment	(3) Value-weighted
Alpha	0.00176**	0.00176**	0.00180**
	(0.00076)	(0.00075)	(0.00088)
Rm-Rf	0.93627***	0.93495***	0.91247***
	(0.05483)	(0.05449)	(0.06703)
SMB	0.42937***	0.44149***	-0.04054
	(0.12380)	(0.12077)	(0.22705)
HML	-0.61191***	-0.59409***	-0.43806**
	(0.16058)	(0.16024)	(0.19788)
RMW	0.39726	0.40348*	0.28947
	(0.24108)	(0.24042)	(0.29120)
CMA	0.24567	0.23487	0.16329
	(0.21065)	(0.20980)	(0.25115)
Observations	100	100	100
R-squared	0.86443	0.86502	0.77631

Notes: This table reports the return of the portfolio formed by informed investors estimated using the Fama-French five-factor model. The portfolio includes the last stocks with buy or strong-buy recommendations from each star winners and failed finalists within 30 days before the award announcement. All Stocks enter the portfolio on the Monday in the notification week. The stocks recommended by failed finalists exit at the opening price on the first trading day after the award announcement, while the stocks recommended by star winners exit at the closing price on the Friday after the award announcement. Individual stock returns are aggregated to the portfolio level using equal-weighted, equal-investment and value-weighted method in column (1), (2) and (3), respectively. Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.8: Main result – Market reaction in notification week and after award
Finalist cutoff

VARIABLES	(1) 2-d BHAR	(2) 2-d BHAR	(3) 2-d BHAR
Panel A: Avg.in notification week			
Win(t)	0.00576**	0.00496**	0.00432*
	(0.00226)	(0.00246)	(0.00243)
Baseline mean	0.00424	0.00424	0.00424
Observations	1,088	1,088	1,088
R-squared	0.022	0.115	0.228
Number of clusters	143	143	143
Panel B: 1st trading day after award			
Win(t)	-0.00797**	-0.00958***	-0.01061***
	(0.00343)	(0.00326)	(0.00323)
Baseline mean	0.00424	0.00424	0.00424
Observations	1,088	1,088	1,088
R-squared	0.019	0.125	0.228
Number of clusters	143	143	143
Year FE	Yes	Yes	Yes
Brokerage FE	No	Yes	Yes
NF Industry FE	No	Yes	Yes
Stock Industry FE	No	No	Yes
IK bandwidth	0.035	0.035	0.035

Notes: The data are at the analyst × stock × date level consisting of the last stock recommended within 1-30 days before the notification week by analysts whose margin of vote share is within 0.035 from the finalist cutoff. Bandwidths are IK bandwidths in the corresponding sample. The dependent variable is the average adjusted two-day BHAR from Monday to Thursday in the notification week (panel A) and the adjusted two-day BHAR on the first trading after the award announcement (panel B). *Win(t)* equals 1 if the analyst is above the finalist cutoff in the year of the award announcement. All regressions are estimated using a linear RD model and triangular weights. Standard errors in parentheses are clustered by NF industry × year. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.9: Mechanism - Institutional investors and information leakage
Finalist cutoff

	(1)	(2)
	Trading commission from mutual fund	
	Above median	Below median
VARIABLES	2-d BHAR	2-d BHAR
Panel A: Avg. in notification week		
Win(t)	0.00613**	0.00168
	(0.00312)	(0.00248)
Baseline mean	0.00385	0.00445
Observations	553	535
R-squared	0.342	0.362
Number of clusters	118	116
Bandwidth	0.035	0.035
Panel B: 1st trading day after award		
Win(t)	-0.01239***	-0.00667
	(0.00396)	(0.00567)
Baseline mean	0.00385	0.00445
Observations	535	535
R-squared	0.347	0.346
Number of clusters	118	116
Bandwidth	0.035	0.035

Notes: The data are at the analyst × stock × date level consisting of the last stock recommended within 1-30 days before the notification week by analysts whose margin of vote share is within 0.035 from the finalist cutoff. We split the stocks by whether the percentage of their recommending analysts' brokerage house's mutual fund trading commission over its operating income in the mid-year before the award is above or below median (0.6 percent). The dependent variable is the average adjusted two-day BHAR from Monday to Thursday in the notification week (panel A) and the adjusted two-day BHAR on the first trading after the award announcement (panel B). *Win(t)* equals 1 if the analyst is above the finalist cutoff in the year of the award announcement. All regressions are estimated using a linear RD model and triangular weights. Standard errors in parentheses are clustered by NF industry × year. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.10: Post-award effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Market reaction	Earnings forecast			Resource allocation		
VARIABLES	2-d BHAR	Error	Bias	Boldness	Teammate error	Size	Lead author
Panel A: Winner Cutoff							
Win(t)	-0.00067	-0.00106	-0.00148	-0.00085	0.00068	0.15529*	0.07061*
	(0.00065)	(0.00186)	(0.00212)	(0.00070)	(0.00188)	(0.09189)	(0.03747)
Outcome mean	0.00461	0.03239	0.02310	0.01873	0.02838	1.91328	0.34602
Observations	34,091	11,858	11,858	11,858	30,072	30,244	30,244
R-squared	0.038	0.289	0.301	0.207	0.521	0.479	0.156
Number of clusters	146	142	142	142	140	146	146
IK bandwidth	0.042	0.02	0.02	0.02	0.02	0.02	0.02
Panel B: Finalist Cutoff							
Win(t)	0.00074	-0.00280*	-0.00370*	-0.00196	-0.00426*	0.18621**	0.03649
	(0.00070)	(0.00165)	(0.00210)	(0.00173)	(0.00232)	(0.10304)	(0.04460)
Outcome mean	0.00427	0.03314	0.02335	0.01919	0.02921	1.82289	0.32665
Observations	42,805	12,048	12,048	12,048	30,239	30,471	30,471
R-squared	0.033	0.296	0.291	0.189	0.549	0.476	0.141
Number of clusters	146	142	142	142	140	146	146
IK bandwidth	0.045	0.018	0.018	0.018	0.018	0.018	0.018

Notes: The sample for column (1) is at the recommendation level consisting of all recommendations issued within 0-365 days after the award announcement from analysts whose margin of vote share is within IK bandwidth. The dependent variable is the adjusted two-day BHAR on the recommendation publication date. The samples for columns (2) to (4) consist of the *last* forecast report for each stock issued within 0-365 days after the award announcement by each analyst whose margin of vote share is within IK bandwidth in the year after the award announcement. The samples for columns (5) to (6) consist of *all* forecast reports issued during the same period. *Teammate error* is the leave-one-out average of coauthors' baseline forecast error excluding the analyst in question (non-solo reports) or the average baseline forecast error among all coauthors of an analyst in the post-award period (solo reports), and smaller error means better quality. *Size* is the number of teammates in each forecast report. *Lead author* equals one if an analyst being listed as the first author and zero otherwise; this variable is coded as missing for solo-author forecasts. Standard errors in parentheses are clustered by NF industry $\times$ year. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 1.11: Heterogeneity - Post-award effects by affiliation

VARIABLES	(1) Forecast error	(2)	(3) Forecast bias	(4)	(5) Boldness	(6)
	Aff=1	Aff=0	Aff=1	Aff=0	Aff=1	Aff=0
Panel A: Winner Cutoff						
Win(t)	0.00152	-0.00085	0.00481	-0.00148	0.00315	-0.00096
	(0.00417)	(0.00223)	(0.00534)	(0.00237)	(0.00387)	(0.00072)
Outcome mean	0.02963	0.03286	0.01918	0.02376	0.01774	0.01881
Observations	1,726	10,132	1,726	10,132	1,726	10,132
R-squared	0.444	0.302	0.456	0.317	0.536	0.210
Number of clusters	135	142	135	142	135	142
IK andwidth	0.02	0.02	0.02	0.02	0.02	0.02
Panel B: Finalist Cutoff						
Win(t)	0.00316	-0.00299*	0.00047	-0.00359	0.00296	-0.00403**
	(0.00312)	(0.00162)	(0.00421)	(0.00221)	(0.00276)	(0.00186)
Outcome mean	0.02911	0.03372	0.01912	0.02396	0.01881	0.01881
Observations	1,520	10,528	1,520	10,528	1,520	10,528
R-squared	0.472	0.308	0.490	0.300	0.566	0.191
Number of clusters	135	142	135	142	134	146
IK bandwidth	0.018	0.018	0.018	0.018	0.018	0.018

Notes: Affiliation (Aff) equals 1 if the brokerage of the analyst has been a underwriter of the stock in question before the forecast report. Samples and dependent variables mirror those in Table 10 columns (2) to (4). *Win(t)* equals 1 if the analyst is above the winner cutoff (panel A) or the finalist cutoff (panel B) in the year of the award announcement. All regressions control for month, day-of-week, brokerage, NF industry, and year-by-stock industry fixed effects. All coefficients are estimated using a linear RD model and triangular weights. Standard errors in parentheses are clustered by NF industry $\times$ year. *** p<0.01, ** p<0.05, * p<0.1.

Chapter 2

Symbolic Awards at Work: A Regression Discontinuity Design

Abstract

This paper studies the effects of a non-pecuniary symbolic award on winners, losers, and their peers, using a regression discontinuity design. We identify newly recruited insurance salespeople who barely won a quarterly "Best Rookie" award and their counterparts who barely missed it in a large insurance company. Our main finding is that barely winners earn less life insurance commission than barely losers in the quarter following the award designation. Interestingly, the performance difference is mainly driven by winners earning less rather than losers' earning more. Several mechanisms, such as signaling, effort reallocation, licensing, mean reversion, conformity preference, and strategic reallocation across time or across teammates, are tested and ruled out. One mechanism, which we have empirical support for, is peer sabotage of winners triggered by the award designation. Finally, we examine spillover effects of the award and find no evidence that coworkers of winners and losers perform differently in any measurable aspects after the award.

2.1 Introduction

Non-pecuniary symbolic awards are prevalent in labor markets. A burgeoning literature has documented positive effects of symbolic awards on winners' subsequent performance in various organizational contexts.[1] Without monetary incentives, symbolic awards can still lead to recipients' performance increase for other reasons.[2] However, there can be downsides to symbolic awards. Papers on peer sabotage suggest that winners may decrease performance due to competitive or jealous peers (Lazear, 1989; Andiappan and Dufour, 2020).

This paper examines a symbolic award in a natural firm setting and presents some of the first causal evidence that symbolic awards can *decrease* winners' subsequent productivity. We provide empirical evidence supporting that peer sabotage is an underlying mechanism. Additionally, we find no spillover effects on winners' coworkers, indicating that the award reduces winners' subsequent performance without improving others' work performance.

The symbolic award studied in this paper is the quarterly "Best Rookie" award given out in the largest branch of a leading insurance firm in China (*the company*). This award recognizes the top ten out of around 800 newly recruited salespeople (*rookies*) in a quarter, based on their commission from selling life insurance in their first quarter in the company (*first quarter life insurance commission*). The award winners in a quarter are recognized at a company-wide meeting held at the beginning of the next quarter. Winners' ranking and their first-quarter life insurance commission are posted on the board during the meeting, while the information on non-winners is unknown to the public.

An ideal experiment to pin down the effects of the award requires observing two equally

[1] See Frey and Gallus (2017) for a review.

[2] First, winning an award facilitates one's access to resources (Chan, Frey, Gallus, and Torgler, 2014). Second, winners can use the award to signal their abilities and look for better career options (Spence, 1973; Dewatripont, Jewitt, and Tirole, 1999; Neckermann and Frey, 2013). Third, utility from positive self image can incentivize agents to work harder (Bénabou and Tirole, 2002; Kolstad, 2013; Breza et al., 2017). Fourth, desire to maintain good social image may incentivize agents to behave as expected or outperform expectations (Ariely, Bracha, Meier, 2009; Kosfeld and Neckermann, 2011; Dellavigna, List, and Malmendier, 2012; Neckermann et al., 2014). Finally, winners may feel a stronger identification with their firms (Akerlof and Kranton, 2005; Gallus, 2016) and reciprocate by working harder (Fehr and Schmidt, 1999; Kube et al., 2012).

accomplished rookies, only one of whom gets the award. Our regression discontinuity (RD) design approximates this. We compare the subsequent performance of two rookies whose first quarter life insurance commission is close to the award threshold — one narrowly winning (*barely winner*) and the other narrowly losing (*barely loser*).[3] Given the large number of rookies scattered across teams and the lack of real-time public rankings, rookies are unlikely to perfectly manipulate their performance to win the award. Supporting the validity of the RD design, we find that barely winners and barely losers are similar in demographics, attrition rate, baseline performance, as well as their teammates' baseline performance and characteristics. Since teammates are assigned *before* each award designation, we can also examine the spillover effects of the award by comparing the post-award performance of barely winners' and barely losers' non-rookie teammates.

Several features of this setting make it uniquely useful in estimating the effects of symbolic awards. First, the award is purely symbolic — winners receive no pecuniary prize or promotion.[4] Second, the award is non-repeated — each salesperson competes for the award only once. The ex-post effects of the award are not contaminated by salespeople's desire to win again. Third, the clearly-defined and closely-knit teams in the company form natural units which enable us to examine spillover effects of the award. Such natural units are usually unavailable in field data. Fourth, non-rookies are ineligible for the award, so the spillover effects on non-rookie teammates are not confounded by competition for the same award. Finally, the award was established years before our sample period. Analyzing it can shed light on how awards function in a well-established field setting.

Our analysis yields two main findings. First, barely winners earn 1,720 CNY (about 250 USD) less in life insurance commission than barely losers in the quarter after the award designation. The difference amounts to over 27 percent of the average first-quarter life insurance commission among top 20 rookies. Interestingly, the difference is almost entirely driven by

[3]A working paper by Larkin (2011) uses a similar identification strategy to examine salespeople's trade off behavior between pulling sales forward to win an award and delaying sales to earn higher commission.

[4]The award does not enter the promotion algorithm or alter salespeople's incentive schemes, and we find no evidence that barely winners receive better promotion or exit the firm earlier than barely losers.

winners earning less rather than losers earning more. Given the average commission rate of 15 percent, the above results imply a loss of 11,400 CNY (about 1,650 USD) in insurance sales per winner per quarter.

Second, the data allow us to explore the underlying mechanisms of the performance decrease. We find little evidence for mechanisms including signaling, effort reallocation, licensing, gaming, mean reversion, conformity preference, or strategic reallocation across teammates. One mechanism which is consistent with our findings and cannot be ruled out is peer sabotage of winners triggered by the award designation. Papers on peer sabotage (Lazear, 1989; Andiappan and Dufour, 2020) indicate that teammates may sabotage winners if they update belief on winners' likelihood to outcompete them for limited resources, or if they are jealous of winners' public recognition.

We then provide empirical evidence for peer sabotage. Firstly, we exploit the fact that performance observability among peers is a necessary condition for any peer effects to take place (Kandel and Lazear, 1992; Mas and Moretti, 2009; Bursztyn and Jensen, 2015). Under peer sabotage, the performance reduction should be more pronounced among barely winners with lower pre-award performance observability, because winning the award raises their observability more and induces a higher increase in peer sabotage. A feature of our setting is that manager's mentioning of top performers in team meetings is usually the only way for salespeople to observe others' performance. We thus proxy a rookie's pre-award performance observability by their rank of life insurance commission in their team before the award. Consistent with our conjecture, the difference in subsequent performance between barely winners and barely losers is larger among rookies who rank low in their team to begin with. Secondly, peer sabotage should be more severe in teams where winners compete directly with teammates for limited resources. We find that the difference in subsequent performance between barely winners and barely losers is indeed larger among rookies in teams with more severe competition. Importantly, the above findings cannot be explained by the pre-existing difference in the observables of rookies and their teammates between sub-samples.

Finally, we find no evidence that the award has any spillover effects on the peers of winners and losers. The non-rookie teammates of barely winners and those of barely losers do not perform differently in any measurable aspects in the quarter after the award designation. In other words, the award designation reduces winners' own performance without improving others' work performance.

This paper makes several contributions. First, we present some of the first causal evidence on the existence and the mechanisms of *negative ex-post* effects of symbolic awards on winners' productivity in the workplace. Most papers have documented *positive ex-ante* effects of awards (Kosfeld and Neckermann, 2011; Ashraf, Bandiera, and Lee, 2014), or *positive ex-post* effects on workers' productivity (Neckermann et al., 2014; Bradler et al., 2016; Chen and Lu, 2020) and retention rate (Gallus, 2016). To our knowledge, Gubler et al. (2016) is the only paper that documents negative *ex-ante* effects of symbolic awards in the workplace. They find that a symbolic attendance award crowds out internal motivation and performance in tasks *not* directly incentivized, despite also having positive effects on directly incentivized tasks among employees who had previously had punctuality problems. By contrast, our paper shows that winning a symbolic award decreases one's own productivity *ex-post* due to peer sabotage without improving others' performance.

In settings other than ordinary workplace, several papers also find negative effects of awards on winners' subsequent performance. Borjas and Doran (2015) study winners of Fields Medal and suggest that winners' publication performance worsens because they reallocate effort from writing papers to exploring new topics. Malmendier and Tate (2009) find that CEOs who win prestigious awards subsequently participate in more activities outside their firms at the expense of their firms' performance. In both papers, the main culprit for performance decrease is effort reallocation. Our paper adds to theirs by presenting a new mechanism, peer sabotage, through which awards lead to negative effects on winners. In the school settings, Robinson et al. (2019) finds that a retrospective attendance award provides moral license to receiving students and worsens their subsequent attendance, while Bursztyn and Jensen (2015) shows that higher-performing students reduce their effort to avoid being listed on an academic leaderboard due to peer pressure. Although the two papers provide valuable evidence that symbolic awards can affect performance negatively, they focus on student population, who are generally immature and whose cost of not properly investing in human capital will not materialize until years later. Our paper contributes by showing that symbolic awards can negatively affect the productivity of *adult* workers who are the main bread-earners in their households and whose productivity reduction takes an *immediate* toll on their take-home income.

Some papers on public performance ranking also find evidence for negative effects.[5] Blader et al. (2016) document that a public performance ranking among truck drivers improves or worsens the drivers' performance depending on whether the work site culture is competitive or collaborative. Ashraf (2017) demonstrates that sweater factory workers who outperform their friends in a public performance ranking decrease performance due to conformity preference. Our paper differs from theirs in that the main mechanism is different and that the award adds an extra layer of public recognition above and beyond the public ranking. For example, Medvec et al. (1995) show that athletes who are in the middle of the podium feel differently from those who barely make the podium. Kaniel and Parham (2017) show that mutual funds which barely make it to the "WSJ Category Kings" lists receive more fund flows than those which barely miss it, even when the *full* rankings are published. The higher public recognition associated with the award could exacerbate the negative effects on performance and incur higher costs for firms.[6]

Finally, this paper is the first to examine spillover effects of symbolic awards in firms, although past papers have studied such effects in non-firm environments (Guryan et al. 2009; Ager, Bursztyn, and Voth, 2016; Bradler et al., 2016; Moreira, 2016; Sequeira et al., 2016;

[5]See Ashraf (2017) for a review on recent papers about public performance ranking.

[6]In addition, the information set is also different — only top rankings are published in our setting while the full rankings are published in the public performance rankings. Therefore, workers below the cutoff face fewer confounding incentives in our setting than under public performance rankings.

Corelissen et al. 2017). Empirical evidence in firms is scarce due to two reasons. First, data on hierarchical relationships and peer groups in firms are hard to obtain (Warzynski, Smeets, and Waldman, 2017). Second, spillover effects can be confounded by direct effects of peers losing the award themselves, since most awards do not have eligibility criteria irrelevant to performance. In this paper, we solve both of these issues and find no spillover effects of the award designation on non-competing peers.

Overall, our findings point out an unexpected channel through which symbolic awards impose costs on firms. While this paper estimates the local average treatment effects pertaining to top performers, our findings have implications for the many firms which hand out symbolic awards. Top performers are important workers to study, because they deliver much higher productivity than average performers and serve as role models for others (Morgenroth, Ryan, and Peters, 2015; Lafortune, Riutort, and Tessada, 2018). Therefore, understanding how widely used incentives affect top performers is a crucial step towards increasing firm's productivity. Our findings suggest that paying attention to team dynamics and work environment where incentives are implemented and addressing peer sabotage properly will be a fruitful way to improve the effectiveness of incentives.

The remainder of the paper is as follows. Section 2 describes the organizational background and data structure. Section 3 outlines the empirical strategy. Section 4 presents the effects of the symbolic award on winners and losers and explores various mechanisms. Section 5 presents the spillover effects of the award. Section 6 concludes.

2.2 Organization Background and Data

2.2.1 Organization Background

The organization. Our testing ground is the largest branch of a leading insurance firm in China. Hereafter, we refer to this branch as "the company". The company is located in a city in

Eastern China. The city covers an area over 10,000 km^2 and has a population of about 7 million in 2013. The company has 12 sub-branches in the city, each consisting of about 30 teams and 400 salespeople on average. Between January 2013 and December 2016, our sample period, the company made a total of 1.78 billion CNY in insurance premium and employed a total of more than 20,000 salespeople (at least 4,000 salespeople in a single quarter).

There are two main job levels in the company: salesperson and manager. There are three sub-levels for salesperson ranging from 1 to 3 and three sub-levels for manager ranging from 4 to 6. Salespeople are responsible for selling insurance and referring new employees to the company. Besides these job tasks, managers are also responsible for managing salespeople in their teams and lower-level managers (e.g., level 5 and 6 managers oversee level 4). A team in the company is defined as a group of salespeople overseen by one manager. The team is formed by existing team members referring new salespeople to the company or by obtaining new salespeople from company job fairs.[7] Since managers' income partly depends on the performance of salespeople in their teams, managers hold regular team meetings to monitor and motivate them. In a typical team meeting, a manager will invite the salespeople who sold the most insurance in the previous week or so to share experience with other teammates. One thing worth pointing out is that there is no public performance ranking in the company, and salespeople do not know the performance of teammates other than the top performers.[8]

Salespeople have zero base salary and earn income from insurance commission and bonuses. They can sell two types of insurance products: life insurance and short-term insurance. Life insurance covers the insured person for the whole of life and pays out to the beneficiary upon the death of the insured. The premium for life insurance is paid annually according to a prearranged schedule; the responsible salesperson earns commission, an insurance-type-specific percentage of the annual premium, at the time of each payment. In the rest of the paper, "life

[7] The team formation is indeed non-random. However, because teammates are determined before each award designation, the non-random assignment will not invalidate our RD design.

[8] There is a social norm against asking others about their earnings, and salespeople rarely discuss earnings with their coworkers. This social norm is not uncommon in many countries, including the United States (Glassdoor, 2016).

insurance commission" refers to the commission from the *first* annual premium payment rather than the commission from later payments. Short-term insurance covers only for a short period of time and pays out for various prearranged contingencies. The premium for short-term insurance is paid in a lump sum when the contract is signed; the responsible salesperson earns a one-time commission as a predetermined percentage of the premium (*other insurance commission*).

Besides insurance commission, salespeople can earn various small bonuses. For example, they earn bonuses based on the number and the performance of employees that they *directly* refer (*referrals*) or directly manage (*subordinates*). But they do not receive bonuses based on the performance of team members who are neither their referrals nor subordinates. In other words, there are no team incentives for individual salespeople except the manager of the team. For an average salesperson, 65 percent of income is from life insurance commission, 10 percent from other insurance commission, and 25 percent from bonuses.

The company's promotion algorithm is based on two metrics: life insurance commission and the number of referrals. Salespeople are assessed at the beginning of each quarter based on their performance in the last quarter. For example, performance in January through March is evaluated at the beginning of April. They will be promoted to the next level if their performance in selling insurance and recruiting new employees in the last quarter is above the level-specific threshold; they will be demoted if their quarterly performance is below certain basic requirements.[9] All salespeople are eager to be promoted. For one thing, they get higher bonuses from their subordinates' insurance sales when they are at a higher job level, although the commission they earn from their own sales is fixed across job levels. For another, when their subordinate reaches a higher job level than themselves, they lose the bonus from the subordinate. Given the value of promotion, salespeople make every effort to understand the promotion algorithm which

[9] More specifically, a salesperson will be promoted from level 1 to 3 if her life insurance commission in a quarter reaches 1,500 CNY. At level 3, there are three possible career paths going forward. First, the salesperson will be promoted to the manager level (level 4) and start her own team if her quarterly life insurance commission is over 4,500 CNY and has at least two referrals in the company. Second, the salesperson will remain at level 3 if her quarterly life insurance commission is at least 1,500 CNY but has not qualified for promotion to level 4. Third, the salesperson will be demoted to level 2 if her quarterly life insurance commission is below 1,500 CNY.

governs the only way of promotion. Out-of-algorithm promotion is strongly discouraged and rarely occurs in the company. In our sample period, only 37 out of 6,707 promotions from rookies to higher job levels are considered inconsistent with the promotion algorithm.

In this company, salespeople sell insurance alone, although juniors often ask senior teammates for advice on selling insurance. Salespeople are responsible for developing their own selling regions and are allowed to sell to any customers. Given the large base of potential customers, it is uncommon for two salespeople to compete for one customer. However, salespeople often compete for internal resources, such as referral assignment and training opportunities, which are assigned based on managers' discretion. The number of referrals is an important criterion for promotion to the manager level. Besides recruiting on one's own, salespeople may be assigned a rookie who is recruited via company job fairs (i.e., not referred by anyone). Managers usually assign the unassigned rookie to their most promising subordinates who satisfy all criteria of promotion except the number of referrals.[10] In addition, the parent firm regularly introduces new insurance products and holds training sections in the headquarter to explain the product details. Only two to three salespeople per team can have the training opportunities, and managers usually pick the top performers.

The "Best-Rookie" award. Starting from the early 2000s, the company implements a quarterly award program to recognize top-performing rookies. Only rookies entering the company in a specific quarter can compete for the award for that quarter.[11] Rookies in each quarter are ranked according to their first quarter life insurance commission. At the beginning of the next quarter, the top ten rookies are presented with the "Best Rookie" award at a company-wide meeting. During the meeting, the rank of award winners and their life insurance commission

[10] For a manager to be promoted to a higher level, she needs at least two lower-level managers overseen by her. Therefore, many managers are better off assigning the unassigned new recruits to their promising subordinates than to themselves.

[11] As long as a rookie's contract starts in quarter t, she competes for *that* quarter's award, regardless of the exact entry date. Rookies who enter earlier in a quarter will have an advantage because they have more time to sell insurance. But rookies do not seem to strategically select the entry dates, as more rookies enter later rather than earlier in a quarter — 30 percent rookies enter in the first month of a quarter, 32 percent in the second, and 38 percent in the third.

are posted on the board, while the information on other rookies is unknown to the public. Given the high bar of winning the award, award winners are scattered across various teams — over 90 percent of managers with winners in their teams experience no more than two winners in the sample period. The award is purely symbolic. It does not come with monetary prizes or factor into the promotion algorithm, and salespeople face the same incentive scheme regardless of their award status.

In the rest of the paper, quarter t represents the first quarter when a rookie joins the company, and quarter $t+\tau$ represents the τth quarter after the rookie's first quarter in the company. Since the award is based on the performance in quarter t, we assign t to variables associated with the award designation, even though the award is physically handed out at the beginning of quarter $t+1$.

2.2.2 Data Source and Sample Construction

The company provides us with data covering all salespeople in the company between January 2013 and December 2016. The data consist of four parts:

1. Individual monthly performance, including insurance commission by detailed categories, total income, and the number of referrals. Since the award is given on a quarterly basis, we aggregate the performance to the quarterly level. We further winsorize all monetary outcomes at 1 percent level to reduce the influence of outliers. Main findings are robust to not winsorizing, and results are available upon request.

2. Personal information, including an anonymized identifier for each salesperson, their gender, age, years of education, urban status, home address, and contract start date. Urban status is one if a salesperson is from urban areas and zero if from rural areas. We define contract end date for a salesperson as the last day of the month after which she does not appear in the data again. If a salesperson leaves the company after the end of our sample period, we

code her contract end date as missing.

3. Hierarchical information, including salespeople's direct managers, referrers, teammates, and subordinates. Salespeople are defined to be teammates in a quarter if they share the same direct manager in that quarter.

4. A list of the winners of the "Best Rookie" award in each quarter. The list matches up perfectly with a list generated by the authors using the raw data on rookies' first quarter life insurance commission.

We assemble two samples for analysis. The first sample is for analyzing the effects of the award designation on winners' and losers' subsequent performance. For rookies hired in each quarter, we merge them with their personal information, hierarchical information, performance in that quarter and in all subsequent quarters, using their identifiers. There are 13,163 quarterly rookies during our sample period. Since we are interested in the effects of the award on subsequent performance, we require rookies to appear in the sample for at least two quarters.[12] We are left with 10,996 rookies (including 151 winners; the extra one is from a tied rank). We then calculate the optimal IK bandwidth using the standardized first quarter life insurance commission as the running variable (Imbens and Kalyanaraman, 2012).[13] For the RD regression, we focus only on the rookies who are within the optimal IK bandwidth (*main RD sample*), which consists of 1,837 rookies (including 115 winners).[14] Table B.2 reports the number of observations and the range of

[12]As a robustness check, we include the dropped rookies and use Heckman two-step to correct for the selection. Results are reported in column (1) Table B.1. Since only 14 out of 1,851 rookies within RD bandwidth dropped out in the second quarter (most are far from award thresholds on the left) and a perfectly exogenous predictor for this selection is hard to come by, we only report regressions *without* Heckman selection in the following sections.

[13]Standardized Commission$_{i,t} = \frac{commission_{i,t} - avg(commission_t)}{se(commission_t)}$. Running variable = standardized commission – the tenth standardized commission. Our findings are robust to using discrete rank as the running variable. Results are available upon request.

[14]The imbalance in the number of winners and losers is due to the award structure which employs a fixed quantity cutoff of 10. Such imbalance should pose little threat to our identification. First, we include triangular weight, so far away winners or losers contribute relatively little to the estimated discontinuity at the cutoff. Second, we conduct a battery of balance tests and find no discontinuities in any baseline characteristics between barely winners and barely losers. Third, non-winner rookies in the main RD sample are on average top 110 out of 800 rookies, who should also be deemed as top performers.

rank included in the main RD sample by year and quarter. In the appendix, we report main results using various samples, e.g., rookies who rank top 20^{th}, rookies who rank between 5^{th} and 15^{th}, and rookies whose baseline performance is within certain bandwidths varying between 2 and 3.5.

Table 2.1 panel A displays the summary statistics for the main RD sample. One thing worth pointing out is that 67 percent of rookies are female in this company, which is a norm in the Chinese insurance industry. Another thing to note is that the average years of education are around 14, equivalent to "some college". Since selling insurance is a job with no requirement on tertiary education, salespeople with college degree selected into this industry should not be viewed as similar to general college graduates. An average rookie joins the company roughly in the middle of a quarter and works for about 32 days in her first quarter in the company (*duration*). On average, the rookies earn 2,470 CNY in life insurance commission in their first quarter in the company and 2,110 CNY in the second quarter.

The second sample in our paper is for examining the effects of the award designation on the subsequent performance of winners' and losers' non-rookie teammates. We first identify the rookies who earn the highest life insurance commission among all rookies in a team, as there can be multiple rookies in each team, and the spillover effects, if any, are more likely to come from the best rookie within each team. These rookies are considered the source of impact and referred to as "participants". We then identify the *non-rookie* teammates of the participants in the participants' first quarter in the company. This process yields 17,409 quarter and salespeople pairs (7,411 unique salespeople led by 378 unique managers). We further require these salespeople to remain in the company in the quarter after the award designation of the participants. This leaves us with 15,471 quarter and salespeople pairs (6,538 unique salespeople led by 378 teams managers). These observations constitute the peer sample.[15] As shown in Table 2.1 panel B, around 60 percent of the salespeople in the peer sample are female, and the average years of

[15]The probability of a non-rookie teammate leaving the company after the award is about 14 percent for barely winners and 11 percent for barely losers. As a robustness check, we include the dropped teammates and use Heckman two-step to correct for the selection. Results are reported in column (2) Table B.1. Since results are similar, we only report regressions *without* Heckman selection in the following sections.

education are around 14. On average, they earn 2,210 CNY in life insurance commission in the corresponding participants' first quarter and 2,320 CNY in the second quarter. Table B.3 report summary statistics for the full rookie sample and the full peer sample with no bandwidth restriction.

2.3 Empirical Strategy

In this section, we explain our empirical strategies and identifying assumptions. We first examine the impacts of the symbolic award on winners and losers' subsequent performance in the main RD sample. The specification is as follows:

$$\begin{aligned} Y_{i,t+\tau} &= \beta_0 + \beta_1 Win_{i,t} + \beta_2 (StdCommission_{i,t} - Cut_t) \\ &\quad + \beta_3 Win_{i,t} \times (StdCommission_{i,t} - Cut_t) + \beta_4 X_{i,t+\tau} + \alpha_{t+\tau} + \varepsilon_{i,t+\tau}, \end{aligned} \tag{2.1}$$

where $Y_{i,t+\tau}$ is the outcome of interest for rookie i in the τth quarter after her first quarter in the company, such as life insurance commission, other insurance commission, the number of new referrals, and so on. τ equals 1 for our main regressions, as we are most interested in the immediate effects of the award designation on rookies' performance. We will extend τ when we analyze the performance dynamics. $Win_{i,t}$ equals 1 if i's life insurance commission in t is top ten among all rookies, and 0 otherwise. $StdCommission_{i,t}$ is the standardized life insurance commission in quarter t, calculated by subtracting a rookie's raw life insurance commission by the average life insurance commission of all rookies in quarter t and dividing the difference by the standard deviation of life insurance commission of all rookies in quarter t. The running variable is $StdCommission_{i,t} - Cut_t$, namely the difference between a rookie's standardized life insurance commission and the standardized life insurance commission at the award threshold (rank tenth) in quarter t. We construct the running variable with the standardized commission, so that we could compare rankings across quarters. We also include the interaction between $Win_{i,t}$ and

$StdCommission_{i,t} - Cut_t$ to allow different slopes of $StdCommission_{i,t} - Cut_t$ on the two sides of the award threshold. $X_{i,t+\tau}$ is a vector of control variables for i measured in $t+\tau$, including gender, age, age squared, urban status, and years of education. $\alpha_{t+\tau}$ is the quarter-by-year fixed effects, which controls for time-varying common shocks to the company. $\varepsilon_{i,t+\tau}$ is the error term.

The regression is estimated using local linear regression with triangular weights and IK bandwidths (Lee and Lemieux, 2010). We report heteroscedasticity-consistent standard errors in the main tables. We do not cluster standard errors by team or by year-and-quarter for the following reasons. First, according to Abadie et al. (2017), we should not cluster standard errors by team because there is no clustering by team in sampling or in the treatment assignment. Rookies in any teams can compete for the award, and the 115 winners in the main RD sample spread across 92 teams. Second, clustering standard errors by year-and-quarter is reasonable because this is the level of treatment assignment and the probability of being a "Best Rookie" varies across time due to the varying number of rookies. However, there are only 15 year-and-quarter cells in the sample period, which are too few to obtain consistent standard error estimates. Therefore, as robustness check, we use wild bootstrap (1,000 times) to obtain robust standard errors clustered by year-and-quarter (Cameron et al., 2008). We also try several other inference methods. Results are reported in Table B.4, and the standard errors are stable across methods.

β_1 is the coefficient of interest, which measures the impacts of award designation on the subsequent performance of winners relative to losers. There are two threats to identification: (1) perfect manipulation, a situation in which a rookie can perfectly manipulate her life insurance commission so that she is certain to win the "Best Rookie" award; (2) differential attrition, a situation where rookies on different sides of the award threshold exit at different rates post award. We provide evidence against these threats in the next section.

Besides the direct effects of the award designation on winners relative to losers, we also examine how the award designation affects their teammates. To conduct this analysis, we estimate the following regression using the peer sample:

$$\begin{aligned} Y_{j(i),t+\tau} = \gamma_0 + \gamma_1 Win_{i,t} + \gamma_2 (StdCommission_{i,t} - Cut_t) \\ + \gamma_3 Win_{i,t} \times (StdCommission_{i,t} - Cut_t) + \gamma_4 X_{j(i),t+\tau} + \alpha_{t+\tau} + \varepsilon_{j(i),t+\tau}. \end{aligned} \tag{2.2}$$

where $Y_{j(i),t+\tau}$ is the outcome of interest for non-rookie teammate j of participant i in the τth quarter after i's award designation. $X_{j(i),t+\tau}$ is a set of control variables for j measured in quarter $t+\tau$, including gender, age, age squared, urban status, years of education, job level, and tenure. Standard errors are clustered by team, because treatment assignment is clustered by team — *all* non-rookie teammates in a team with award winners are coded as treated. Unless otherwise noted, all else remains the same as in regression (1).

γ_1 is the coefficient of interest, which measures the impacts of participants' award status on their teammates. As long as there is no perfect manipulation among the participants or differential attrition among their teammates after the award designation, the RD design will identify the local average treatment effects of the award on winners' teammates relative to losers'. The non-random teammate assignment does not affect the validity of the RD design, as teammates of participants were determined *before* the designation.

2.4 Effects on Winners and Losers

2.4.1 Validity of RD

In this section, we examine the identifying assumptions of the RD design. Namely, (1) the award designation is as good as random around the award threshold, and (2) no differential post-award attrition exists around the threshold.

One threat to the randomness of the award assignment is perfect manipulation, a situation in which rookies know the award threshold and can manipulate their life insurance commission to be just above the threshold.[16] However, perfect manipulation is unlikely to occur in this setting. There are on average 800 rookies recruited in each quarter, and they spread out in over 290 teams located in 12 geographically dispersed sub-branches. Even though top performers within each team are known to their teammates, not all promising rookies are top performers in their teams. Moreover, the information in a team is hard to spread given the number and the dispersion of the teams. In addition, the award threshold varies substantially from one quarter to another (Figure B.1), and award winners spread across many teams — over 90 percent of teams with winners experience no more than two winners. Taking all the above into consideration, it is unlikely that salespeople can perfectly predict the award threshold and manipulate their commission accordingly.

To lend further support to the quasi-randomness of the award assignment, we compare the demographics and the baseline performance for rookies at the two sides of the thresholds. In Table 2.2 columns (1)-(5), we show the results regarding barely winners' and barely losers' gender, age, education, urban status, and the duration in their first quarter in the company. None of the outcomes exhibit discontinuities at the threshold, as indicated by the small and insignificant estimates on *Win* dummy. We further compare barely winners' and barely losers' first-quarter life insurance commission. Figure 2.1 plots the outcome on the y-axis and the running variable — the difference between a rookie's standardized first quarter life insurance commission and the standardized life insurance commission at rank tenth in the corresponding quarter — on the x-axis. The outcome changes smoothly across the award threshold. Regression results reported in Table B.5 column (1) tell the same story. But the non-existence of discontinuity at the threshold is not so surprising, because the running variable is essentially an affine transformation of the outcome. Therefore, we run additional placebo tests using other performance measures in the first quarter as outcomes, such as other insurance commission, number of referrals, and total income.

[16]Traditional manipulation test is not well suited in our setting, because the main sample is from the right tail of performance distribution where observation is sparsely distributed. The sparsity becomes even worse, as we have a fixed-quantity cutoff rather than a fixed-value cutoff. Manipulation test via STATA command "rddensity" won't pass even for fake cutoffs at rank 1-9 or 11-20.

Table B.5 columns (2)-(4) show that none of the above variables exhibit discontinuities at the award threshold.

Since rookies often ask senior teammates for advice on selling insurance, it is also important to examine the baseline performance and characteristics of these teammates. Overall, we find no significant difference in any of the examined variables. We will further explain these when we discuss the mechanisms in section 4.3 and the spillover effects in section 5.

Finally, Table 2.2 column (6) shows that there is no significant difference in the probability of exiting by the end of the first quarter between barely winners and barely losers.[17] In fact, only 14 out of 1,851 rookies within the RD bandwidth drop out in the quarter after the award, and most of them are on the far left side of the award cutoff. Given the low exit rate and the triangular weight centered at the award cutoff, the attrition should have a minimum impact on the RD estimates. As a robustness check, we use entry time, whether a rookie is referred by her manager, and rookie's distance to her manager's address to predict a rookie's early exit, and then apply Heckman two-step correction to the basic regression. The corrected estimates in Table B.1 are similar to our main estimates.

2.4.2 Main Results

In this section, we discuss the effects of the symbolic award on barely winners and barely losers' subsequent performance. Figure 2.2 displays their life insurance commission in the quarter following the award designation. The y-axis plots the outcome, and the x-axis plots the running variable. Though the slope of the fitted line on both sides of the award threshold is positive, there is a significant dip right above the threshold. In other words, though rookies who rank higher in their first quarter generally perform better in the second quarter, those who are just above the award threshold perform worse than those just below.

[17] Note that the sample size in column (6) is greater than the sample size for our main RD regression, because we include both rookies who have left and those who remain in quarter t+1 in column (6).

Table 2.3 reports the corresponding regression results. Our preferred specification in column (3) shows that barely winners' life insurance commission in the quarter after the award designation is 1,720 CNY lower than barely losers'. The difference amounts to over 27 percent of the top 20 rookies' average first-quarter life insurance commission.[18] The estimate remains stable as we adjust the controls from column (1) to (3). We report heteroscedasticity-consistent standard errors in Table 2.3, but the standard errors barely change under other inference methods as shown in Table B.4. We also conduct robustness checks by restricting our sample to rookies who rank top 20^{th}, rookies who rank between 5^{th} and 15^{th}, and rookies within bandwidths varying between 2 and 3.5, as well as using local quadratic regression in Table B.6, Table B.7, and Table B.8, respectively. Almost all estimates on Win dummy are within a third of standard deviation from the estimates in the main RD sample.

Since the RD estimates are *relative* in a cross-sectional sense, the discontinuity may result from winners slacking, losers working harder, or both. To understand which is the driving force, we plot the level change and the percentage change in the life insurance commission from rookies' first quarter to their second quarter in Figure 2.3.

Figure 2.3 shows that both the level and the percentage change among barely winners are significantly negative, while the change among barely losers centers tightly around zero.[19] To be precise, the average change in life insurance commission from rookies' first to second quarter is -2,439 CNY or -39.7 percentage points (*p* value¡0.01) among barely winners, whereas the change is 54 CNY or 0.22 percentage points (*p* value¿0.1) among barely losers.[20] These findings suggest that the performance difference between barely winners and barely losers is, surprisingly, driven by winners slacking off rather than by losers working harder. Simply put, barely winners respond to the award designation by reducing their life insurance commission, while barely losers do

[18]We use the average among top 20 rookies (6,209 CNY) rather than among all rookies in the main RD sample (2,469 CNY) as the benchmark, because the former is more relevant to the discontinuity at rank 10.

[19]Note that the changes in life insurance commission of the winners on the far right are above zero. We restrain from over-interpreting these changes, because the sample size there is small and the precision is low.

[20]We regress the change in life insurance commission on a constant with triangular kernel separately among barely winners and barely losers. The stated numbers are the estimated constants.

not respond much.[21] Given the average commission rate of about 15 percent, the performance difference in the insurance commission implies a revenue loss of 11,400 CNY ($= 1,720/0.15$) per winner per quarter.

A noteworthy point is that the minimum life insurance commission requirement for promotion to the manager level is 4,500 CNY. Since the average first-quarter life insurance commission among the winners in the main RD sample is over 6,220 CNY, they can fulfill the part of promotion requirement on insurance commission even when their commission decreases by 1,720 CNY in the second quarter.[22]

2.4.3 Plausible Mechanisms

What could be the mechanisms driving barely winners' performance decrease? In this section, we propose several non-mutually exclusive mechanisms and examine whether they are consistent with our empirical findings. We will start by examining the mechanisms that take place via winners' *own* behavioral change.

Signaling. Barely winners may use the award to signal high ability to outside firms and look for better jobs, which decreases their total effort on the job and lowers their life insurance sales. If this is the case, we expect barely winners' exit rate in the quarters immediately following the award designation to be higher than barely losers'. However, Table 2.4 columns (1) through (5) show that no significant difference exists in the cumulative exit rate between barely winners and barely losers in the six quarters after the award designation. Barely winners and barely losers are equally likely to have already left the company by the end of each quarter. Column (6) further shows that their total duration in the company within the sample period is similar to each other.[23]

[21] Some may argue that losers work harder to offset a common downward trend from the first to the second quarter while winners' effort remains unchanged, resulting in losers' unchanged performance while a decrease in winners'. We will explain why this interpretation is less likely than our preferred one in section 4.4.

[22] The other part of the promotion requirement is having at least two referrals in the company.

[23] There are two caveats: (1) we do not observe salespeople after they leave the company, so we cannot compare the quality of their next jobs; (2) we cannot pin down the total duration for salespeople who leave the company after the end of our sample period.

Effort reallocation. Tournaments have been shown to affect participants' post-award effort allocation (Malmendier and Tate, 2009; Borjas and Doran, 2015). Even if barely winners' total effort on the job remains constant, they may reallocate effort to tasks other than selling life insurance, such as selling short-term insurance and recruiting new salespeople, which are the two most important job tasks besides selling life insurance. If this is the driving force, we expect barely winners to perform better in these tasks. However, as shown in Table 2.5, barely winners perform no better in selling short-term insurance or referring new salespeople. Moreover, by selling less life insurance and not performing better in other tasks, barely winners make 2,032 CNY less in total income than barely losers in the following quarter, which amounts to 30 percent of the baseline mean.

Apart from the important job tasks specified in the labor contract, barely winners may reallocate effort to other (unobservable) tasks beneficial to their performance in the long term, such as on-the-job training. Although we cannot quantify the exact effort reallocation to each of these tasks, we can study the efficacy of the reallocation by comparing barely winners and barely losers' long-term performance in the company. We use salespeople's highest job level and probability of being promoted to manager in the sample period as an indicator for their long-term performance. Table 2.6 indicates that barely winners' highest job level ever reached is significantly lower than barely losers', and the two groups are not significantly different in the likelihood of being promoted to managers. Note that barely winners' poorer future performance cannot be explained by their early exit from the company, as barely winners and barely losers stay in the company for a similar amount of time (Table 2.4 column (7)). The above evidence suggests that barley winners do not reallocate their effort to tasks which can improve their performance in the long run.

Gaming the system. It is also possible that rookies game the award system by selling insurance to themselves and canceling the contract after winning. Since life insurance commission is the original amount in a quarter less canceled commission in that quarter, and whoever games the system is likely to cancel in the quarter right after the award, the gaming behavior could lead to barely winners' poor performance following the award. To examine this possibility, we compare the life insurance cancellation of barely winners' and barely losers' in each quarter from t to t+3 (quarter t being rookie's first quarter). Table 2.7 shows that barely winners experience 225 CNY less cancelation than barely losers in quarter t+1 (significant at 10 percent level) and the coefficients on Win dummy are all negative in the following quarters albeit insignificant. Given that the absolute magnitude of the coefficients on *Win* dummy is less than 300 CNY and the sign is opposite from what the gaming mechanism suggests, gaming is unlikely to be the main mechanism.

Licensing and rest on the laurels. Research on licensing suggests that when people feel that they have fulfilled their obligations to behave in certain ways, they subsequently become less likely to perform that behavior (Mullen and Monin, 2016). Winners of the award may feel licensed to reduce their effort going forward (see Robinson et al., 2019) or simply rest on the laurels. We would then expect winners who earn more in the first quarter to reduce their effort more, as they surpass the expectation of "being a good rookie" by a higher amount and should feel more licensed to reduce their effort or rest on the laurels. However, Figure 2.3 shows that the reduction is mainly driven by barely winners who exceed the cutoff by a smaller amount while those exceeding by larger amount do not reduce performance much. Therefore, licensing or rest on the laurels is inconsistent with our findings.

Strategic reallocation across salespeople. Besides winners themselves, their *teammates'* behavioral change may be the culprit. One possibility is that rookies' teammates pass on sales to rookies to help them win the award. The subsequent decreases in performance among barely winners are merely due to the winners returning the sales. Among all teammates, referrers and direct managers are the most likely to pass on sales, because having one's subordinates or referees named as the "Best Rookie" sends a positive signal of one's own ability. Under this mechanism, we would expect a decrease in life insurance commission among the referrers and the managers of barely winners in the quarter leading up to the award and an increase in their commission in the quarter afterwards. We thus use the quarterly life insurance commission of each rookie's referrer and manager as outcomes and repeat the RD regressions separately for each quarter from t-1 to t+3. We further control for the job level of the referrer and the manager and an indicator of whether the referrer is also the manager. Table 2.8 shows that the referrers and the managers of barely winners and those of barely losers do not perform differently in any quarters.[24] In addition to managers and referees, other senior members in teams can also pass on sales, so we examine their performance in Table 2.8 panel C. Again, the senior teammates of barely winners and those of barely losers do not perform differently in any quarters either. Results based on *monthly* life insurance commission exhibit the same pattern as shown in Table B.9. Overall, strategic reallocation across salespeople cannot explain our main findings.

Mean reversion and other mechanisms. Finally, one may argue that factors which are unrelated to rookies' or their teammates' behavioral change, such as mean reversion or other external shocks, could cause the decreases in winners' performance. However, this is unlikely. First, there is no explicit protocol in the company that assigns harder-to-sell regions to good performers; salespeople independently develop their own customer base. Second, the rank is based on three months of performance, so randomness in sales, a common cause of mean reversion, should have balanced out. Thirdly and most importantly, we have shown strong empirical evidence that the winning status is quasi-randomly determined around the award cutoff. Therefore, any mean reversion and external shocks that are uncorrelated with the quasi-random winning status should affect rookies right above and those right below the award cutoff similarly. As a result, these factors cannot generate the discontinuity we see in the data.

[24]Note that the smaller sample size in columns (1), (4), and (5) can be either due to mechanical truncation of sample period or due to managers and referrers not being in the company (either because they have not entered in column (1) or because they have already left in columns (4) and (5)).

2.4.4 Peer Sabotage

We have ruled out several mechanisms that could explain barely winners' performance decrease following the award designation. However, there is one mechanism that has not been ruled out and is consistent with our main findings: peer sabotage. Papers like Lazear (1989) and Andiappan and Dufour (2018) indicate that teammates may refuse to help or even sabotage winners.[25] Economically, teammates may perceive the winners as likely to overtake them and compete for limited resources such as referral assignment and training opportunities in the near future. Psychologically, teammates may be jealous that the winners stand out in front of the whole company, whereas they did not when they were rookies.

Performance observability among peers is a necessary condition for any peer effects (Kandel and Lazear, 1992; Mas and Moretti, 2009). In our setting, there is no public performance ranking within a team or in the company. The most common way for one to observe others' performance is team meetings where managers ask top performers to share selling experience. For a rookie who is already a top performer in her team, winning the award should not materially change her performance observability. In other words, top performers right above and right below the award cutoff should face similar sabotage, if any, and we do not expect to see discontinuity in their post-award performance. In contrast, for a rookie who is not a top performer in her team, winning the award reveals her performance publicly and can drastically change how her teammates perceive and treat her. As a result, low performers right above the award cutoff can face much higher sabotage than those right below the award cutoff, leading to discontinuity in post-award performance.

Because a team consists of both rookies and non-rookies, winning the "Best Rookie" award does not guarantee to be top in a team. We can rank salespeople in a team by their average monthly life insurance commission in the baseline quarter and proxy a rookie's pre-award

[25]It is common for junior teammates to ask seniors for information about insurance products or advice on selling insurance. One way of sabotage will be for senior teammates to stop doing so for winners. Another possibility is for senior teammates to delay transmitting work-related information to winners.

performance observability among teammates by this within-team rank.[26] We then split the main RD sample into rookies who rank top three in their teams (*high-rank*) and those who rank fourth and below (*low-rank*). As shown in Figure 2.4, the discontinuity in subsequent performance between barely winners and barely losers is indeed larger in the low-rank sample. There is a visually discernible dip in the post-award life insurance commission right above the award threshold in the low-rank sample (panel B) but not in the high-rank sample (panel A). In Table 2.9 panel A, we report the regression results. Barely winners earn 2,970 CNY less than barely losers after the award designation in the low-rank sample (column 1), while the difference is economically and statistically insignificant in the high-rank sample (column 2). The estimates in the two samples are statistically different (p value=0.027). As we change the cutoff rank from rank three to rank two and one in Table 2.9 panels B and C, the estimates in the high-rank sample converge to zero, consistent with a decreasing marginal increase in observability from winning when rookies' pre-award rank improves. In contrast, the estimates in the low-rank sample remain large and significant throughout.[27]

One thing to note is that the non-existence of discontinuity between barely winners and barely losers in the high-rank sample does *not* mean that their performance remains unchanged from the first to the second quarter. Getting frequent mentions during team meetings may also induce peer sabotage on these high-rank rookies, reducing their subsequent performance. But the key is that the difference in the reduction, if any, between winners and losers among the high-rank rookies should be much smaller. In untabulated results, we show that high-rank rookies experience slight decrease in life insurance commission from first to second quarter, and the reduction among winners is statistically indistinguishable from losers'.

[26]We use average *monthly* life insurance commission rather than the quarterly total to correct for rookies' differences in entry time.

[27]The estimates between high- and low-rank samples are not statistically significant at the conventional level in panel C. This is partially due to the tiny sample size in the high-rank group. But since the pattern in panel C is consistent with the other panels, we do not over-interpret this statistical insignificance. Also, we do not tabulate the results from sample splits using rank cutoff at fifth or below; the number of winners in the low-rank samples becomes too small to be meaningful.

We further show evidence that the heterogeneities by pre-award rank are not driven by pre-existing differences in rookies' performance, their characteristics, or their teammates' characteristics. To do so, we plot the RD graph using first quarter life insurance commission as placebo outcome (Figure B.2), examine the distribution of running variables (Figure B.3), and compare rookies' and their teammates' baseline characteristics in low- and high-rank samples separately (Tables B.11 and B.12). All validity checks pass.[28] As a robustness check, we split the main RD sample by rookies' performance percentile within their team rather than by rank. The results are qualitatively similar (Table B.10).

The above heterogeneities also refute the possibility that conformity preference drives winners' performance reduction. Theories on conformity (Bernheim, 1994; Akerlof, 1997) indicate that winners may reduce performance to avoid social punishment or psychological disutility from standing out. A key condition of conformity preference is peers' ability to observe winners' performance *after* the award designation (Bursztyn and Jensen, 2015). In other words, winners have no reasons to voluntarily reduce their performance to avoid social punishment or to feel better if their performance reduction cannot be observed by peers who they care about. Since there is a social norm against asking others for salaries and no public performance ranking in the company, and the performance of salespeople who rank low within a team will not be revealed later on, this condition is *not* met among the low-rank winners who reduce their performance the most. Therefore, conformity preference cannot explain our findings.

Next, we provide more evidence to corroborate the existence of peer sabotage. The first piece of evidence utilizes the variation in the level of competition across teams — peer sabotage should be stronger in teams with higher level of competition. We proxy the level of competition by whether winners compete directly with other teammates for limited resources such as referral assignment. The number of referrals is an important criterion for promotion to the manager level,

[28]We only tabulate the validity checks in subsamples split at rank fourth, because this split is the most even in terms of the number of winners and total observations. Validity checks using other rank cutoffs are available upon request.

but it is not easy to recruit on one's own. On average, it takes about 8 quarters before one refers their first salesperson to the company. Therefore, salespeople often compete to be assigned a referral who is recruited via company job fairs.[29] Managers have sole discretion in the assignment and often assign the unassigned new recruits to the most promising subordinates who satisfy all promotion criteria for team managers except the number of referrals.[30] Senior teammates on the verge of promotion thus have strong incentives to sabotage winners who are also on the verge of promotion by the second quarter in the company, because these winners will compete for the referral assignment with the senior teammates.

We thus split the main RD sample by whether the team has at least one senior teammate who are qualified for promotion except for the number of referrals.[31] Consistent with our conjecture, both Figure 2.5 and Table 2.10 show that the discontinuity in subsequent performance between barely winners and barely losers mainly comes from teams with senior competitors. The difference in estimates between the samples with and without senior competitors is significant at the 1 percent level. To ensure our findings are not driven by the pre-existing difference of rookies in different teams, we examine the baseline characteristics and performance of barely winners and barely losers in each sub-samples in Table B.13 and Figure B.4, respectively. There are no significant discontinuities in any of the examined variables in either sub-sample.

The above heterogeneities also speak against an alternative explanation mentioned in section 4.2 — losers work harder and winners do not change effort while both experience a downward trend. This is because the motivation effects from losing the award should not correlate with the degree of peer sabotage. If anything, losers will not want to work harder in teams with more severe peer sabotage, because they are also likely to be sabotaged if they stand out too much.

[29] On average, there is about one such unassigned new recruit in each team in each quarter.

[30] When a salesperson makes over 4,500 CNY in life insurance commission in a quarter and has at least two referrals in the company, she will be promoted to the manager level in the next quarter.

[31] More specifically, these senior competitors are at job level 3 and have the same number of referrals as the competing rookie at the beginning of quarter t+1 (either zero or one referral). In our analysis sample, about 35 percent of the rookies have at least one competitor at the beginning of quarter t+1.

The second piece of corroborating evidence exploits the difference in relationship between managers and rookies. Since managers can mitigate the peer sabotage, winners who are directly referred by managers should face less peer sabotage and experience less performance reduction. This is exactly what we find. Table B.14 shows that the discontinuity in performance between barely winners and barely losers who are direct referrals of their managers is only one half the size of the discontinuity between those who are not direct referrals.

Finally, the lack of a direct measure of peer sabotage limits our ability to quantify the level of peer sabotage. But overall, our findings support the notion that peer sabotage is a driving force behind winners' worse post-award performance.

2.4.5 Performance Dynamics

We have shown that barely winners earn less life insurance commission than barely losers in the quarter after the award designation. We now explore how this gap evolves over time.

Table B.15 reports the difference in quarterly life insurance commission between barely winners and barely losers from their first quarter in the company (t) to the fourth quarter (t+3). The gap in performance is no longer significant after quarter t+2, though it remains negative throughout. To understand the change better, we break down the quarterly performance to monthly performance and repeat the RD regressions using monthly life insurance commission from three months before the award ceremony to nine months after. The estimates and the 95 percent confidence intervals for the coefficient on Win dummy from separate RD regressions are plotted in Figure B.5. The plot indicates that barely winners' performance drops the most in the month following the award and the worse performance persists until the fourth month after the award. From the fifth month onwards, the difference between the two groups becomes small and insignificant.

One caveat for the dynamics is salespeople's endogenous exit. Though only 0.75 percent (14/1,851) of rookies within RD bandwidth exit by the end of their first quarter (quarter t), this cumulative exit rate increases to about 30 percent by the end of quarter t+3. Moreover, the longer the rookies stay before exiting, the closer they are to the initial award cutoff, and their exit matters more for the RD estimates. As a robustness check, we repeat the above regressions in a sample with always-stayer rookies. Results are quantitatively similar.

2.5 Spillover Effects on Peers

So far, we have established that winning the symbolic performance award induces peer sabotage and takes a toll on winners' subsequent job performance and take-home income. But peer sabotage in the form of refusing to help may benefit others, if the seniors spend the extra time on their own job tasks or on helping other junior teammates. Therefore, in order to understand the overall ex-post effects of the award designation, we also need to examine how the award designation affects the teammates of barely winners and barely losers.

To establish causality, the teammate assignment should be uncorrelated with the award designation. Recall that teammates are determined in a rookie's first quarter in the company *before* the corresponding award designation. As long as the award is quasi-randomly assigned around the award threshold, whether a rookies' teammates end up on the winning or the losing side of the threshold will also be quasi-random. To further verify the RD validity, we conduct balance tests using the baseline characteristics of participants' teammates. In Figure B.6, panels A through F, we plot teammate's various characteristics on the y-axis and the running variables of their corresponding participants on the x-axis. Teammates' characteristics, such as age, gender, education, urban status, job level, and firm tenure in the participants' first quarter in the company, all change smoothly across the award threshold. We also plot their performance in the participant's first quarter as placebo tests (Figure B.7). Performance such as life insurance commission, short-term insurance commission, number of new referrals, total income, and the likelihood of being promoted also change smoothly across the award threshold. The key takeaway is that teammates of winning participants are ex-ante similar to those of the losing ones except for the participants' award status.

We now move on to discuss the spillover effects of the award designation on participants' non-rookie teammates. Figure 2.6 plots the life insurance commission of participants' teammates in the quarter after the award designation on the y-axis and the corresponding participants' running variables on the x-axis. No significant discontinuity exists at the award threshold. Table 2.11 column (1) formalizes the RD graphs and shows that the coefficients on *Win* dummy are small and insignificant. We also find no significant difference in other performance measurements, such as other insurance commission, number of referrals, total income, and probability of being promoted in the quarter following the award designation in Table 2.11 columns (2) to (5). Overall, teammates of barely winning participants do not seem to perform differently from those of barely losing participants in the quarter after the award designation. In other words, there is little evidence that winners' performance reduction has any positive effects on their teammates' performance. Overall, peer sabotage induced by the award designation causes winners' worse post-award performance without improving their teammates' performance.

2.6 Conclusion

This paper studies the effects of a non-pecuniary symbolic award using an RD design. The main finding is that barely winners' performance worsens relatively to barely losers' in the quarter after the award designation. Our findings support the notion that peer sabotage triggered by the award designation is a main driving force for winners' performance decrease. In addition, we find no evidence for spillover effects of the award designation on the peers of barely winners and barely losers.

Our findings point out an unexpected channel through which symbolic awards impose costs on firms. While the effects estimated here are local average treatment effects pertaining to top performers, the findings have implications for the many firms that hand out symbolic awards. Top performers are important to study, because they deliver much higher productivity than average performers and serve as role model to motivate others (Morgenroth, Ryan, and Peters, 2015; Lafortune, Riutort, and Tessada, 2018). Understanding how widely used incentives like symbolic awards affect top performers is a crucial step towards increasing firm's productivity.

Finally, we cannot estimate the ex-ante incentive effects of the symbolic award, as the data at hand only cover periods when the award is in effect. We are thus unable to pin down the net effects of the symbolic award, which include both the ex-ante and the ex-post effects. But given the importance of top performers and the economically significant decrease in their post-award performance, the ex-post effects estimated here are still an important margin to study. We hope our findings encourage firms to pay attention to team dynamics and the general environment where incentives like symbolic awards are implemented. Properly addressing peer sabotage triggered by the symbolic awards can be a fruitful way to improve the effectiveness of the incentives in the future.

Acknowledgements

Chapter 2 is coauthored with Teng Li and has been submitted for publication. The dissertation author was the principal author on this chapter.

2.7 Figures and Tables

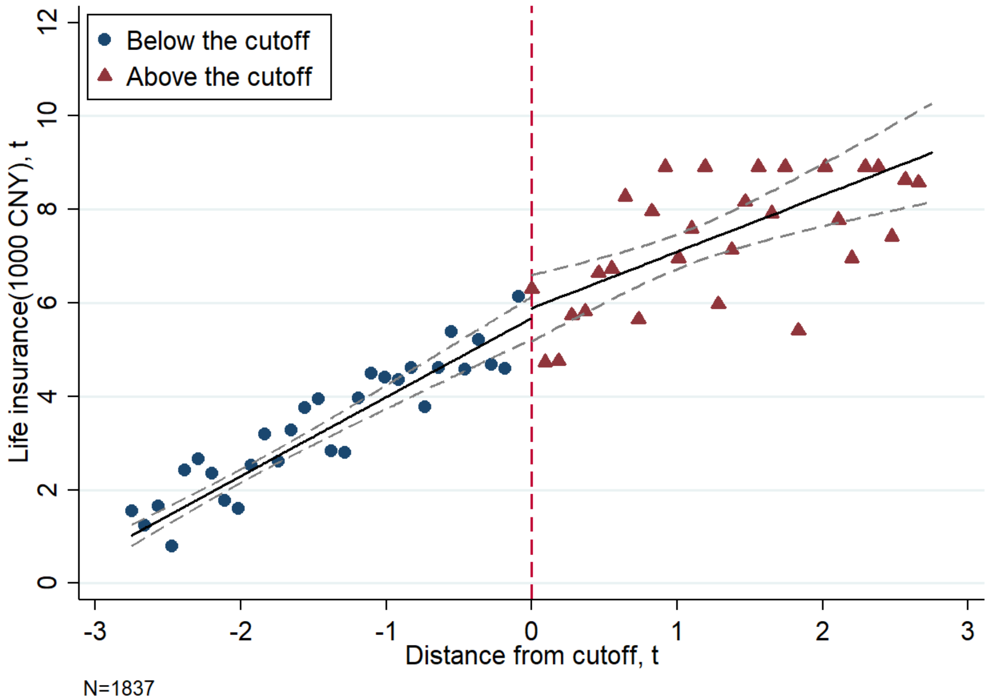


Figure 2.1: Life insurance commission in baseline quarter t

Notes: Each observation is the average life insurance commission in a rookie's first quarter in the company among rookies (main RD sample) in a 0.09 bin based on their standardized first quarter life insurance commission. Dashed vertical line denotes the 10th standardized first quarter life insurance commission in each quarter (normalized to 0). The solid lines are estimated using a linear regression based on individual-level data using triangular weights. The dashed lines denote the 95% confidence interval based on the heteroscedasticity-consistent standard errors.

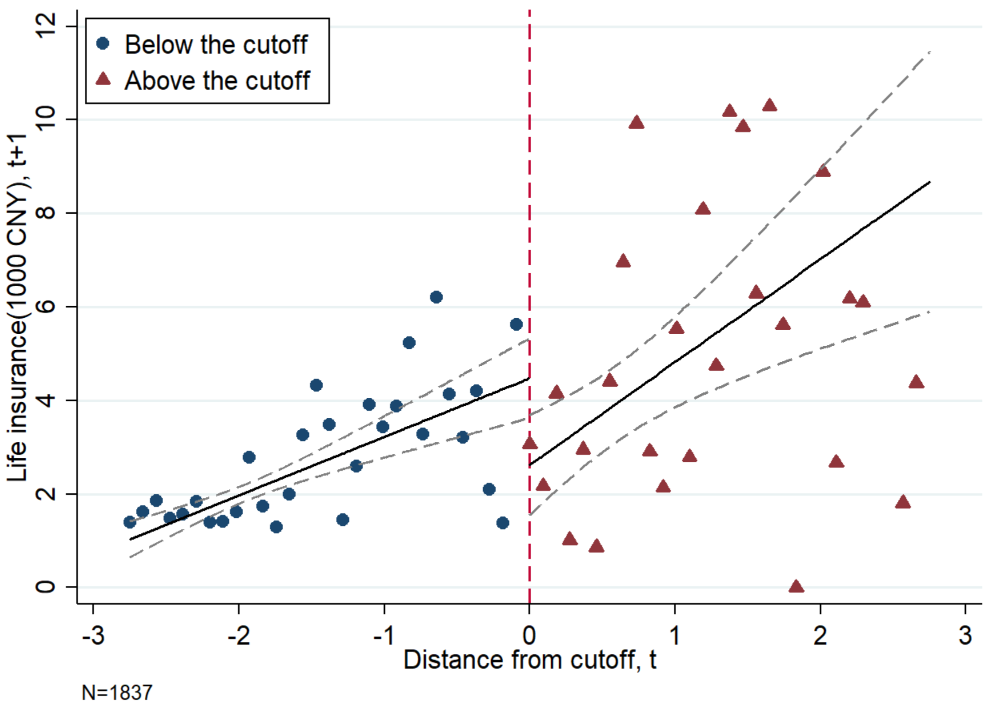
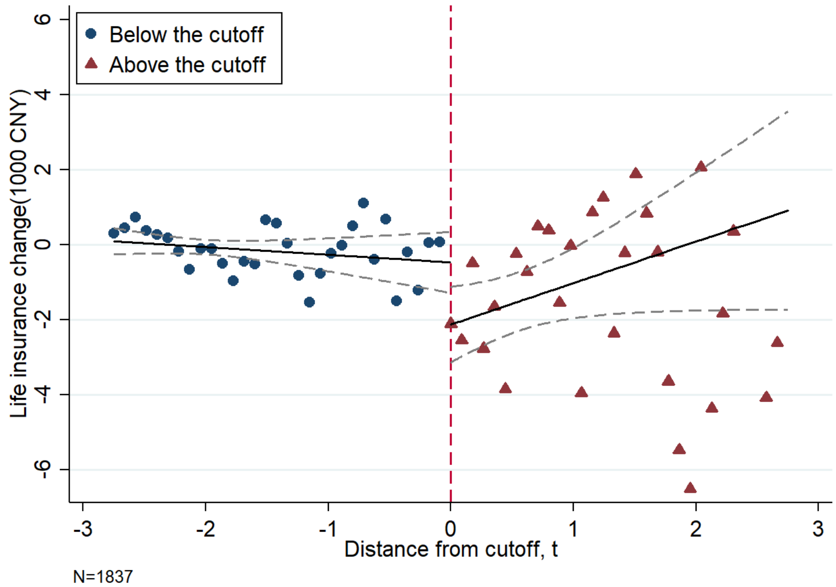
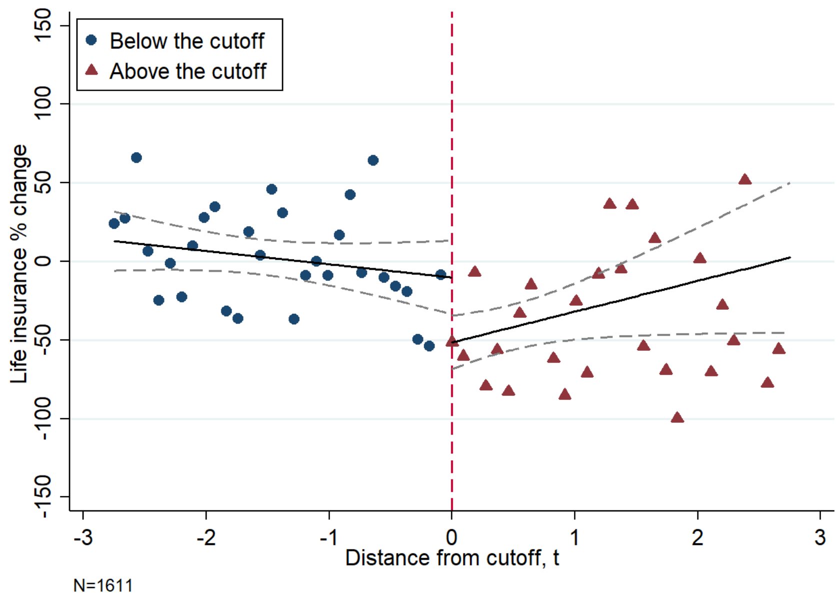


Figure 2.2: Main result - Life insurance commission in quarter t+1

Notes: Each observation is the average life insurance commission in the quarter after an award designation among rookies (main RD sample) in a 0.09 bin based on their standardized first quarter life insurance commission. Dashed vertical line denotes the 10th standardized first quarter life insurance commission in the previous quarter (normalized to 0). The solid lines are estimated using a linear regression based on individual-level data using triangular weights. The dashed lines denote the 95% confidence interval based on the heteroscedasticity-consistent standard errors.



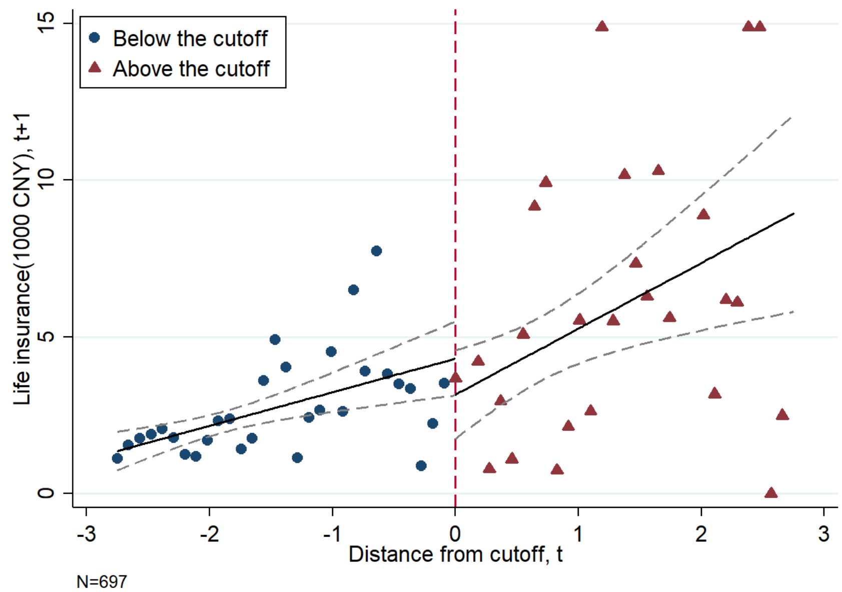
A: Level change



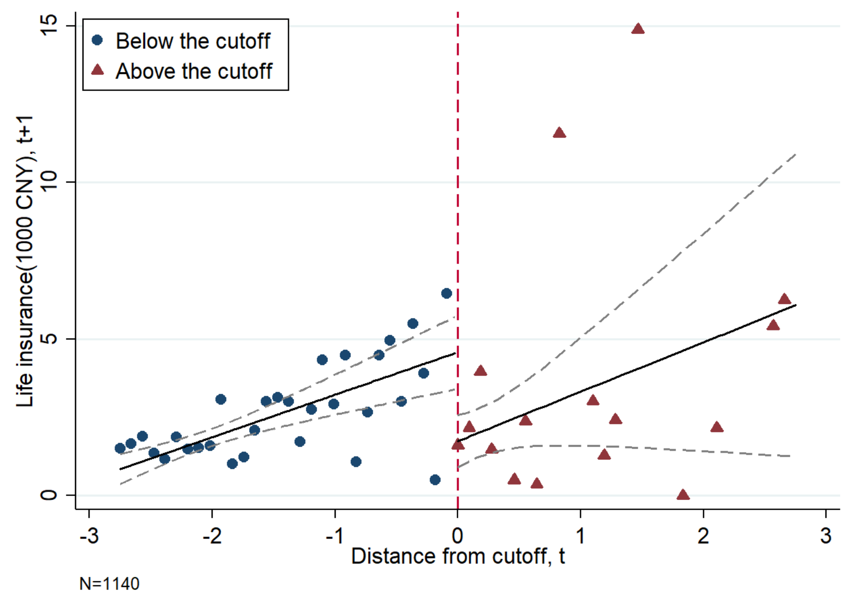
B: Percentage change

Figure 2.3: Interpretation - Change in life insurance commission from quarter t to t+1

Notes: Each observation is the average level change (panel A) or percentage change (panel B) in life insurance commission from the first to the second quarter of rookies (main RD sample) in a 0.09 bin based on their standardized first quarter life insurance commission. Dashed vertical lines denote the 10th standardized first quarter life insurance commission in the previous quarter (normalized to 0). The solid lines are estimated using a linear regression based on individual-level data using triangular weights. The dashed lines denote the 95% confidence interval based on the heteroscedasticity-consistent standard errors.



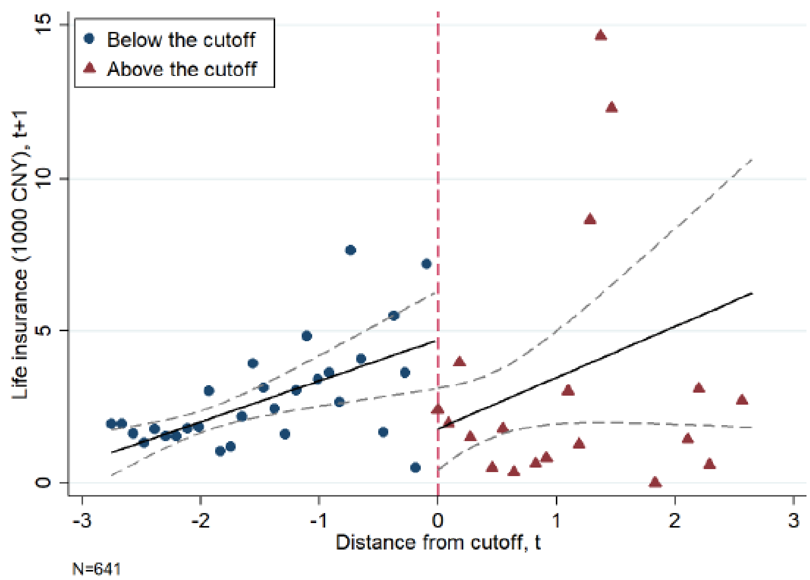
A: Rank=1-3



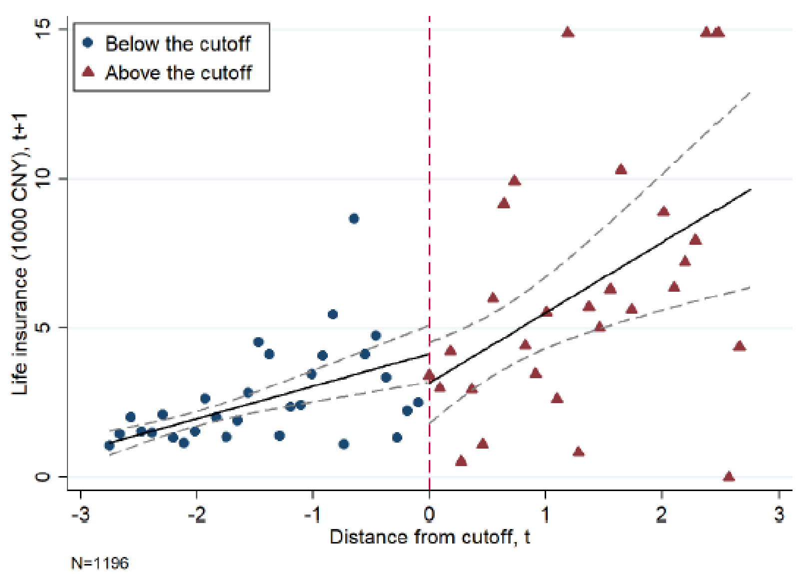
B: Rank=4+

Figure 2.4: Peer sabotage - Impact of award by within-team rank in quarter t

Notes: Each observation is the average life insurance commission in the quarter following an award designation of rookies (main RD sample) who rank 1st to 3rd (panel A) and 4th and worse (panel B) within a team in their first quarter in a 0.09 bin based on their standardized first quarter life insurance commission. Dashed vertical lines denote the 10th standardized first quarter life insurance commission in each quarter (normalized to 0). The solid lines are estimated using a linear regression based on individual-level data using triangular weights.The dashed lines denote the 95% confidence interval based on the heteroscedasticity-consistent standard errors.



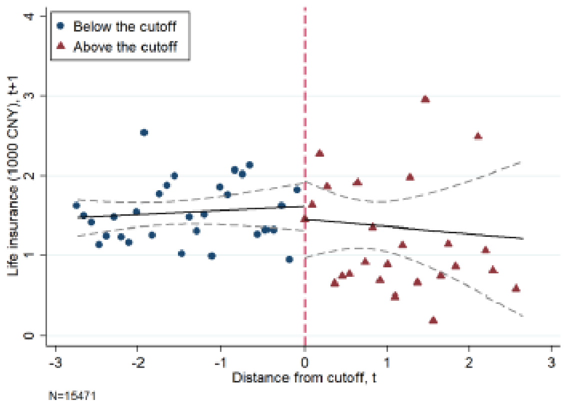
A: With Competitors in teams



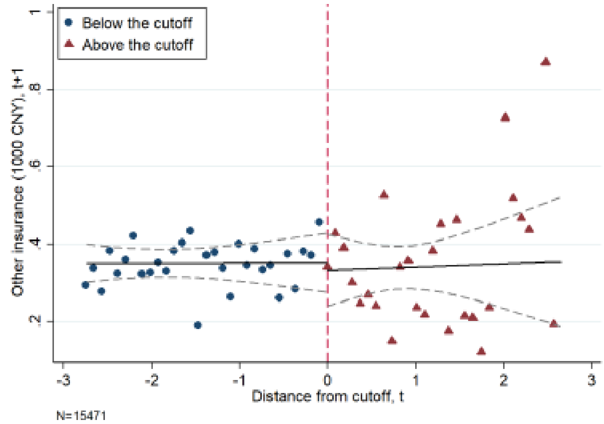
B: Without Competitors in teams

Figure 2.5: Peer sabotage - Impact of award by level of competition in quarter t+1

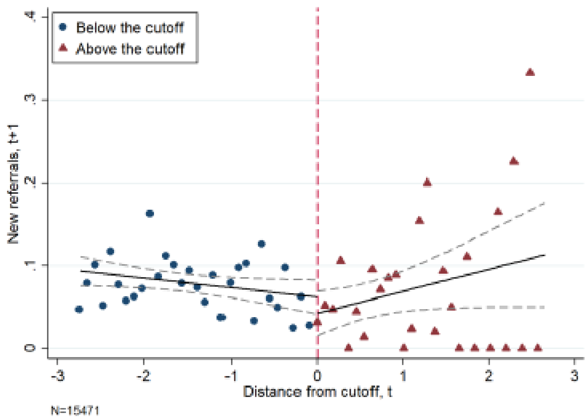
Notes: Each observation is the average life insurance commission in the quarter after an award designation among rookies (main RD sample) who have at least one competitor (panel A) and no competitors (panel B) in their teams in a 0.09 bin based on their standardized first quarter life insurance commission. "Competitors" are qualified senior teammates who are at job level 3 and have the same number of referrals as the competing rookie in the beginning of quarter t+1 (either zero or one referral). Dashed vertical lines denote the 10th rank standardized first quarter life insurance commission in each quarter (normalized to 0). The solid lines are estimated using a linear regression based on individual-level data using triangular weights. The dashed lines denote the 95% confidence interval based on the heteroscedasticity-consistent standard errors.



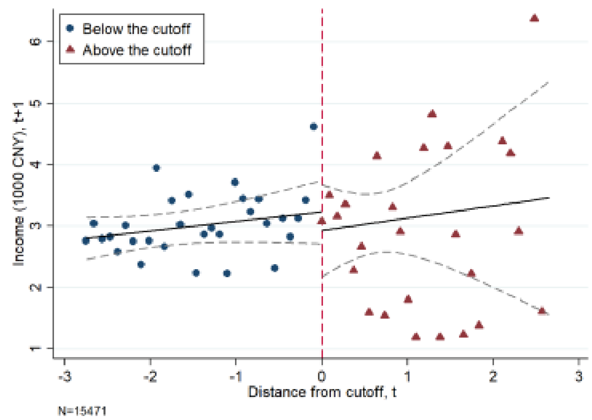
A: Life insurance commission



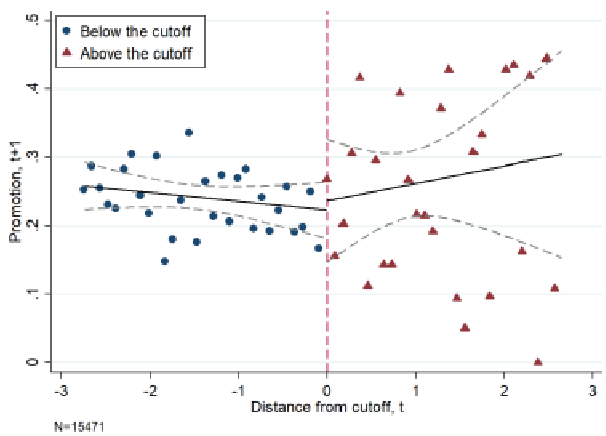
B: Other insurance commission



C: The number of new referrals



D: Income



E: Promotion

Figure 2.6: Spillover effects - Teammates' performance in quarter t+1

Notes: Each observation is the average performance in the quarter following an award designation of participants' non-rookie *teammates* in a 0.09 bin based on *participants'* standardized first quarter life insurance. Dashed vertical line denotes the 10th standardized first quarter life insurance commission in each quarter (normalized to 0). The solid lines are estimated using a linear regression based on individual-level data using triangular weights.The dashed lines denote the 95% confidence interval based on the standard errors clustered by team.

Table 2.1: Summary statistics

Variable	N	Mean	Std. Dev.	Minimum	Maximum
Panel A: Main RD Sample					
Life insurance commission (t+1)	1,837	2.11	3.38	0.00	14.88
Other insurance commission (t+1)	1,837	0.27	0.46	0.00	2.52
Number of referrals (t+1)	1,837	0.19	0.65	0.00	6.00
Income (t+1)	1,837	4.07	4.60	0.00	20.20
Exit (t+1)	1,837	0.06	0.21	0.00	1.00
Life insurance commission (t)	1,837	2.47	2.37	0.00	8.91
Other insurance commission (t)	1,837	0.20	0.30	0.00	1.85
Number of referrals (t)	1,837	0.07	0.36	0.00	6.00
Income (t)	1,837	3.87	3.1	0.00	12.06
Male	1,837	0.33	0.47	0.00	1.00
Age	1,837	34.72	7.91	18.00	54.00
Education	1,837	13.78	1.40	9.00	19.00
Urban	1,837	0.49	0.50	0.00	1.00
Duration (t)	1,837	32.22	18.44	1.00	64.00
Panel B: Peer Sample					
Life insurance commission (t+1)	15,471	1.48	4.27	0.00	39.61
Other insurance commission (t+1)	15,471	0.35	0.64	0.00	4.68
Number of referrals (t+1)	15,471	0.08	0.41	0.00	14.00
Income (t+1)	15,471	2.96	6.21	0.00	64.78
Exit (t+1)	15,471	0.10	0.30	0.00	1.00
Promotion (t+1)	15,471	0.25	0.43	0.00	1.00
Life insurance commission (t)	15,471	1.49	3.89	0.00	36.47
Other insurance commission (t)	15,471	0.35	0.58	0.00	3.33
Number of referrals (t)	15,471	0.10	0.56	0.00	29.00
Income (t)	15,471	2.98	4.70	0.00	27.58
Promotion (t)	15,471	0.20	0.40	0.00	1.00
Male	15,471	0.33	0.47	0.00	1.00
Age	15,471	39.70	9.71	18.00	73.00
Education	15,471	13.76	1.59	9.00	19.00
Urban	15,471	0.49	0.50	0.00	1.00
Job level (t)	15,471	2.09	0.77	1.00	3.00
Tenure (t)	15,471	17.66	18.28	2.00	80.00

Notes: Main RD sample consists of rookies whose standardized first quarter life insurance commission is within 2.75 to the award threshold each quarter. Peer sample consists of non-rookie teammates of the rookies in the main RD sample. t refers to a rookie's first quarter in the company. Monetary variables are in the unit of 1,000 CNY and are winsorized at 1% level in the full sample without bandwidth restriction. *Number of referrals* is the number of new recruits referred by a salesperson. $Promotion(t+1)$ is dummy equalling 1 if one gets promoted by the end of quarter t+1; 0 otherwise. $Exit(t+1)$ is dummy equalling 1 if a rookie has exited the firm by the end of quarter t+1 and is not observed in quarter t+2; 0 otherwise. $Male$ is an indicator of being male. Age is the age in years. $Education$ is the years of education received. $Urban$ is an indicator of coming from urban areas. *Duration* is the number of working days in a quarter. *Job level* ranges from 1 (lowest) to 3 (highest).

Table 2.2: Validity of RD - Baseline characteristics

VARIABLES	(1) Male	(2) Age	(3) Education	(4) Urban	(5) Duration(t)	(6) Exit(t)
Win	0.103	-1.745	-0.282	-0.038	-3.099	0.003
	(0.088)	(1.311)	(0.218)	(0.088)	(2.889)	(0.011)
Observations	1,837	1,837	1,837	1,837	1,837	1,851
R-squared	0.026	0.066	0.132	0.029	0.155	0.011
Top 20 baseline mean	0.349	36.572	13.937	0.532	40.092	0.007
Year×Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	2.75	2.75	2.75	2.75	2.75	2.75

Notes: All coefficients are estimated using a linear RD model and triangular weights using rookies whose standardized life insurance commission is within 2.75 to the award threshold in each quarter. Exit(t) is a dummy equalling 1 if a rookie has exited by the end of quarter t and is not observed in quarter t+1; 0 otherwise. Definitions for other variables are described in the notes to Table 1. Top 20 baseline mean refers the mean of outcome variables among the top 20 rookies in their first quarter in the company. Heteroscedasticity-consistent standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.3: Main result - Life insurance commission in quarter t+1

VARIABLES	(1) Life insurance	(2) Life insurance	(3) Life insurance
Win	-1.857***	-1.803***	-1.720***
	(0.696)	(0.646)	(0.655)
Observations	1,837	1,837	1,837
R-squared	0.078	0.214	0.229
Top 20 baseline mean	6.209	6.209	6.209
Year*Quarter FE	No	Yes	Yes
Demographic controls	No	No	Yes
Bandwidth	2.75	2.75	2.75

Notes: The dependent variable is the life insurance commission earned in the quarter following an award designation (measured in 1,000 CNY). All coefficients are estimated using a linear RD model and triangular weights using rookies whose standardized life insurance commission is within 2.75 to the award threshold in each quarter. Column (1) has no control variables, column (2) includes year-by-quarter fixed effects, and column (3) further controls for gender, age, age squared, education, and urban status, which are all described in the notes to Table 1. Heteroscedasticity-consistent standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.4: Signaling mechanism - Cumulative exit rate by quarter

VARIABLES	(1) t+1	(2) t+2	(3) t+3	(4) t+4	(5) t+5	(6) Tenure
Win	0.005	0.032	0.018	0.033	0.053	-0.192
	(0.010)	(0.043)	(0.051)	(0.065)	(0.071)	(0.483)
Observations	1,837	1,837	1,837	1,837	1,837	1,837
R-squared	0.383	0.346	0.323	0.358	0.350	0.313
Top 20 baseline mean	0.059	0.160	0.268	0.383	0.472	7.230

Notes: The dependent variables are dummies equalling 1 if a rookie has exited the company *by* the end of 1st, 2nd, ..., and 5th quarter after an award designation and is not observed in the following quarter, 0 otherwise, in column (1), (2), ..., and (5). Note that the exit rate is cumulative rather than per quarter. Tenure is the total length of stay in the company before a rookie leaves the company or before our sample period ends (measured in quarters). Specifications mirror the one in Table 3 column (3). Heteroscedasticity-consistent standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.5: Effort reallocation mechanism - Other performance in quarter t+1

VARIABLES	(1) Other insurance	(2) Referral	(3) Income
Win	-0.091	-0.138	-2.032**
	(0.061)	(0.128)	(0.854)
Observations	1,837	1,837	1,837
R-squared	0.076	0.144	0.176
Top 20 baseline mean	0.405	0.491	6.693

Notes: The dependent variables in columns (1)-(3) are other insurance commission, the number of new referrals, and income in the quarter following an award designation, respectively. Monetary values are measured in 1,000 CNY. Specifications mirror the one in Table 3 column (3). Heteroscedasticity-consistent standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.6: Effort reallocation mechanism - Future performance

VARIABLES	(1) Highest job level	(2) Ever being a manager
Win	-0.391*	-0.026
	(0.216)	(0.063)
Observations	1,837	1,837
R-squared	0.148	0.070
Top 20 baseline mean	3.424	0.149

Notes: The dependent variables in columns (1) and (2) are the highest job level reached before a rookie leaves the company or before our sample period ends and the probability of a rookie ever being promoted to manager (i.e., level 4-6), respectively. Specifications mirror the one in Table 3 column (3). Heteroscedasticity-consistent standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.7: Gaming mechanism - Cancelled life insurance commission by quarter

VARIABLES	(1) t	(2) t+1	(3) t+2	(4) t+3
Win	0.253	-0.225*	-0.033	-0.198
	(0.180)	(0.132)	(0.156)	(0.201)
Observations	1,837	1,837	1,716	1,526
R-squared	0.231	0.111	0.070	0.090
Top 20 mean	0.632	0.387	0.282	0.282

Notes: The dependent variables in columns (1)-(4) are the life insurance commission *cancelled* in the corresponding quarter. "Top 20 mean" is the outcome mean among top 20 rookies (based on first quarter ranking). The number of observations decreases due to the exit of salespeople from the company. Specifications mirror the one in Table 3 column (3). Heteroscedasticity-consistent standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.8: Strategic reallocation across salespeople - Teammates' life insurance commission by quarter

VARIABLES	(1) t-1	(2) t	(3) t+1	(4) t+2	(5) t+3
Panel A: Referrers					
Win	-1.146	0.872	-0.000	-2.306	-1.556
	(2.250)	(2.597)	(2.278)	(2.088)	(2.359)
Observations	1,513	1,831	1,833	1,710	1,514
R-squared	0.282	0.247	0.283	0.183	0.200
Top 20 mean	12.377	13.067	14.16	12.109	11.101
Panel B: Managers					
Win	-0.090	0.682	0.791	1.479	1.554
	(3.704)	(3.039)	(2.620)	(5.424)	(3.210)
Observations	1,480	1,729	1,729	1,611	1,425
R-squared	0.165	0.122	0.163	0.144	0.096
Top 20 mean	14.860	13.896	16.799	15.199	13.323
Panel C: Senior teammates					
Win	0.090	0.331	-0.121	-0.298	-0.640
	(0.393)	(0.446)	(0.451)	(0.499)	(0.463)
Observations	3,645	4,306	4,306	4,051	3,803
R-squared	0.071	0.088	0.079	0.079	0.078
Top 20 mean	3.359	3.010	2.905	3.129	3.193

Notes: The dependent variables in columns (1)-(5) are the life insurance commission in the second quarter before the award (t-1) to the third quarter after the award (t+3) for rookies' referrers (panel A), managers (panel B), and senior teammates (panel C). Senior teammates are defined as those whose job level at t=0 were equal to 3. "Top 20 mean" is the outcome mean among the referrers, managers, or senior teammates of top 20 rookies (based on first quarter ranking). The number of observations decreases due to the exit of rookies, managers, referrers, or senior teammates from the company. Specifications mirror the one in Table 3 column (3) and also control for the job level of the referrers, managers, or senior teammates and an indicator of whether the referrer is the same as the manager. Heteroscedasticity-consistent standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.9: Peer sabotage - Impact of award by within-team rank in quarter t

	(1)	(2)	(3)
		Panel A	
VARIABLES	Rank=4+	Rank=1-3	Prob.>*chi2*
Win	-2.970***	-0.359	0.027
	(0.866)	(0.828)	
Observations	1,140	697	
R-squared	0.214	0.283	
No. of winners	30	85	
Top 20 baseline mean	5.583	6.579	
		Panel B	
VARIABLES	Rank=3+	Rank=1-2	Prob.>*chi2*
Win	-2.486***	0.357	0.022
	(0.742)	(1.022)	
Observations	1,333	504	
R-squared	0.244	0.284	
No. of winners	44	71	
Top 20 baseline mean	5.683	6.722	
		Panel C	
VARIABLES	Rank=2+	Rank=1	Prob.>*chi2*
Win	-2.124***	0.144	0.119
	(0.808)	(1.275)	
Observations	1,595	242	
R-squared	0.234	0.260	
No. of winners	64	51	
Top 20 baseline mean	5.953	6.752	

Notes: This table splits the main RD sample by a rookie's rank of average monthly life insurance commission among her teammates' in the quarter before an award designation. Panels A-C split the sample at within-team rank 4th, 3rd, and 2nd, respectively. The dependent variable is the life insurance commission in the quarter following the award designation. "No. of winners" refers to the number of award winners in each subsample. Specifications in columns (1) and (2) mirror the one in Table 3 column (3). Column (3) reports the p-value for Chow test on null hypothesis that the coefficients in columns (1) and (2) are statistically indistinguishable. Heteroscedasticity-consistent standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.10: Peer sabotage - Impact of award by level of competition in quarter t+1

VARIABLES	(1) With Competitors	(2) Without Competitors	(3) Prob.>*chi2*
Win	-3.621***	-0.201	0.007
	(1.008)	(0.782)	
Observations	641	1,196	
R-squared	0.264	0.253	
No. of winners	40	75	
Top 20 baseline mean	7.197	7.348	

Notes: This table splits the main RD sample by whether rookies have at least one competitor who competes with them for internal resources in the quarter after an award designation. "Competitors" are qualified senior teammates who are at job level 3 and have the same number of referrals as the competing rookie in the beginning of quarter t+1 (either zero or one referral). The dependent variable is the life insurance commission in the quarter following the award designation. "No. of winners" refers to the number of award winners in each subsample. Specifications in columns (1) and (2) mirror the one in Table 3 column (3). Column (3) reports the p-value for Chow test on null hypothesis that the coefficients in columns (1) and (2) are statistically indistinguishable. Heteroscedasticity-consistent standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 2.11: Spillover effects - Teammates' performance in quarter t+1

VARIABLES	(1) Life insurance	(2) Other insurance	(3) Referral	(4) Income	(5) Promotion
Win	-0.063	0.014	-0.027	-0.262	-0.010
	(0.229)	(0.040)	(0.018)	(0.381)	(0.034)
Observations	15,471	15,471	15,471	15,471	15,471
R-squared	0.233	0.343	0.040	0.269	0.108
Baseline mean	1.693	0.374	0.136	3.283	0.199
Year*Quarter FE	Yes	Yes	Yes	Yes	Yes
Demographic controls	Yes	Yes	Yes	Yes	Yes
Bandwidth	2.75	2.75	2.75	2.75	2.75

Notes: The dependent variables in columns (1)-(5) are life insurance commission, short-term insurance commission, the number of new referrals, total income, and likelihood of being promoted to higher job level ("promotion") in the quarter following an award designation, respectively. Baseline mean refers to the outcome mean of non-rookie teammates of the top 20 participants in each quarter. In all specifications, we control for year-quarter fixed effects, gender, age, age squared, education, job level, and urban status, which are all described in the notes to Table 1. Standard errors reported in parentheses are clustered by team. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Chapter 3

When Weed is Legalized Next Door: How Colorado's Recreational Marijuana Legalization Affects Neighboring States

Abstract

I examine the effect of Colorado's recreational marijuana legalization (RML) on the illegal marijuana use and the burden of police to enforce marijuana laws in its neighboring states. I use a difference-in-differences (DID) design with distance to Colorado border as treatment intensity. I find that Colorado's RML increased marijuana possession offenses and arrests among adult males in police agencies closer to the Colorado border relative to those farther away. I further provide evidence that marijuana possession offenses shifted to locations near highways and roads. The amount of marijuana seized in these locations also increased, whereas that seized in other locations did not. The findings add to the heated policy debate over the pros and cons of RMLs, and alert the states considering RML to take the spillover effect into account when calculating the costs of the policy.

3.1 Introduction

On November 8, 2016, five states in the United States voted on whether to legalize the recreational use of marijuana. Four of these five states passed RMLs, resulting in a total of eight states including the District of Colombia where those aged 21 and over can legally buy recreational marijuana[1].

RML states may gain benefits from legalizing marijuana, such as an increase in tax revenue from legal sales of marijuana and savings of police resources on controlling marijuana-related crimes (Adda, McConnell & Rasul, 2014; Miron, 2010). In contrast, the spillover effects from nearby RML states concern states that prohibit marijuana, particularly with their residents' illegal marijuana use and the burden on their local law enforcement. Because there is no residency requirement to buy marijuana in RML states, out-of-state residents can also buy marijuana legally. Anecdotal evidence suggests a booming industry of marijuana tours to Colorado for buying marijuana (Feuer, 2016). These consumers may bring marijuana back to their home states where marijuana possession is still illegal. In fact, Nebraska and Oklahoma filed a federal lawsuit against Colorado in 2014, claiming that RML in Colorado has increased their costs of enforcing marijuana laws and detracted their efforts and expenditure away from tackling more serious crimes[2]. Given the current "Green Rush" of RML in the United States and RML's potential negative effects on other states, understanding how RML in one state affects the illegal marijuana use and the burden on police department in its neighboring states is timely and necessary. This paper aims to answer the question using RML in Colorado as a case study.

The spillover mechanism is as follows. First, learning that one's neighboring state already passed or is likely to pass RML may change one's attitude towards and perceived risks

[1]Colorado and Washington passed RMLs in November 2012 which became effective in December 2012. Alaska, Oregon, and the District of Columbia passed RMLs in November 2014 that became effective in February 2015 (for Alaska and the District of Columbia) and July 2015 (for Oregon). California, Maine, Massachusetts, and Nevada passed RMLs in November 2016 that became effective in December 2016 (for Massachusetts) and January 2017 (for Maine and Nevada), and that will become effective in January 2018 (for California).

[2]The Supreme Court eventually declined to hear this case (https://www.supremecourt.gov/opinions/15pdf/144orig_6479.pdf).

of consuming marijuana.[3] Residents in non-RML states may thus increase their demand for marijuana and buy it either from local suppliers or suppliers in nearby RML states. According to Latané (1981) and Latané et al. (1995), the social impact of a source on a receiver decreased with increasing physical distance. Therefore, the change in perception and the increase in demand may be larger in regions closer to RML states. One thing to note is that this channel may take effect shortly before the passage of RML.[4] Second, RML may change the quality and the price of (legal or illegal) marijuana, though the direction of change is theoretically ambiguous.[5] Anderson, Hansen, and Rees (2013) collected price information from High Times from 1990 to 2011, and documented that MMLs gradually led to lower price of high-grade marijuana in MML states. This channel can take effect even before recreational marijuana stores are opened in RML states. While out-of-state residents cannot buy legal marijuana before the stores are opened, they may access cheaper illegal marijuana on the street in the RML state right after it passed RML.[6] Third, after recreational marijuana stores open, purchasing marijuana in RML states is safer and easier for out-of-state people, because most RML states do not require proof of residency to purchase. Lured by better price or quality of marijuana and easier and safer access to marijuana, residents in neighboring non-RML states may cross the border to RML states to buy marijuana. They are more likely to do so when the cost of cross-border shopping is lower, e.g. closer in physical distance and shorter in travel time to RML states. Taking the above three channels together, the increase in marijuana use and the increase in law enforcement's costs on enforcing marijuana laws are likely to be higher in non-RML regions closer to RML states.

[3]Khatapoush and Hallfors (2004) find that people in California perceived less harm from smoking marijuana after medical marijuana legalization (MML).

[4]I find suggestive evidence that the percentage of juveniles who perceive no great risk smoking marijuana once per month increased in 2012 among Colorado's neighboring states relative to the United States average. Wall et al. (2011) also documented that states with MMLs had lower adolescent perception of marijuana riskiness compared to states without from 2002 to 2008, and the difference existed around one year before MML passages.

[5]I only have weed price data from December 2013 through July 2015 (retrieved July 17, 2017 from https://github.com/frankbi/price-of-weed and priceofweed.com), and I am not able to empirically examine the change in quality and price in this paper.

[6]According to *Denver Post*, the marijuana black market still thrives in Colorado even after recreational marijuana stores were opened (http://www.denverpost.com/2014/12/19/after-pot-legalization-focusing-on-a-new-kind-of-black-market/); let alone before the stores opened.

Several features make Colorado and its neighboring states a particularly useful setting to test the spillover effect. First, Colorado passed RML (Colorado Constitutional Amendment 64) in December 2012, making it and Washington the first two states to allow the sales and consumption of legal recreational marijuana. Such an early starting date provides a longer post-treatment period to analyze the spillover effect of RML. Second, RML in Colorado allows anyone over 21 to buy up to one ounce of marijuana from licensed dispensaries regardless of residency status; thus, out-of-state residents can also purchase marijuana in Colorado. Third, Colorado has the highest number of neighboring states that have not passed either RML or MML and do not border other RML states. In this study, I consider Utah, Wyoming, Nebraska, Kansas, Oklahoma, and Texas neighboring states of Colorado. Although Texas does not border Colorado, I include it for its proximity.[7] I excluded Arizona and New Mexico because they passed MML in 2011 and 2008, respectively, and including them may contaminate the results.[8] I do not include Wyoming because police agencies in Wyoming do not report to the National Incident-Based Reporting System (NIBRS). While Washington is one of the first two states to pass RML, I do not include its neighboring states in this paper. The reason is that Idaho is Washington's only neighboring state that has not passed MML or RML, but Idaho also borders Oregon. Oregon passed MML in 1998, and the 2010 *State vs. Berringer* case prompted Oregon to clarify that out-of-state residents were allowed to obtain a medical marijuana registration card and buy medical marijuana in Oregon.[9] Including Idaho in the study may confound the estimated effects.

This paper focuses on illegal marijuana use among adult males and the burden on police

[7] I further restrict all agencies to be within 400 miles of Colorado, so Mexico should have little impact on Texas agencies. My results remain similar when I drop Texas from the sample.

[8] MMLs should, in theory, increase both the supply of marijuana and the demand for marijuana, unambiguously leading to an increase in consumption (Pacula et al., 2010). Due to the prohibitive costs of ensuring that only patients can access medical marijuana, diffusion to non-patients is likely to occur. Chu (2014) used illegal marijuana possession arrests and treatment admissions to rehabilitation facilities as proxies for marijuana use among non-patients, and documented that both measures increased after MMLs. Wen et al. (2015) used restricted-access individual-level National Survey on Drug Use and Health (NSDUH) data and found that MMLs increased the probability of daily marijuana use, marijuana abuse, and marijuana dependence among adults aged 21 and above.

[9] Refer to http://www.doj.state.or.us/wp-content/uploads/2017/06/op2010-2.pdf for details. The Oregon Health Authority stopped issuing cards to patients without Oregon addresses in January 2016.

agencies to enforce marijuana laws in Colorado's non-RML and non-MML neighboring states. More specifically, I use agency-level marijuana possession offenses from NIBRS for years 2009 to 2015 to proxy illegal marijuana use, and supplement them with agency-level marijuana possession arrests from Uniform Crime Reports (UCR) to better compare with existing literature.[10] Offenses and arrests do not measure marijuana use directly, as they represent frequencies rather than individuals, and they are combination of responses from both drug users and police officers. But conceptually offense and arrest data can capture changes not only at the extensive but also at the intensive margin (Chu, 2014). Also, offense and arrest data represent objective measures, and do not suffer from the self-reporting bias common in survey data.[11]

In addition, I construct agency-level "stand-alone" marijuana possession arrests, i.e. no other drug-irrelevant arrests are reported in the same incident, using NIBRS data for years 2009 to 2015 to better proxy the burden on police agencies to enforce marijuana laws. Drug possession arrests sometimes occur as a byproduct of regular search during other arrests (Miron, 2010). If the drug arrests are byproducts, the added costs from handling such arrests apart from the other arrests will be small. Therefore, stand-alone arrests can better proxy the burden on police agencies. This exercise is only feasible using NIBRS data, because it records all arrests associated with one crime incident up to ten. Arrests in themselves do not equal police agencies' total costs in tackling illegal marijuana, but they are a direct and important factor in calculating the total costs (Miron, 2010). In addition to stand-alone marijuana possession arrests, I also report results using stand-alone marijuana sale and manufacture arrests as supplement.[12]

In this paper, I adopt a difference-in-differences (DID) research design with distance from a police agency in the neighboring states to the Colorado border as treatment intensity of

[10]For example, Hao and Cowen (2017) used marijuana possession data from UCR. Estimates using offense data can differ from those using arrest data because marijuana possession is a minor offense and may not lead to an arrest.

[11]Miller and Kuhns (2012) documented that people might respond more honestly about marijuana use in surveys after MMLs.

[12]Ideally, I should multiply the proportion of arrests due to marijuana possession, and marijuana sale and manufacture with total police expenditure as in Miron (2010), but the Justice Employment and Expenditure (JEE) data series stopped in 2012.

RML on the area covered by the agency. My main specification controls for agency and state by year fixed effects. I identify the effects of RML from the change in the difference of marijuana possessions (or marijuana sale and manufacture) between nearby and far away agencies after subtracting the common annual shock on marijuana in each state and the time-invariant property of the police agency itself. Drawing inference from marijuana possession offenses and arrests, I find that Colorado's RML increased illegal marijuana usage as well as the burden on police agencies to enforce marijuana laws in neighboring states. NIBRS data show that Colorado's RML increased marijuana possession offense rate in agencies closer to Colorado by 72 per 100,000 (100k) residents among adult males than farther-away agencies, or about 32 percent of baseline mean. NIBRS stand-alone marijuana possession arrest also increased around 30 percent of baseline mean after Colorado's RML. However, marijuana sale and manufacture offenses and arrests did not seem to respond to the RML, which suggests that Colorado's RML mainly affected neighboring states through inducing resident's demand rather than increasing local drug dealers' supply. Moreover, marijuana possession estimated with NIBRS data shifted to locations like highway and street, and that the amount of possessed marijuana increased in those locations for closer agencies relative to father-away ones. I also examine the effect of RML on marijuana use by racial and by age groups. The effect mainly concentrates in white adult males and the effect decreases as age rises.

I examine the validity of my DID design using an event study, where I allow spillover effect of RML to vary from year to year. The marijuana possession in closer police agencies did not increase relative to farther away agencies before 2012 (Colorado passed RML in December 2012). The strong increase only began in 2013 and peaked in 2015 (recreational marijuana stores started operating in January 2014). As a robustness check, I include the number of medical marijuana patients in Colorado to control for the potential ramping-up effect of Colorado's MML, but results change very little after the inclusion.

Findings in the paper add to the heated policy debate on the pros and cons of RMLs by showing evidence for increased illegal marijuana use among residents and increased burden on police enforcement in states bordering Colorado. Given these findings, states should prepare for negative spillover from their RML neighbors. This paper also urges the states that are considering RML to take the potential spillover effects into account when conducting cost-and-benefit analysis.

The paper proceeds as follows: Section 2 presents a brief literature review. Section 3 describes the main data sets and sample construction. Section 4 and 5 discuss the empirical strategy and main results. Section 6 explains RML's potential impact on law enforcement and presents corresponding robustness checks. Section 7 concludes.

3.2 Literature

This paper contributes directly to the literature studying the effect of marijuana legalizations on marijuana use. Many of these studies used MMLs rather than RMLs as the policy shock.[13] They generally found that MMLs increased illegal marijuana use in the MML states. With transaction-level information from marijuana purchases made by arrestees, Pacula et al. (2010) found evidence supporting the conclusion that a reduction in sanctions on marijuana use, like MMLs, would increase use of marijuana. Chu (2014) used illegal marijuana possession arrests and treatment admissions to rehabilitation facilities as proxy for marijuana use among non-patients, and documented that both measures increased after MMLs. Wen et al. (2015) utilized restricted-access individual-level NSDUH data and found that MMLs increased the probability of daily marijuana use, marijuana abuse, and marijuana dependence among adults aged 21 and above.[14]

[13]Studies likely used MMLs because more states have passed MMLs and longer post-periods are available for MMLs. As of October 2017, 29 states plus the District of Columbia have MMLs, while only seven states plus the District of Columbia have RMLs; the earliest MML was passed in 1996 while the earliest RML was passed in 2012.

[14]There are an even larger literature examining the public health impact of MMLs, though results are mixed. Model (1993) showed that marijuana decriminalizations were associated with fewer emergency room episodes involving drugs other than marijuana. Bachhuber et al. (2014) found that MMLs lowered stat opioid overdose mortality rate as well as heroin treatments and cocaine/heroin arrests. Anderson et al. (2013) found that MMLs led to a reduction in drunk driving fatalities. However, Kelly and Rasul (2014) showed that a policing experiment that

The majority of the papers in this literature employed a state-level DID, and considered all MML states as one homogeneous treatment group and all other states (or a weighted average of all other states) as the control group.[15] One common assumption is that the passage (or the timing of passage) of MMLs is exogenous. Even though many MMLs were passed by lawmakers rather than by a general vote by the electorate, the passages may still reflect the will of the general public. By focusing on the effect of Colorado's RML on other states, this paper bypasses this exogeneity assumption and exploits the relative exposure to Colorado's RML between closer and farther away police agencies in the neighboring states to identify the effect of RML. Another assumption made by these papers is that states in the control group are not treated. Violation of this assumption will lead to underestimate of the true effect. In this paper, I will empirically examine whether the no-spillover-effect assumption holds.

This study also relates to papers quantifying the financial gains and losses associated with marijuana legalizations (Caulkins, 2010; Gieringer, 2009; Miron, 2010). These papers estimated government's expenditure on enforcement of marijuana prohibition in the legalizing states from three aspects: police resources from elimination of drug arrests, prosecutorial and judicial resources from elimination of drug prosecutions, and correctional resources from elimination of drug incarcerations. My findings show that increased marijuana possession arrests in neighboring states can be another source of costs to include in the estimates.

In a more general sense, this paper adds to a large literature on how one region's policies affect other regions in various contexts. For example, tax changes in one locality will induce consumers' cross-border shopping for cigarettes (Goolsbee et al., 2010; Lovenheim, 2008; Merriman, 2010) and alcohol (Stehr, 2007), and the degree of the effect depends on the distance from the consumers to the alternative shopping locations. In addition, Lovenheim and Slemrod (2010) found that an increase in a state's minimum legal drinking age increased fatal accidents

de-penalized the possession of small quantities of cannabis in London raised hospital admissions related to hard drugs among men.

[15] Some exceptions exist. Pacula et al. (2015) differentiated states with different MML policy frameworks and found some evidence that the differences in the details of MMLs could imply different legalization effects.

among 18 and 19-year-olds living close to regions with a lower legal drinking age. Bharadwaj (2015) also showed that the 1957 amendment to the Mississippi marriage law, which raised minimum marriage age, reduced the marriage rate and increased school enrollment in neighboring counties.

Given the brief history of RMLs in the United States, very few papers examine the spillover effect of RMLs. Two recent papers touched on this topic. Ellison and Spohn (2015) found that Nebraska border counties experienced significant growth in marijuana-related arrests and jail admissions after Colorado's MML. However, they did not have a control group and that they used data from 2000 to 2004 and 2010 to 2013 but not data from 2005 to 2009. The paper most closely related to mine is the one by Hao and Cowan (2017).[16] The authors examined the spillover effect of RML in both Colorado and Washington on their neighboring states respectively using arrest data from UCR for years 2009 through 2014. They found that RML led to an increase in marijuana possession arrests in border counties relative to non-border counties in the neighboring states. Our papers differ from each other in several aspects. First, I use agency-level data with agency fixed effects and state by year fixed effects, while they use county-level data and control for county fixed effects and year fixed effects. They may not be able to control for time-variant state-specific changes regarding marijuana. Second, I compare arrest and offense data from NIBRS and arrest data from UCR, while they used arrest data from UCR.[17] Third, I exclude New Mexico because New Mexico's MML law came into effect in 2008, and the MML rules were revised in 2010, whereas they included it; I include Texas, while they did not. Fourth, I examine the change in marijuana possession by age, by race, and by location types, whereas they differentiated between adults and juveniles. Finally, I extend the post-treatment period to 2015, during which I found the largest effect .

[16]I wrote my paper simultaneously with theirs. We only discovered each other's paper at an advanced stage of writing up the manuscript.

[17]Minor offenses such as marijuana possession may not result in arrests. Also, with incident-level NIBRS data, I can examine the effect of RML at a detailed level, such as the location of the marijuana possession and the amount of possessed marijuana.

3.3 Data and Sample Construction

3.3.1 Data Sources

I collect data on crimes from two sources. The first dataset is the agency-level UCR yearly summarized data on arrests for years 2009 through 2015. The UCR arrest data reports annual arrest counts by age-sex and by race subgroups for each UCR offense code reported by each UCR reporting police agency. The second dataset is the incident-level NIBRS offense and arrest data for years 2009 through 2015. While UCR assigns a specific code ("187") for illegal marijuana possession and a code ("182") for illegal marijuana sale and manufacture, NIBRS does not. Therefore, I define illegal marijuana possession as offense = "35A" (drug/narcotic violation), criminal activity = "P" (possessing/concealing), property type = "6" (seized by police), and suspected drug type = "E" (marijuana); I define illegal marijuana sale and manufacture as offense = "35A", criminal activity = "C" (cultivating/manufacturing/publishing) or "D"(distributing/selling), property type = "6" , and suspected drug type = "E". Since drug crimes, especially drug possessions, sometimes are byproducts of more serious non-drug crimes, and counting such arrests as added costs for police to enforce marijuana laws will overstate the true costs (Miron, 2010), I construct stand-alone marijuana possession and marijuana sale and manufacture by restricting incidents to have only drug-related crimes (offense code starting with "35"). I then aggregate the incident-level data to annual counts for each NIBRS reporting agency by adding up the number of offenders and arrestees in each category in each year, respectively.[18]

Compared to UCR arrest data, NIBRS offense data has two advantages. First, NIBRS reports crime incidents at a much more detailed level. I can examine the crimes from more angles, such as by location types and the amount of possessed marijuana. I can also calculate stand-alone marijuana possession and marijuana sale and manufacture arrests, because NIBRS reports all

[18] I did not include any offenses with missing information in any of the above categories in the final counts. I recognize that this algorithm may understate actual illegal marijuana possessions because of the missing information. Even though this exclusion results in an underestimation, the estimates using NIBRS are still much larger in both level and in log than those using UCR data.

crime types up to ten associated with one incident. Second, NIBRS reports all offenders in a crime incident irrespective of whether an arrest has been made. Because marijuana possession is a relatively minor offense and not all offenses will become arrests, UCR arrest data may understate the true level of marijuana possessions. However, NIBRS has more selection into reporting issues than UCR. The population covered by reporting agencies in UCR represents more than 97.7 percent of the total United States population in 2014, while the population covered by reporting agencies in NIBRS covers only 30.3 percent (FBI CIUS, 2014; FBI UCR, 2014).

Following the convention of criminology literature, I focus on crimes committed by adult males only. I consider NIBRS adult male marijuana possession offenses per 100k residents covered by a police agency as my main proxy of illegal marijuana use, and supplement this measure with UCR marijuana possession arrests per 100k residents to better compare with existing papers.[19] For proxy of the burden on police agencies in enforcing marijuana law, I mainly use NIBRS stand-alone adult male marijuana possession arrests per 100k residents, and I supplement it with adult male marijuana sale and manufacture arrests per 100k residents.

I collect county-level unemployment rate from the Bureau of Labor Statistics and county-level age and racial composition data from the United States Census Bureau population estimates.

3.3.2 Sample Construction

I define the neighboring states of Colorado in this study as Utah, Nebraska, Kansas, Oklahoma, and Texas. I do not include Arizona and New Mexico so I minimize the confounding effect from their MMLs passed in 2011 and 2008, respectively.[20] Wyoming is not in the sample because agencies in Wyoming do not report to NIBRS. I include Texas for its proximity to Colorado, but my results change very little when I drop Texas.

I first require police agencies to be within 400 miles of Colorado.[21] I do not include police

[19]For example, Hao and Cowen (2017) used UCR marijuana possession arrests.
[20]MML could increase marijuana use (Chu, 2014; Pacula et al., 2010; Wen et al., 2015).
[21]I calculated the distance using the Haversine formula from the Federal Information Processing Standard (FIPS)

agencies that are farther away, because RML is not very likely to affect these agencies, since they are at least five hours's drive from Colorado (assuming a driving speed of 75 miles per hour). Furthermore, using an overly large sample selection distance will include Texas agencies that are too close to the Mexican border, where a lot of marijuana smuggling occurs.

Because participation in the UCR and NIBRS program is largely voluntary, agencies sometimes do not report every month or every year, and they may not report data for all offense categories. While distinguishing a true zero from missing data is difficult, the Federal Bureau of Investigation communicates with large city agencies to ensure data quality (Akiyama & Propheter, 2005), and most missing data are from small agencies that do not report for an entire year (Lynch & Jarvis, 2008). Therefore, I focus on city police agencies with larger than 2,500 covered residents that report for more than six months during a year. For an agency satisfying all the above selection criteria, I take their missing offense and arrest categories as true zeros.[22]

I apply the above sample selection criteria to the NIBRS data and then match the selected agency-year observations back to the UCR data. My final NIBRS sample consists of 1,490 agency-year observations from 251 agencies in 148 counties, and the UCR sample consists of 1,468 agency-year observations from 250 agencies in 148 counties (22 agency-year observations do not appear in the UCR data). Figure 3.1 depicts the location of police agencies in the NIBRS sample, which is the main sample of the paper.

3.3.3 Descriptive Statistics

Table 3.1 presents the baseline summary statistics for main variables by distance to the Colorado border. An immediate observation is that adult male marijuana possession offenses (arrests) per 100k residents are generally higher in agencies within 150 miles of Colorado

place code of a police agency in the neighboring states to the closest FIPS place code in Colorado.

[22] As robustness checks, I report the results using various population cutoffs as well as results when I drop all observations with missing values in the appendix. All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's website and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

than in those farther away, except among black adult males and those near highway and roads. NIBRS stand-alone marijuana sale and manufacture arrest rate is also higher in agencies closer to Colorado, while the corresponding arrest rate calculated with UCR data is slightly lower in closer regions. Considering this discrepancy, I include agency fixed effects to account for pre-existing difference between agencies, and I also report results using the natural log of arrest and offense rate as outcomes.

Compared across data sets, the adult male marijuana possession rate is the highest in NIBRS offense, followed by NIBRS stand-alone arrest, and the lowest in UCR arrest. However, adult male marijuana sale and manufacture rate is slightly lower when measured with NIBRS stand-alone arrest than with UCR arrest. The difference between offense and arrest data reflects the fact that not all marijuana possession offenses are turned into arrests, and it also stresses the importance to use offense data in addition to arrest data. The difference between NIBRS stand-alone arrest data and UCR arrest data may come from the different algorithms in defining marijuana possession and marijuana sale and manufacture incidents.[23]

Finally compared within the same distance category, marijuana possession arrests per 100k residents among black adult males is much lower than that among white adult males, which is largely due to the smaller black population in Colorado's neighboring states.

Figure 3.2 displays the time trend of the average adult male marijuana possessions per 100k residents among NIBRS reporting police agencies in different groups. The "<=150mi CO neibors" group includes police agencies in Colorado's neighboring states that are within 150 miles of the closest FIPS place code in Colorado, and are the agencies included in the final regression sample; "other CO neibors" group contains the agencies between 150 to 400 miles of the closest FIPS place code in Colorado; "CO" group refers to all agencies in Colorado; "US" group refers to all agencies in the United States. Overall, the prevalence of marijuana possession decreases over time for US, CO, and other CO neibor group. While the marijuana possession

[23]UCR yearly summarized data has readily defined codes associated with marijuana possession and marijuana sale and manufacture respectively, while NIBRS does not.

for agencies within 150 miles of the Colorado border was like other CO neibor group which decreased slightly during years 2009 through 2011, the former reversed its trend starting in 2012. More specifically, with NIBRS data, adult male marijuana possession in close agencies increased slightly in 2012, almost tripled the growth in the previous year from 2012 to 2013, continued to increase in 2014, and reached the peak in 2015. With UCR arrest data, adult marijuana possession in close agencies jumped from 2011 to 2012, stayed relatively flat during years 2012 through 2014, and experienced a big jump from 2014 to 2015.

Because Colorado passed RML in December 2012 and the first recreational marijuana store did not open until January 2014, a natural question to ask is why the marijuana possession rate had already jumped in 2012. One possibility is that residents in neighboring states of Colorado began to change their perception of harmfulness and their consumption of marijuana when they learned that RML was likely to pass in Colorado.[24] Figure C.1 depicts the 2-year moving average of percentage of juveniles aged 12 to 17 who perceive no great risk consuming marijuana once per month from the publicly available State Behavioral Barometer.[25] The percentage of juveniles who think consuming marijuana is not of great risk in Colorado's neighboring states exhibited a parallel trend compared to the United States average in 2009 to 2011. However, starting in 2012, the perceived harmfulness in the neighboring states began to increase at a higher speed than its United States counterpart; in 2014, the gap between the two groups shrink to almost half of that in 2011. The change in perception of marijuana harmfulness among juveniles may be suggestive evidence of the perception change among adults.

Another explanation is the ramping-up effect of Colorado's MML during years 2009 through 2011, when the number of registered medical marijuana patients in Colorado soared (Breathes, 2012). As MML became increasingly relaxed in Colorado, out-of-state residents may

[24]Since purchasing marijuana was still illegal for out-of-state residents in Colorado in 2012, the jump is more likely caused by residents turning to local marijuana supply or illegal marijuana on the street in Colorado.

[25]It would be nicer to plot perception change among adult males, the main subjects of this paper, but data about them is not publicly available. All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's website and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

find purchasing marijuana easier in Colorado as well, though still illegal. In Table C.3, I show that the results change very little when I control for the number of registered medical marijuana patients in the nearest Colorado counties as a proxy for the degree of MML.[26]

The final possibility is that police agencies in neighboring states that are closer to Colorado anticipated the spillover effect from Colorado's RML, so they exerted more effort in cracking down marijuana possessions ahead of time. I will examine this possibility in detail in the robustness check section. In the rest of paper, I still define 2013 to 2015 as treatment periods to err on the conservative side.

3.4 Empirical Strategy

My main difference-in-differences specification takes the following form:

$$\begin{aligned} y_{gicst} = {} & \beta_1 Treat_{ics} * Post_t + \gamma_1 UnemploymentRate_{cst} + \gamma_2 RacialComposition_{cst} \\ & + \alpha_i + \alpha_{st} + \varepsilon_{gicst} \end{aligned} \tag{3.1}$$

where y_{gicst} is the outcome of interest in demographic group g covered by agency i in county c, state s, year t. The main outcomes are the offense or arrest counts per 100k residents. I use level rather than growth rate as the main outcome, because the level speaks more directly to the incremental costs of RML on neighbouring states. But I will also present results with log ratio. *Treat* denotes whether a police agency is defined as a "treated" agency. The first way to define treatment is to consider all agencies within 150 miles to Colorado as treated and those beyond as control. Then, *Treat* is simply a dummy variable, equaling 1 for treated agencies and 0 for controls. The second way is to use the distance from an agency to Colorado as a continuous measure of its exposure to RML, because the effect of RML is unlikely to be discontinuous at a certain distance cutoff. Then, *Treat* is the distance to Colorado in units of 100 miles (*Distance*).

[26] All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's website and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

Post is a dummy equaling 1 for years 2013 through 2015, and 0 otherwise. $UnemploymentRate$ is the county-level unemployment rate. $RacialComposition$ is the county-level ratio of black males aged 20 and over. α_i denotes the police agency fixed effects, which control for agency-specific time-invariant characteristics, such as overall public security of the area covered by the agency. α_{st} refers to the state by year fixed effects, which control for time-variant and state-specific factors, such as marijuana-relevant policy in a state in a certain year. This regression is estimated by OLS, and robust standard errors are clustered at the county level. The coefficient of interest is β_1, which measures the average effect of RML on treated agencies relative to controls.[27]

One important difference between my specification and that in Hao and Cowen (2012) is that I control for state by year fixed effects, whereas they control for year fixed effects. Not only do I control for the average annual shock to marijuana possessions among all Colorado's neighbouring states, I also allow the shock to be different from state to state.

3.5 Main Results

3.5.1 Effect of RML on Marijuana Possessions

Table 3.2 displays the main results of how Colorado's RML affected adult male marijuana possessions per 100k residents in agencies covering areas closer to Colorado relative to those farther away. Overall, the closer police agencies experienced a significant increase in adult male marijuana possessions after RML compared to their farther away counterparts.

Table 3.2 panel A, B, and C present the results on marijuana possession estimated with NIBRS offense data, NIBRS stand-alone arrest data, and UCR arrest data, respectively. Column 1 controls for agency and year fixed effects, while column 2 controls for agency and state by year fixed effects, which is my preferred specification. After Colorado passed the RML,

[27] I also show results controlling for agency-specific time trend for certain outcomes. The estimates become smaller and not significant at the convention level.

police agencies within 150 miles of Colorado saw a significant increase in adult male marijuana possession offenses by 72 per 100k covered residents, or about 32 percent of the pre-2012 mean, in contrast with their farther away counterparts (panel A column 2). Similarly, stand-alone adult male marijuana possession arrest increased by around 52 arrests per 100k residents in closer agencies, also around 32 percent of the pre-2012 mean (panel B column 2). With NIBRS data, the estimates change very little when I move from column 1 to 2. With UCR data, however, the estimated effect is cut by 30 percent and no longer significant. The difference between offense and arrest data speaks to the fact that not all marijuana possession offenses are turned into arrests. The discrepancy between NIBRS stand-alone arrest data and UCR arrest data may reflect the difference in construction of arrest counts, and it also stresses the importance of comparing results across data sets. In Table C.1, I show the effect of RML on stand-alone marijuana sale and manufacture arrests. Even though the magnitude of estimate is large compared to baseline mean, I do not have enough precision. These findings suggest that the spillover effects might mainly channel through the demand side among residents in neighboring states.

Because the choice of cutoff distance for defining treatment can affect the estimates, in Figure 3.3, I plot the estimates and their 95 percent confidence intervals with the same specification in column 2 of Table 3.2 but with various cutoff distances. The estimates for the effect on marijuana possessions are fairly stable when cutoff distance is between 75 and 110 miles, but the estimates fall gradually thereafter as more farther away agencies are defined as treated. The pattern of estimates are similar between NIBRS and UCR. [28]

Considering the pattern of estimates, I use continuous distance in the unit of 100 miles to measure police agencies' relative exposure to Colorado's RML. The results are in column 3 and 4 of Table 3.2. On average, for each 100 miles closer in distance to Colorado, RML led to an

[28]Studies like Hansen, Keaton, Weber (2017) suggests that marijuana sale and consumption might be affected within very short distance from the border of RML state, like 25 miles. I did not plot the estimates for cutoff distance below 75 miles, because the number of police agencies soon drops to below 10, and the estimates become much more volatile. The pattern of estimates is largely similar when I use log incident rate as outcome, so the pattern is not driven by closer agencies always having higher level of marijuana possession.

increase of 45 per 100k residents in adult male marijuana possession offenses and 31 per 100k residents in stand-alone possession arrests, both around 20 percent of their baseline means (panel A and B column 4). For UCR data, however, the increment in arrest rate per 100 miles closer to Colorado is 13 per 100k residents and only 10 percent of the baseline mean. In Table C.1, I report results for stand-alone marijuana sale and manufacture arrests with continuous distance as treatment intensity, and estimates are still insignificant.

Now, I present some robustness checks for the effect of RML on marijuana possessions. In Table C.3, I control for the number of marijuana patients in the closest Colorado county, and the results remain similar. In Table C.4, I present the results when I drop observations with missing counts and drop observations from Texas (column 2 to 3 and column 6 to 7). Estimates either change very little or become larger in magnitude. To account for agencies moving into and out of UCR and NIBRS program from year to year, I also report the results estimated with only agencies appearing all seven years (column 4 and 8). The estimates decrease slightly but remain significant when I restrict to strongly balanced sample. I also report results when I control for agency-specific linear time trend, but the estimates are no longer significant at the convention level (column 1 and 5). In Table C.5 column 1 and 2, I use inverse hyperbolic sine and natural log of offense (arrest) rate as outcome variables. On average, Colorado's RML increased NIBRS adult male offense and stand-alone arrest rate by 15 to 28 percent for each 100 miles decrease in distance to the Colorado border. However, there is no significant effect of RML on adult male arrest rate estimated with UCR data. Moreover, in Table C.6 and Table C.7, I try various population cutoffs from 10th percentile (872 people) to 90th percentile (25,476 people) of the covered population of agencies which pass all selection criteria other than the population cutoff. The magnitude of estimates fluctuates slightly and the standard errors increase as the population cutoff increases.[29]

[29]When the cutoff population reaches 11,000 (80th percentile) and 250,000 (90th percentile), only a few hundred of observations remain, and I no longer have power to detect the effect of RML. All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's website and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

Since marijuana possession rate jumped slightly in 2012 in the time series plot, I carry out event study in Table 3.3, in which I allow treatment effect to differ from year to year. Overall, little evidence shows that marijuana possession started to increase in or before 2012, and it only started to rise in and after 2013. With NIBRS offense data, estimates became significant in 2013, increased even more in 2014 when the first recreational marijuana stores opened in Colorado, and reached the peak in 2015. With NIBRS stand-alone arrest data, the pattern is similar, but the increase in 2013 relative to 2012 was much smaller in magnitude compared to that in 2014 and 2015 (though not significantly different), especially using distance dummy as treatment. With UCR arrest data, effect did not show up until 2015. For marijuana sale and manufacture arrests, the only significant treatment effect appears in 2009. I report the results in Table C.2.[30]

Given the potential freedom in choosing whichever cutoff distance that works best for me to define distance dummy, I use continuous distance as my main measurement of exposure to RML in the following analysis.[31] Since RML did not seem to affect stand-alone marijuana sale and manufacture arrests, I will focus on marijuana possessions from now on.

3.5.2 Heterogenous Effects of RML

Because one of the most common methods of transporting drugs within the United States is via passenger vehicles (U.S. Department of Justice, 2010), illegal marijuana possession is likely to increase disproportionally near highways and roads in Colorado. In Table 3.4, I restrict marijuana possessions to those occurring near highways/roads/streets/sidewalks, using detailed locations of crime incidents in NIBRS. Not only did marijuana possession offenses and arrests near these locations increase around 26 and 24 respectively per 100k residents for each 100 miles decrease in distance to the Colorado broader (column 1), the marijuana possession offenses also

[30]All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's website and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

[31]Results using distance dummy as treatment show similar pattern are are generally larger in magnitude. Results are available upon requests.

shifted disproportionately to these locations (column 2). Moreover, the amount of possessed marijuana near these locations increased 0.33 kilograms per 100 miles decrease in distance to Colorado on average (column 3), while the amount of possessed marijuana in other locations is smaller and not significant (though the two estimates are not statistically different).[32] Whether this disproportional increase near highways and roads is due to more people driving and walking around with marijuana or due to police pulling over more people is subject to further investigation.

Table 3.5 presents the heterogenous effect of RML for black and white adult males. Colorado's RML significantly increased marijuana possessions among white adult males in police agencies closer to the Colorado border relative to those farther away, but not among black adult males.[33]

Table 3.6 shows the effect of RML on adult males of various age groups. Overall, the effect of RML decreases as age increases. The largest effect concentrates in males aged between 18 and 20 and between 21 and 24, which together account for over half of the increase in all adult male marijuana possessions after RML. Since Colorado's RML did not legalize people aged below 21 to buy and consume marijuana, the increase in possession rate among younger males is likely to come from a thriving black market which functions in a gray area of RML in Colorado.[34] One point worth noting, while younger adult males experienced large and significant increases in marijuana possession offenses and arrests, slightly older males did not. This difference is suggestive evidence that the effect is not mainly driven by police officers making more traffic stops, because police officers should not be able to distinguish males in close age groups when

[32]Results regarding amount of possessed marijuana in other locations are in Table C.10. The number of observations is smaller in column 2 than that in column 1 because total marijuana possession offenses or arrests can be zero. The number of observations is smaller in column 3 because NIBRS sometimes report amount of marijuana in units that cannot be transferred into kilograms, and I take them as missing values in column 3.

[33]A more precise measure of offense and arrest rate among black adult males will be the number of black adult male offenders or arrests per 100k black adult males, since they only take up around 1 to 1.5 percent of population in Colorado's neighboring states. But neither UCR nor NIBRS report the covered population by race.

[34]A recent investigation by *Gazette* showed that RML in Colorado allows for up to six recreational plants and six medicinal per resident, but loopholes via extended plant counts and co-ops left wiggle room for up to about 500 plants, far beyond that of other states with legalized cannabis (http://gazette.com/state-of-marijuana/marijuana-black-market). Large-scale, multinational crime organizations have exploited Colorado laws, rented multiple residential properties for large-scale cultivation sites (Colorado House Bill 17-1220, 2017).

making the stops. In Table C.7, I further examine the effect of RML on juveniles who are younger than 18-years old.[35] The estimates are all small and insignificant, which corroborates the above argument.

3.6 Robustness Checks

Marijuana possession offenses and arrests are the results of interaction between criminals and police officers. The increase in marijuana possessions among agencies closer to Colorado after RML can be due to more residents in the neighboring states crossing the border to buy marijuana. But police agencies closer to Colorado might also anticipate the spillover effect, so they may have exerted more effort in cracking down marijuana possession, such as hiring more police officers and making more traffic stops. Alternatively, police officers may be more tolerant to low-level marijuana possessions, since they do not have the necessary resources to prosecute every single marijuana possession incident they come across. The endogenous reaction of police officers can bias the estimates upward or downward, and at the very least change the interpretation of the results.

In Table 3.7, I examine the number of officers per 100k residents with an event study design. I do not find evidence that the size of police agencies closer to Colorado increased relative to farther away agencies either immediately before or after RML in Colorado. The number of police officers does not directly measure the resources that police agencies put into cracking down on certain criminal behaviors. Better measurements are police expenditure, the number of shifts, the length of shifts, and the number of traffic stops. But JEE data stopped in 2012, and data on traffic stop is not publicly available in Colorado's neighboring states.

If police officers closer to Colorado did make more search and traffic stops in response to the RML, possession offense and arrests for younger age groups and for other illicit drugs could

[35] All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's website and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787.

also increase. Table C.8 shows that RML did not affect marijuana possessions among juveniles. Table C.9 further shows that illicit drug possessions, like cocaine and heroin, did not change after Colorado's RML.[36]

In Table C.5 column 3 and 4, I use two offense (arrest) ratios as outcome variables: the ratios of marijuana possession offenses (arrests) to all offenses (arrests) among adult males, and the ratios of marijuana possession offenses (arrests) to all drug possession arrests among adult males. These two measures of arrest ratios have the advantage that they can partially account for unobserved changes in available legal resources and measurement errors from estimated populations (Chu, 2014). However, my results using these ratios lack precision, and the composition of offenses and arrests in closer agencies did not seem to change relative to farther away agencies after RML.

3.7 Discussion and Conclusion

In this paper, I estimate the effects of Colorado's RML on neighboring states' illegal marijuana use and the burden on police agencies to enforce marijuana laws based on marijuana possession offenses (arrests), as well as marijuana sale and manufacture arrests. I find that Colorado's RML increased marijuana possessions among adult males by 20 to 30 percent of the baseline mean in police agencies closer to the Colorado border relative to those farther away. But I find little evidence that marijuana sale and manufacture increased after RML. These findings suggest that the spillover effects might mainly channel through the demand side among residents in neighboring states. I further show that adult male marijuana possession offenses shifted to

[36]However, as documented by Bachhuber et al. (2014) and Chu (2015), MML lowered state opioid overdose mortality rates as well as heroin treatments and cocaine and heroin arrests. The decrease in the use of illicit drugs, such as cocaine and heroin, among residents in neighboring states possibly offset the increase resulting from search and stop by police officers. Given the still-under-debate relation between marijuana use and the use of other illicit drugs, this test is not clean. All appendices are available at the end of this article as it appears in JPAM online. Go to the publisher's website and use the search engine to locate the article at http://www3.interscience.wiley.com/cgi-bin/jhome/34787

locations near highways and roads. The amount of marijuana seized near these locations also increased, while the amount seized in other locations did not. Finally, to examine potential changes in law enforcement in response to RML, I examine the effects of RML on law enforcement size, marijuana possessions among juveniles, and other illicit drug possessions. I do not find evidence that RML changed any of the above measures.

There are several limitations of this study. First, as already discussed in the paper, potential endogenous reaction of police could bias the estimates in unknown direction. Second, the arrest and offense data are not able to identify whether the increase in use comes from initiation or increased demand among existing users.[37] Third, this paper is not able to identify the source of illegal marijuana in neighboring states. Namely, whether the source is a legal purchase in recreational marijuana stores or an illegal purchase on the street in Colorado. Answers to this question can help RML states and their neighboring states better cope with the spillover effect. Last, this paper uses Colorado's RML as a case study because of its early passage and unique geographical location. States should take caution in generalizing the estimates in this paper to RMLs in other states, since RMLs are not homogeneous across states.

Taken together, the findings add to the heated policy debate over the pros and cons of RMLs. However, due to the early stage of the literature on RMLs, this paper by itself is far from providing definitive conclusions. Rather, this paper provides evidence that some indicators of marijuana use in the neighboring states do respond to RML in Colorado, and that burden on police departments in these states do increase.

Acknowledgements

Chapter 3 is being prepared for publication. The dissertation author was the sole author on this chapter.

[37] As discussed in Chu (2014), literature generally suggests a small or nonexistent effect on the extensive margin. But since estimates in existing studies often come with large estimated standard errors, this conclusion should be treated with some caution.

3.8 Figures and Tables

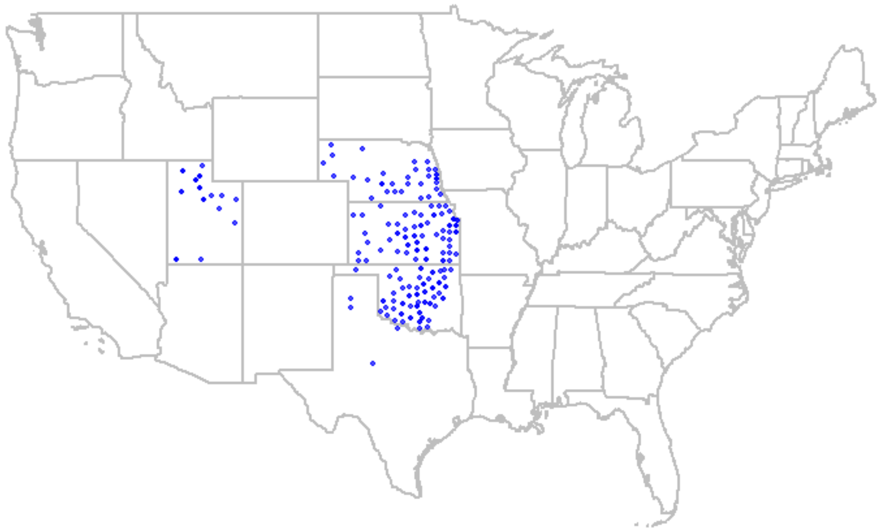
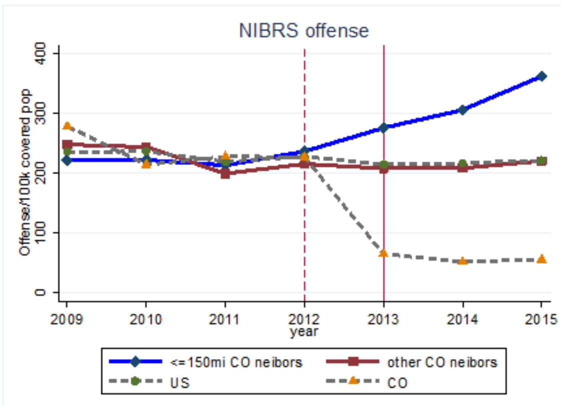
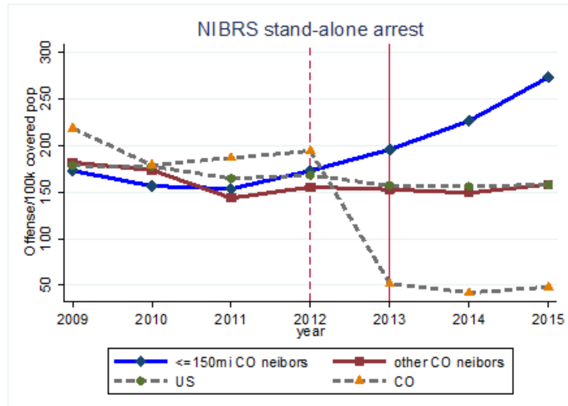


Figure 3.1: Location of police agencies

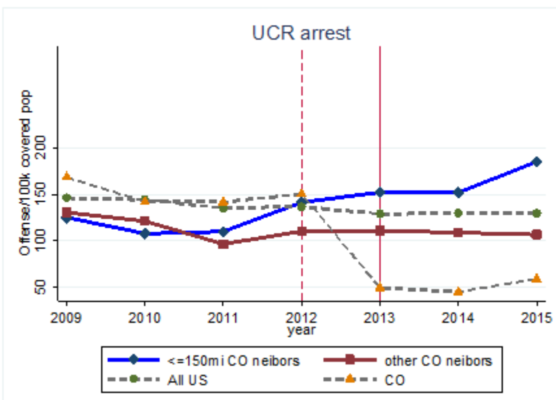
Notes: Each dot represents the FIS placing code assigned to a specific police agency in the NIBRS sample (1490 agency×year observations and 448 unique police agencies).



Panel A



Panel B



Panel C

Figure 3.2: Time trend of adult male marijuana possession

Notes: Dashed vertical line marks the time when Colorado passed RML (Dec 2012), and solid vertical line marks the time when the effect of RML was supposed to appear (i.e. one year after the passage of RML). Marijuana possession rates are counts of marijuana possession offenses or arrests per 100k residents covered by a police agency. "<=150mi CO neibor" is the average of marijuana possession rate among police agencies in neighboring states within 15 miles to the Colorado border; "Other CO neibors" is the average among other police agencies in the neighboring states of Colorado; "CO" is the average of all agencies in Colorado; "All US" is the average of all agencies in the U.S. Only agencies reporting to NIBRS are included in the figures.

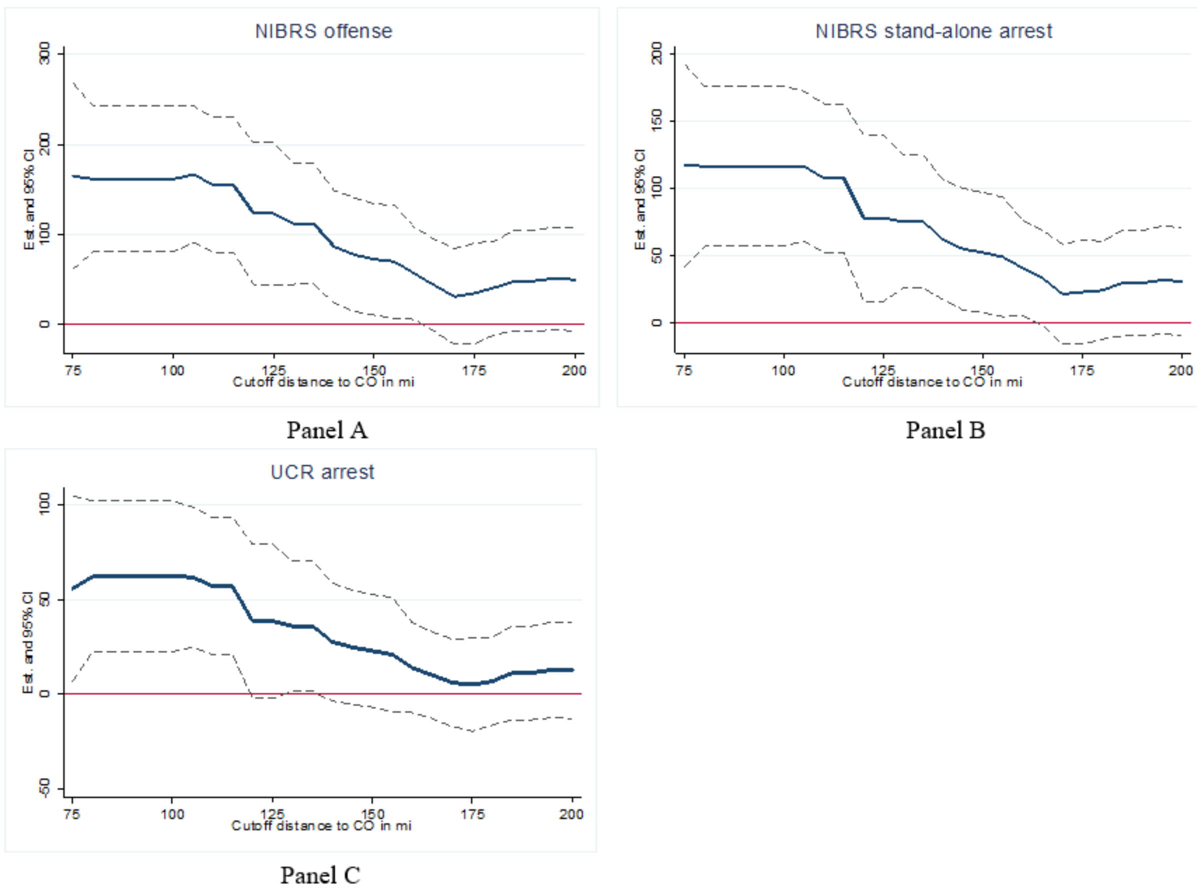


Figure 3.3: Varying distance cutoff - Effect of RML on adult male marijuana possession

Notes: The estimates are from regressions using the preferred specifications (column (2) in Table 2) with various distances from CO (5 miles apart from 75 miles all the way to 200 miles) as cutoffs for defining treated agencies.

Table 3.1: Baseline summary statistics

	≤150mi to CO		150-400mi to CO		All agencies	
VARIABLES	Mean	SD	Mean	SD	Mean	SD
NIBRS offense						
MJ possession counts(per 100k pop.)						
Juvenile	55.54	57.18	50.89	49.87	51.53	50.94
Adult	230.63	175.53	225.31	194.83	226.05	192.17
White 18+	197.58	150.10	189.70	164.07	190.79	162.14
Black 18+	18.94	26.29	26.89	46.10	25.79	43.96
Highway 18+	103.31	90.10	119.84	131.20	117.55	126.38
MJ possession hwy prop.(%)	45	21	52	25	51	24
MJ quantity hwy(kg)	0.03	0.07	0.27	4.24	0.24	3.95
NIBRS stand-alone arrest						
MJ possession counts(per 100k pop.)						
Juvenile	41.29	45.32	34.61	40.34	35.53	41.10
Adult	170.2	135.97	162.46	163.38	163.54	159.81
White 18+	145.61	117.35	137.01	136.83	138.20	134.26
Black 18+	13.47	20.24	19.20	37.16	18.41	35.35
Highway 18+	83.12	75.05	96.43	119.75	96.43	114.68
MJ possession hwy prop.(%)	51	23	58	26	58	26
MJ quantity hwy(kg)	0.03	0.11	0.26	4.06	0.28	3.96
MJ sale/manufacture(per 100k pop.)	19.93	32.56	17.14	30.78	17.52	31.02
UCR arrest						
MJ possession counts(per 100k pop.)						
Juvenile	35.24	40.61	28.68	37.91	115.70	131.97
Adult	127.48	122.02	113.82	133.48	29.58	38.33
White 18+	137.76	130.78	121.69	143.86	123.90	142.17
Black 18+	10.73	18.51	16.48	35.85	15.69	34.04
MJ sale/manufacture(per 100k pop.)	23.98	34.36	28.96	58.30	28.28	55.63
Agency						
Distance to CO(100mi)	0.97	0.39	2.75	0.79	2.50	0.97
Covered pop.(100k)	0.22	0.40	0.18	0.37	0.19	0.37
Officer(per 100k pop.)	173.46	48.02	183.73	57.54	182.30	56.40
County						
MJ patient	90.30	97.73	61.24	58.96	65.27	66.40
Unemployment(%)	5.05	1.62	6.31	1.72	6.14	1.76
Black male 20+(%)	0.75	0.98	1.44	1.55	1.35	1.51

Notes: Summary statistics are calculated using pre-treatment years 2009-2012. "MJ possession hwy prop." = marijuana possession counts near highway (NIBRS location code "13")/total marijuana possession counts. "Covered pop." is the number of residents covered by an agency. "MJ patient" is the number of medical marijuana registered patients during December of each year in the closest CO county to an agency. "Num of obs in regression sample" is referring to the sample with NIBRS data.

Table 3.2: Main result - Effect of RML on adult male marijuana possession

	(1)	(2)	(3)	(4)
Panel A: NIBRS offense				
Distance to CO≤150mi*Post	74.045**	72.256**		
	(30.324)	(31.558)		
Distance to CO*Post			-37.447***	-45.047***
			(11.240)	(13.864)
Baseline mean	226.05	226.05	226.05	226.05
Observations	1,490	1,490	1,490	1,490
R-squared	0.025	0.083	0.035	0.092
Panel B: NIBRS stand-alone arrest				
Distance to CO≤150mi*Post	56.712**	52.226**		
	(21.743)	(22.773)		
Distance to CO*Post			-28.231***	-31.457***
			(9.189)	(10.510)
Baseline mean	163.54	163.54	163.54	163.54
Observations	1,490	1,490	1,490	1,490
R-squared	0.024	0.091	0.033	0.096
Panel C: UCR arrest				
Distance≤150mi*Post	30.079**	22.902		
	(15.139)	(15.166)		
Distance to CO*Post			-16.937**	-12.912*
			(6.548)	(7.134)
Baseline mean	115.70	115.70	115.70	115.70
Observations	1,468	1,468	1,468	1,468
R-squared	0.029	0.129	0.035	0.130
Agency FE	Yes	Yes	Yes	Yes
Year FE	Yes	No	Yes	No
State*Year FE	No	Yes	No	Yes
Number of clusters	148	148	148	148

Notes: Marijuana possession rates are the counts of marijuana possession per 100k residents covered by an agency. Baseline mean is the mean of y during 2009-2012 across all agencies in the regression sample. Distance dummy equals to 1 if an agency is within 150 miles from Colorado. Distances are measured in the unit of 100 miles. Post equals to 1 for year 2013-2015. All specifications include the county unemployment rate and the county population ratio of black male aged 20+. Standard errors (in parentheses) are clustered at county level. *** p<0.01; ** p<0.05; * p<0.1.

Table 3.3: Event study - Effect of RML on adult male marijuana possession

	Treatment: distance to CO≤150mi			Treatment: distance to CO		
VARIABLES	NIBRS offense	Stand-Alone arrest	UCR arrest	NIBRS offense	Stand-Alone arrest	UCR arrest
Treat*2009	23.316	31.952	18.011	21.599	11.324	11.437
	(28.643)	(26.788)	(21.146)	(18.424)	(16.499)	(13.447)
Treat*2010	20.099	13.356	-9.419	14.227	11.574	12.079
	(33.499)	(28.895)	(22.648)	(18.166)	(15.173)	(11.649)
Treat*2011	23.059	13.385	1.431	-1.509	-1.963	3.221
	(25.932)	(22.337)	(19.103)	(11.795)	(10.033)	(7.481)
Treat*2013	63.240**	38.010*	13.552	-28.077**	-17.020	-2.917
	(26.521)	(19.734)	(11.815)	(12.502)	(10.319)	(6.678)
Treat*2014	77.815*	61.849**	20.336	-40.539**	-29.954**	-6.510
	(44.052)	(28.635)	(13.153)	(18.683)	(13.004)	(7.788)
Treat*2015	128.655***	103.752***	43.801**	-46.738***	-36.576***	-12.757
	(43.259)	(32.089)	(18.918)	(16.573)	(12.881)	(8.824)
Baseline mean	226.05	163.54	115.70	226.05	163.54	115.70
Observations	1,490	1,490	1,468	1,490	1,490	1,468
R-squared	0.086	0.095	0.131	0.095	0.099	0.132
Agency FE	Yes	Yes	Yes	Yes	Yes	Yes
State*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Num of clusters	148	148	148	148	148	148

Notes: Year 2012 is the omitted year. Marijuana possession rates are the counts of marijuana possession per 100k residents covered by an agency. Baseline mean is the mean of y during 2009-2012 across all agencies in the regression sample. Distance dummy equals to 1 if an agency is within 150 miles from Colorado. Distances are measured in the unit of 100 miles. Post equals to 1 for year 2013-2015. All specifications include the county unemployment rate and the county population ratio of black male aged 20+. Standard errors (in parentheses) are clustered at county level. *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Table 3.4: Heterogeneity - Effect of RML on adult male marijuana possession near highway

VARIABLES	Mj poss/100k	Mj poss near hwy(%)	Mj poss (kg)
Panel A: NIBRS offense			
Distance to CO*Post	-26.469***	-0.008*	-0.323*
	(8.563)	(0.005)	(0.194)
Baseline mean	117.55	0.51	0.24
Observations	1,490	1,385	1,273
R-squared	0.085	0.048	0.016
Number of clusters	148	144	144
Panel B: Stand-Alone arrest			
Distance to CO*Post	-23.971***	0.001	-0.346*
	(7.409)	(0.006)	(0.205)
Baseline mean	96.43	0.58	0.26
Observations	1,490	1,336	1,244
R-squared	0.093	0.034	0.017
Number of clusters	148	143	143
Agency FE	Yes	Yes	Yes
State*Year FE	Yes	Yes	Yes

Notes: *** p<0.01; ** p<0.05; * p<0.1. Highway status is defined as offense with NIBRS location code "13" (near highway/road/alley/street/sidewalk). Marijuana possession rates are the counts of marijuana possession per 100k residents covered by a police agency near highway. Proportion of marijuana possession near highway = counts near highway / total counts. Amount of possessed marijuana is in unit KG. Baseline mean is Baseline mean is the mean of y during 2009-2012 across all agencies in the regression sample. Distances are measured in the unit of 100 miles. Post equals to 1 for year 2013-2015. All specifications include the unemployment rate and the population ratio of black male aged 20+ at county level. Standard errors (in parentheses) are clustered at county level.

Table 3.5: Heterogeneity - Effect of RML on adult male marijuana possession by racial group

VARIABLES	White	Black
Panel A: NIBRS offense		
Distance to CO*Post	-40.858***	-2.846
	(11.492)	(2.841)
Baseline mean	190.79	25.79
Observations	1,490	1,490
R-squared	0.089	0.048
Number of clusters	148	148
Panel B: NIBRS stand-alone arrest		
Distance to CO*Post	-29.829***	-0.717
	(8.944)	(2.158)
Baseline mean	138.20	18.41
Observations	1,490	1,490
R-squared	0.094	0.048
Number of clusters	148	148
Panel C: UCR arrest		
Distance to CO*Post	-14.011*	-0.091
	(7.991)	(2.110)
Baseline mean	123.90	15.69
Observations	1,468	1,468
R-squared	0.128	0.049
Number of clusters	148	148
Agency FE	Yes	Yes
State*Year FE	Yes	Yes

Notes: Marijuana possession rates are the counts of marijuana possession per 100k residents covered by an agency. Baseline mean is the mean of y during 2009-2012 across all agencies in the regression sample. Distances are measured in the unit of 100 miles. Post equals to 1 for year 2013-2015. All specifications include the unemployment rate and the population ratio of black male aged 20+ at county level. Standard errors (in parentheses) are clustered at county level. *** p<0.01; ** p<0.05; * p<0.1.

Table 3.6: Heterogeneity - Effect of RML on adult male marijuana possession by age group

VARIABLES	18-20	21-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59
Panel A: NIBRS offense									
Distance to CO*Post	-15.605***	-8.408**	-4.719	-5.771***	-1.917	-2.141	-3.452***	-1.448	-1.429**
	(5.301)	(3.620)	(2.931)	(1.726)	(1.663)	(1.306)	(1.180)	(0.966)	(0.616)
Baseline mean	76.66	51.12	37.17	19.98	13.25	9.53	8.51	5.71	2.79
Observations	1,490	1,490	1,490	1,490	1,490	1,490	1,490	1,490	1,490
R-squared	0.071	0.047	0.036	0.060	0.091	0.053	0.039	0.029	0.030
Panel B: Stand-Alone arrest									
Distance to CO*Post	-11.810***	-6.125**	-2.407	-4.310***	-1.585	-1.732	-1.459*	-0.780	-1.352***
	(4.077)	(2.923)	(2.132)	(1.371)	(1.227)	(1.107)	(0.792)	(0.806)	(0.466)
Baseline mean	19.26	13.28	10.26	13.98	9.04	6.80	6.02	3.86	1.97
Observations	1,490	1,490	1,490	1,490	1,490	1,490	1,490	1,490	1,490
R-squared	0.075	0.047	0.042	0.060	0.090	0.053	0.035	0.028	0.042
Panel C: UCR arrest									
Distance to CO*Post	-5.416	-1.942	-0.500	-1.682	0.117	-1.176	-0.451	-0.980	-0.906**
	(3.378)	(2.100)	(1.572)	(1.213)	(1.001)	(0.785)	(0.663)	(0.704)	(0.355)
Baseline mean	41.14	27.3	18.98	9.86	6.05	4.34	3.75	2.36	1.33
Observations	1,468	1,468	1,468	1,468	1,468	1,468	1,468	1,468	1,468
R-squared	0.088	0.056	0.067	0.039	0.116	0.059	0.038	0.035	0.029
Agency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State*Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Marijuana possession rates are the counts of marijuana possession per 100k residents covered by an agency. Baseline mean is the mean of y during 2009-2012 across all agencies in the regression sample. Distances are measured in the unit of 100 miles. Post equals to 1 for year 2013-2015. All specifications include the unemployment rate and the population ratio of black male aged 20+ at county level. Standard errors (in parentheses) are clustered at county level, and number of clusters are 148 for all regressions. *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Table 3.7: Robustness check - Effect of RML on size of law enforcement agencies

VARIABLES	Officer/100k
Distance to CO*2009	-5.421**
	(2.388)
Distance to CO*2010	0.120
	(1.649)
Distance to CO*2011	-0.113
	(2.146)
Distance to CO*2013	-0.584
	(1.543)
Distance to CO*2014	1.298
	(1.881)
Distance to CO*2015	0.630
	(2.155)
Baseline mean	182.30
Observations	1,490
R-squared	0.073
Agency FE	Yes
State*Year FE	Yes
Number of clusters	148

Notes: Year 2012 is the omitted year. Outcome is the number of officers per 100k residents covered by the agency in 2012 (from LEOKA). Baseline mean is the mean of y during 2009-2012 across all agencies in the regression sample. Distances are measured in the unit of 100 miles. All specifications include the unemployment rate and the population ratio of black male aged 20+ at county level. Standard errors (in parentheses) are clustered at county level. *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Appendix A

Chapter 1 Appendix

A.1 Variable Definitions

Vote share. The vote share for analyst i participating in NF industry j in year t is:

$$X_{ijt} = \frac{Score_{ijt}}{\sum_{i=1}^{15} Score_{ijt}}. \tag{A.1}$$

Win(t). Win(t) equals one if an analyst is a NF winner (or finalist) in year t, and zero otherwise.

Adjusted 2-day Buy-and-hold Abnormal Return.

$$BHAR_{st} = \prod_{\tau=t}^{t+1}(1+R_{s\tau}) - \prod_{\tau=t}^{t+1}(1+R_{s\tau}^{DGTW}), \tag{A.2}$$

where day τ is the day of the measurement. $R_{s\tau}$ is the return of stock s on day τ, calculated with cash dividend reinvestment. $R_{s\tau}^{DGTW}$ is the return on a benchmark portfolio with similar size, book-to-market, and momentum characteristics as stock s on day τ. We follow Daniel et al. (1997) to construct the benchmark portfolios. First, in each month, we classify the universe of A share stocks listed on the Shanghai and the Shenzhen Stock Exchanges into three quintile groups, based firm size, book-to-market ratio, and momentum (the return of the stock in the previous 12 months). Stocks in each of the 125 (5×5×5) cells form a passive stock portfolio (DGTW portfolio). We calculate the return of each DGTW portfolio using equal weights. The DGTW portfolio where stock s belongs will be the benchmark portfolio for stock s. The 2-day buy-and-hold abnormal return of stock s on day t is then the difference between the realized return R_s between day t and $t+1$ and the expected return R_s^{DGTW} between day t and $t+1$ We further assign a minus sign to the abnormal return if the associated stock rating is neutral, sell, or strong sell.

Adjusted 2-day Fama-French Buy-and-hold Abnormal Return.

$$BHAR_{st}^{FF} = \prod_{\tau=t}^{t+1}(1+R_{s\tau}) - \prod_{\tau=t}^{t+1}(1+R_{s\tau}^{FF}) \tag{A.3}$$

$$R_{s\tau}^{FF} = R_{f\tau} + \hat{\alpha}_s + \hat{\beta}_s(R_{m\tau} - R_{f\tau}) + \hat{s}_s SMB_\tau + \hat{h}_s HML_\tau + \hat{r}_s RMW_\tau + \hat{c}_s CMA_\tau, \tag{A.4}$$

where $R_{s\tau}$ is the raw return of stock s on day τ. $R_{s\tau}^{FF}$ is the expected return of stock s on day τ computed using the Fama-French five-factor model (Fama and French, 2015). Factor loadings for each stock s are estimated in a one-year period ending 90 days before the date of measurement τ. The 2-day buy-and-hold abnormal return of stock s on day t is then the difference between the realized return R_s between day t and $t+1$ and the expected return R_s^{FF} between day t and $t+1$. We further assign a minus sign to the abnormal return if the associated stock rating is neutral, sell, or strong sell.

2-day Cumulative Abnormal Turnover.

$$CAT_{st} = \sum_{\tau=t}^{t+1}(log(Turnover_{s\tau}+\delta) - \frac{\sum_{\tau=t-60}^{\tau=t-30} log(Turnover_{s\tau}+\delta)}{30}) \tag{A.5}$$

$$Turnover_{s\tau} = \frac{Trading\ Volume_{s\tau}}{Total\ Outstanding_{s\tau}}, \tag{A.6}$$

where trading volume$_{s\tau}$ is the number of shares traded for stock s on day τ and total outstanding$_{s\tau}$ is the total number of outstanding shares for stock s on day τ. Since the time series of daily turnover is not stationary, we follow Lo and Wang (2000) and take natural log. To avoid problems caused by zero daily trading volume, we follow Loh and Atulz (2011) and add a small constant δ (0.00000255) to the turnover before taking logs. To ensure we have enough sample to calculate the moving average between $t-60$ and $t-30$, we require the stock to have at least 15 valid daily turnovers in this period. Otherwise, we code the abnormal turnover as missing. About 5.69 percent of observations are coded as missing because of this reason.

Forecast Error.

$$Forecast\ Error_{ist} = \frac{\mid EPS\ Forecast_{ist} - Actual\ EPS_{st} \mid}{Book\ Value_{s,t-1}}, \tag{A.7}$$

where $EPS\ Forecast_{ist}$ is the last forecast on the year-end earnings per share for stock s in year t issued by analyst i. $Actual\ EPS_{st}$ is the realized year-end earnings per share for stock s in year t. $Book\ Value_{s,t-1}$ is the year-end book value of equity per share for stock s in year $t-1$.

Forecast Bias.

$$Forecast\ Bias_{ist} = \frac{(EPS\ Forecast_{ist} - Actual\ EPS_{st})}{Book\ Value_{s,t-1}}, \tag{A.8}$$

where $EPS\ Forecast_{ist}$ is the last forecast on the year-end earnings per share for stock s in year t issued by analyst i. $Actual\ EPS_{st}$ is the realized year-end earnings per share for stock s in year t. $Book\ Value_{s,t-1}$ is the year-end book value of equity per share for stock s in year $t-1$.

Forecast Boldness.

$$Forecast\ Boldness_{ist} = \frac{\mid EPS\ Forecast_{ist} - EPS\ Forecast_{-ist} \mid}{Book\ Value_{s,t-1}}, \tag{A.9}$$

where $EPS\ Forecast_{-ist}$ is the average forecast on the year-end earnings per share for stock s in year t issued by all analysts except analyst i. $EPS\ Forecast_{ist}$ is the last forecast on the year-end earnings per share for stock s in year t issued by analyst i. $Actual\ EPS_{st}$ is the realized year-end earnings per share for stock s in year t. $Book\ Value_{s,t-1}$ is the year-end book value of equity per share for stock s in year $t-1$. Forecast boldness measures how far an analyst's forecast deviates from the consensus forecast of all other analysts simultaneously covering the same stock.

Teammate quality. Teammate quality for analyst i's coauthored forecast report is the average baseline forecast error among all authors *excluding* analyst i. Teammate quality for analyst i's solo-author forecast report is the average baseline forecast error among all analysts ever collaborated with analyst i in the post-award period. The baseline forecast error for each analyst is the average forecast error among all last forecast reports for each stock covered by the analyst in the year before the award.

Market capitalization (Market cap). Market cap is the total number of outstanding shares times the share price.

Book-to-market ratio (B/M ratio). B/M ratio is a stock's book value divided by its market capitalization.

Undervalued. Undervalued equals one if the 3rd-quarter B/M ratio of a stock is ≥ 1, and zero otherwise. B/M ratio ≥ 1 means that the cost to replace a firm's assets is greater than the value of its stock, implying that the stock is undervalued by the market.

Price-to-earnings ratio (P/E ratio). P/E ratio is the stock price divided by the stock's earnings per share. It represents the dollar amount an investor needs to invest in a company to receive one dollar of the company's earnings. P/E ratio is coded as missing if it is negative. Higher P/E ratio implies that investors expect higher earnings growth in the future.

Beta. Beta is estimated using a CAPM model based on the daily stock return in the 250-trading-day period before the event day. Daily return of market portfolio is the value-weighted daily return among all A-share stocks. Daily risk-free rate is calculated from fixed annual interest rate.

A.2 Portfolio Construction for Informed Investors

Figure 1.5 shows that the return of stocks recommended by both analysts just above and below the winner cutoff increases during the notification week but the return for those recommended by analysts just below the cutoff decreases right after award announcement. One story consistent with this price pattern is that the list of finalists is leaked and traded on before the announcement.

Suppose an informed investor is told that the prices for five out of seven stocks will increase on the coming Monday. A viable trading strategy is to purchase all seven stocks now and sell the two stocks whose prices do not increase on the next Monday. Informed investors' selling of the two stocks pushes down the prices, while uninformed investors extract signal from the drop in prices and follow suit, further lowering the prices.

To quantify the gain for this trading strategy, we construct a calendar portfolio consisting of the last stocks with buy or strong buy recommendations from both winners and failed finalists within 30 days before the award announcement. All stocks enter the portfolio on the Monday in the notification week. The stocks recommended by failed finalists exit at the opening price on the Monday after the award announcement and those recommended by winners exit at the closing price on the Friday after the award announcement. The individual stock returns are then aggregated to the portfolio return using equal-weighted, equal-investment, or value-weighted method.[1]

The portfolio return is measured using Fama and French (2015) five-factor model,

$$R_{jt} - R_{ft} = \alpha_j + \beta_j(R_{mt} - R_{ft}) + s_j SMB_t + h_j HML_t + r_j RMW_t + c_j CMA_t + \varepsilon_{jt}, \qquad \text{(A.10)}$$

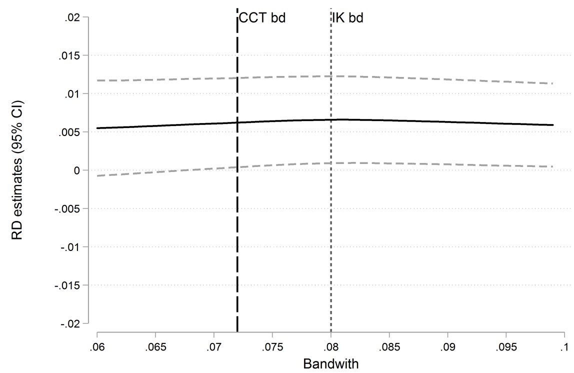
where the dependent variable is the daily return on a portfolio j of recommendations less the

[1] Under equal-weighted, the weight for the return of each stock is one over the total number stocks in the portfolio. Under equal-investment, the weight is the compounded return for each stock since the beginning of the portfolio till the day before. Under value-weighted, the weight is market capitalization of the stock on the same day.

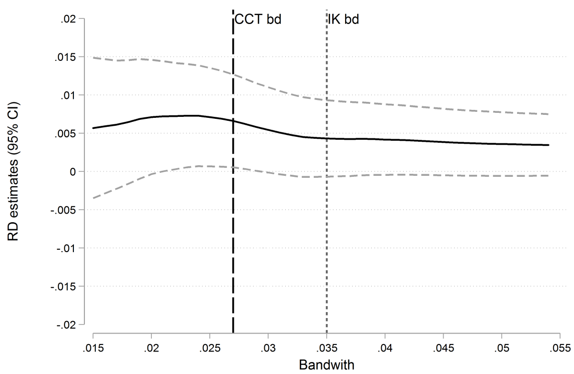
risk-free rate on day t, and the right-hand-side variables are the return on the value-weighed market index less the risk-free rate ($R_{mt} - R_{ft}$), the daily return on a portfolio of small-cap stocks less the return on a portfolio of large-cap stocks (SMB_t), the daily return on a portfolio of high book-to-market stocks less the return on a portfolio of low book-to-market stocks (HML_t), the daily return on a robust operating profitability portfolio less the return on a weak operating profitability portfolio (RMW_t), and the daily return on a conservative investment portfolio less the return on a weak operating profitability portfolio (CMA_t). The factors and the risk-free interest rate are from CSMAR. The average daily portfolio return is measured by the intercept ($\hat{\alpha}$). As shown in Table 1.7, the portfolio earns a significant average daily return of 18 basis point in the 10 days around the award announcement.[2]

[2] There is short-sell restriction in China, so we do not present the return for the Sell portfolio.

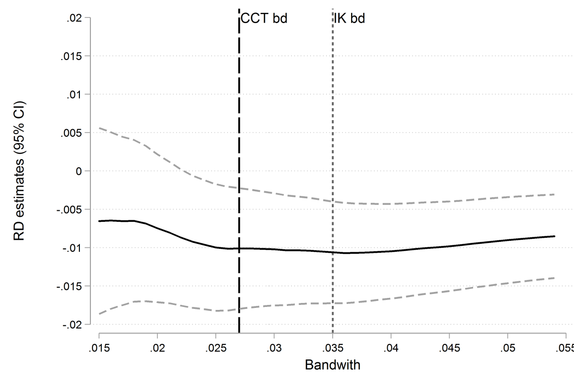
A.3 Additional Figures



A: Winner, 1^{st} trading day after award



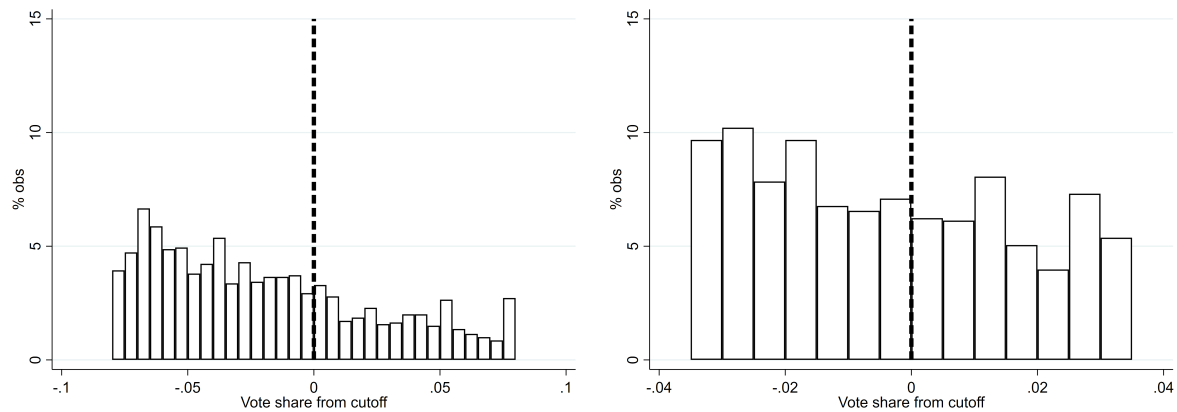
B: Finalist, notification week



C: Finalist, 1^{st} trading day after award

Figure A.1: Robustness check – Main results under various bandwidths

Notes: The figure plots the estimates of coefficients on $Win(t)$ and their 95% confidence intervals for regressions in Table 4 column (3) (panel A), Table 7 panel A column (3) (panel B) and Table 7 panel B column (3) (panel C) under various bandwidths. The vertical long-dashed linse denote the CCT bandwidths (0.072 for the winner cutoff and 0.027 for the finalist cutoff); the vertical short-dashed lines denote the IK bandwidths (0.08 for the winner cutoff of and 0.035 for the finalist cutoff). All specifications and regression methods mirror the ones in the corresponding tables.



A: Winner Cutoff **B:** Finalist Cutoff

Figure A.2: Validity of RD - Density of centered vote share

Notes: The figure plots the density of centered vote share for the analysts in the main RD samples at the winner cutoff (panel A) and at the finalist cutoff (panel B). The dashed vertical lines denote the centered vote share at rank fifth (panel A) and rank seventh (panel B) in each industry each year. The distribution excludes the observations with centered vote share equaling zero, because the way we center the vote share creates a mechanical spike at zero. Refer to section 3.2 for a detailed discussion.

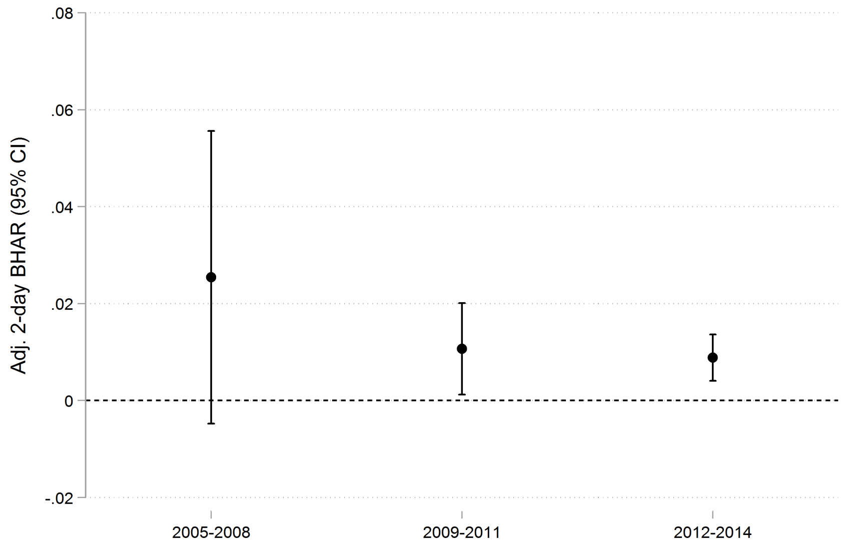


Figure A.3: Heterogeneity – Main results by sample period
Winner cutoff

Notes: The figure plots the coefficients on $Win(t)$ and 95% confidence intervals from the regressions mirroring that in Table 4 column (3) but based on stock recommendations issued in three sub-periods of the whole sample period, i.e., 2005-2008, 2009-2011 and 2012-2014.

A.4 Additional Tables

Table A.1: Sample selection - Probability of entering base samples from raw sample

VARIABLES	(1) Sample with valid analyst ID	(2) Main RD recommendation	(3) Post-award recommendation	(4) Post-award forecast
Panel A: Winner Cutoff				
Win(t)	-0.01646	-0.01362	-0.01663	0.00956
	(0.01159)	(0.04438)	(0.02148)	(0.02224)
% Raw sample	0.97	0.57	0.85	0.85
Observations	2,992	2,992	2,992	2,992
R-squared	0.134	0.123	0.055	0.141
Number of clusters	146	146	146	146
Bandwidth	0.080	0.080	0.08	0.08
Panel B: Finalist Cutoff				
Win(t)	0.00466	0.00929	0.01239	0.03875
	(0.01432)	(0.06184)	(0.03276)	(0.02995)
% of Raw sample	97	64	85	85
Observations	1,748	1,748	1,748	1,748
R-squared	0.124	0.156	0.086	0.169
Number of clusters	146	146	146	146
Bandwidth	0.035	0.035	0.035	0.035

Notes: The data are at the analyst $\times$ year level consisting of analysts whose margin of vote share is within 0.08 from the winner cutoff (panel A) or within 0.035 from the finalist cutoff (panel B). Dependent variables are the probability of an analyst $\times$ year pair entering different regression samples. *Win(t)* equals 1 if the analyst is above the winner cutoff (panel A) or the finalist cutoff (panel B) in the year of the award announcement. *% of Raw sample* refers to the proportion of analyst $\times$ year pairs entering a certain regression sample from the raw sample. All regressions control for year, brokerage house, and NF industry fixed effects and are estimated using a linear RD model and triangular weights. Standard errors in parentheses are clustered by NF industry $\times$ year. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.2: Sample selection - Summary statistics in raw and base samples

VARIABLES	(1) Raw sample	(2) Sample with valid analyst ID	(3) Main RD rec winner cutoff	(4) Main RD rec finalist cutoff	(5) Post base recommendation	(6) Post base forecast
Male	0.72371	0.72277	0.72128	0.72490	0.72069	0.72004
	(0.44722)	(0.44770)	(0.44848)	(0.44666)	(0.44873)	(0.44905)
		[0.9296]	[0.8446]	[0.9201]	[0.7842]	[0.7389]
Master/above	0.90626	0.90805	0.92052	0.91820	0.91542	0.91159
	(0.29151)	(0.28900)	(0.27056)	(0.27412)	(0.27830)	(0.28393)
		[0.7988]	[0.0732]	[0.1197]	[0.1974]	[0.4556]
Experience	3.19461	3.18954	3.26222	3.26335	3.19239	3.20429
	(2.3922)	(2.40058)	(2.41758)	(2.41685)	(2.38763)	(2.38539)
		[0.9281]	[0.3015]	[0.2747]	[0.9695]	[0.8672]
Win(t-1)	0.27471	0.27378	0.294173	0.28530	0.27979	0.28236
	(0.44643)	(0.44596)	(0.45578)	(0.45165)	(0.44896)	(0.45021)
		[0.9281]	[0.1110]	[0.3656]	[0.6382]	[0.4782]
Finalist(t-1)	0.37570	0.37521	0.39286	0.38385	0.38541	0.38774
	(0.48437)	(0.48424)	(0.48850)	(0.48642)	(0.48677)	(0.48731)
		[0.9651]	[0.1932]	[0.5196]	[0.4064]	[0.3029]
Ttoal obs	3,753	3,638	2,128	2,415	3,181	3,198
Unique analyst	1,633	1,600	1,157	1,249	1,412	1,423

Notes: The data are at the analyst × year level consisting of all analysts regardless of their margin of vote share. Variables are defined the same as those in Table 1. Standard errors are in the parentheses. *P* values for the difference in means between the raw sample and other samples are reported in the square brackets.

Table A.3: Placebo test – Market reaction in the notification week
Winner cutoff

	(1) Avg. 2-d BHAR	(2) Avg. 2-d BHAR	(3) Avg. 2-d BHAR
Win(t)	-0.00004	-0.00099	0.00029
	(0.00225)	(0.00242)	(0.00238)
Baseline mean	0.00474	0.00474	0.00474
Observations	1,535	1,535	1,535
R-squared	0.024	0.101	0.200
Number of clusters	146	146	146
Year FE	Yes	Yes	Yes
Brokerage FE	No	Yes	Yes
NF industry FE	No	Yes	Yes
Stock industry FE	No	No	Yes
Bandwidth	0.080	0.080	0.080

Notes: This table reports the placebo tests for the main results. Samples, dependent variables and specifications correspond to the ones in Table 4 column (3). The only difference is that the dependent variable is now the average adjusted two-day BHAR between Monday and Thursday in the notification week. Standard errors in parentheses are clustered by NF industry $\times$ year.

Table A.4: Validity of RD – Brokerage characteristics

VARIABLES	(1) Diversity	(2) Win(t-1)	(3) Total asset	(4) Net profit	(5) Listed	(6) Number of analysts	(7) Held by mutual fund
Panel A: Winner Cutoff							
Win(t)	0.04390	-0.04170	5.04647	0.32280	0.00734	2.05625	0.02388
	(0.03948)	(0.05721)	(4.56138)	(0.32944)	(0.05911)	(2.75858)	(0.05628)
Outcome mean	0.82915	0.25000	38.53665	1.84330	0.40823	49.82209	0.33146
Observations	234	1,068	932	931	1,068	1,068	1,068
R-squared	0.407	0.175	0.099	0.100	0.098	0.167	0.151
Number of clusters	26	146	146	146	146	146	146
Bandwidth	4.34 (rank)	0.080	0.080	0.080	0.080	0.080	0.080
Panel B: Finalist Cutoff							
Win(t)	-0.01856	-0.10190	-2.10261	-0.36321	0.02575	-3.79834	-0.09037
	(0.04577)	(0.07464)	(4.40056)	(0.33779)	(0.07288)	(2.79356)	(0.05999)
Outcome mean	0.87521	0.33284	36.26664	1.67445	0.37810	48.75328	0.30364
Observations	182	685	596	596	685	685	685
R-squared	0.393	0.184	0.135	0.126	0.119	0.210	0.172
Number of clusters	26	143	142	142	143	143	143
Bandwidth	3.81 (rank)	0.035	0.035	0.035	0.035	0.035	0.035

Notes: The data are at the NF industry $\times$ rank level in column (1) and at the brokerage $\times$ NF industry $\times$ year level in columns (2) through (8) for brokerage houses with analysts whose margin of vote share is within 0.08 from the cutoff of winner (panel A) and 0.035 from the cutoff of finalist (panel B). Sample size varies due to missing dependent variables. *Diversity* for a rank in a NF industry is the number of distinct brokerages divided by the number of occurrence for that rank in that industry across years (the lower the value, the less variety in brokerages at the rank). *Win(t-1)* is one if a brokerage has at least one analysts above the corresponding cutoff in the same NF industry in the year before the award announcement. *Total asset* is the total asset of a brokerage at the end of the year before the award announcement (in billion CNY). *Net profit* is the net profit of a brokerage at the end of the year before the award announcement (in billion CNY). *Listed* is one if a brokerage is publicly listed. *Number of analysts* is the number of analysts employed by a brokerage. *Held by fund* is one if a brokerage has mutual funds as its shareholders. All variables are measured in year of the award announcement, unless noted otherwise. Regressions control for NF industry fixed effects in column (1) and year and NF industry fixed effects in columns (2) through (8). Standard errors in parentheses are clustered by NF industry in column (1) and by NF industry $\times$ year in columns (2) through (8). *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.5: Robustness check - Stock recommendation strategy
Winner cutoff

VARIABLES	(1) Within7	(2) Newly covered	(3) Market cap	(4) Under-valued	(5) P/E ratio	(6) Beta
Win(t)	-0.06077	-0.00300	0.97136	0.01133	2.04316	-0.01295
	(0.05124)	(0.00624)	(0.88331)	(0.01791)	(8.44354)	(0.01200)
Win(t)× Post			0.08015	-0.00090	4.19924	0.00311
			(0.41655)	(0.01480)	(6.56172)	(0.00749)
Post			-1.16314	-0.01283	-14.09577	0.00604
			(0.69429)	(0.02539)	(9.91485)	(0.01327)
Outcome mean	0.29318	0.00470	4.41855	0.40371	53.58179	1.11793
Observations	1,702	1,702	1,702	1,702	1,702	1,702
R-squared	0.133	0.051	0.643	0.697	0.153	0.754
Number of clusters	146	146	146	146	146	146
Bandwidth	0.080	0.080	0.080	0.080	0.080	0.080

Notes: The data are at the analyst × year level in columns (1) and (2) and at the analyst × year × period level in columns (3) through (6). $Within7$ equals 1 if an analyst makes any recommendations within 1-7 days before the announcement. *Newly covered* equals 1 if the stock recommended within 1-7 days before the announcement was not recommended by the analyst before. *Market cap* is stock's 3rd-quarter market capitalization measured in billion CNY. *Under-valued* equals 1 if stock's 3rd-quarter book-to-market ratio ≥ 1. *P/E ratio* is stock's 3rd-quarter price-to-earnings ratio. *Beta* is stock's risk factor estimated from CAPM regression using daily data in the past 250 trading days starting from October 31^{st}. *Return* is stock's return with cash dividend reinvestment in October. *post* equals 1 if the stock is recommended within 1-7 days before the announcement, or 0 if 8-60 days before the announcement. *Win(t)* equals 1 if the analyst is above the winner cutoff. All regressions control for year, brokerage house, NF industry fixed effects and are estimated using a linear RD model and triangular weights. Standard errors in parentheses are clustered by NF industry × year. *** p<0.01, ** p<0.05, * p<0.1.

Table A.6: Robustness check – BHAR based on Fama-French five-factor model

VARIABLES	(1) 2-d BHAR	(2) 2-d BHAR
Panel A: Winner, 1st trading day after award		
Win(t)	0.00722**	0.00709**
	(0.00330)	(0.00347)
Baseline mean	0.00492	0.00507
Observations	1,535	1,329
R-squared	0.207	0.221
Number of clusters	146	145
Bandwidth	0.080	0.069 (IK)
Panel B: Finalist, Avg. in notification week		
Win(t)	0.00517*	0.00413*
	(0.00281)	(0.00222)
Baseline mean	0.00469	0.00453
Observations	1,088	1,614
R-squared	0.250	0.203
Number of clusters	143	143
Bandwidth	0.035	0.043 (IK)
Panel C: Finalist, 1st trading day after award		
Win(t)	-0.01032***	-0.00961***
	(0.00375)	(0.00284)
Baseline mean	0.00469	0.00453
Observations	1,088	1,614
R-squared	0.237	0.194
Number of clusters	143	145
Bandwidth	0.035	0.043 (IK)

Notes: This table reports the main results estimated using the adjusted 2-day BHAR calculated from a Fama-French five-factor model as the outcome. Panel A corresponds to column (3) in Table 4, panel B corresponds to panel A column (3) in Table 8, and panel C corresponds to panel B column (3) in Table 8. Column (1) uses the same bandwidth as the main regressions while column (2) uses the IK bandwidth calculated using the new BHAR. Standard errors in parentheses are clustered by NF industry $\times$ year. *** p<0.01, ** p<0.05, * p<0.1.

Table A.7: Ability signaling - Probability of issuing new recommendations
Winner cutoff

VARIABLES	(1) 6-35 days	(2) 36-65 days	(3) 66-95 days
Panel A: All winners			
Win(t)	-0.05622	-0.03904	-0.01506
	(0.05079)	(0.04861)	(0.04604)
Outcome mean	0.55005	0.62779	0.63545
Observations	1,698	1,698	1,698
R-squared	0.108	0.134	0.149
Number of clusters	146	146	146
Bandwidth	0.080	0.080	0.080
Panel B: Repeated winners			
Win(t)	0.00707	0.04710	0.06539
	(0.05835)	(0.05049)	(0.05166)
Outcome mean	0.55081	0.63255	0.64036
Observations	1,407	1,407	1,407
R-squared	0.110	0.131	0.153
Number of clusters	144	144	144
Bandwidth	0.080	0.080	0.08

Notes: The data are at the analyst $\times$ year level. Panel A includes recommendations from all winners and non-winners whose margin of vote share is within 0.08 from the winner cutoff. Panel B restrict winners' recommendations to those from repeated winners. The dependent variable is an indicator of whether an analyst issues any new recommendations within 6-20, 6-35, 36-65, and 66-95 days after the award announcement. *Win(t)* equals 1 if the analyst is above the winner cutoff in the year of the award announcement. All regressions control for year, NF industry, and brokerage fixed effects, and are estimated using a linear RD model and triangular weights. Standard errors in parentheses are clustered by NF industry $\times$ year. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.8: BHAR reversal – Winner cutoff

VARIABLES	(1) 5-day	(2) 10-day	(3) 15-day	(4) 20-day	(5) 25-day	(6) 30-day
Panel A: All stocks						
Win(t)	0.01297***	0.01130***	0.00568	0.00320	0.00191	0.00228
	(0.00283)	(0.00399)	(0.00524)	(0.00548)	(0.00574)	(0.00552)
Observations	1,535	1,535	1,535	1,535	1,535	1,535
R-squared	0.189	0.197	0.216	0.196	0.188	0.200
Number of clusters	146	146	146	146	146	146
IK bandwidth	0.080	0.080	0.080	0.080	0.080	0.080
Panel B: No firm news						
Win(t)	0.01199***	0.01280***	0.00681	0.00381	0.00066	0.00107
	(0.00325)	(0.00419)	(0.00583)	(0.00621)	(0.00682)	(0.00693)
Observations	1,177	1,177	1,177	1,177	1,177	1,177
R-squared	0.192	0.215	0.205	0.195	0.193	0.216
Number of clusters	146	146	146	146	146	146
IK bandwidth	0.080	0.080	0.080	0.080	0.080	0.080

Notes: The data are at the analyst × stock × date level. Panel A consists of the last stock recommended within 1-30 days before the award announcement by each analyst whose margin of vote share is within 0.08 from the winner cutoff. Panel B excludes the stocks have any earnings announcements within the 30 days after the award announcement. Dependent variables are BHARs of varying length starting from the first trading day after the award announcement. *Win(t)* equals 1 if an analyst is above the winner cutoff in the year of award announcement. Specifications mirror the one in Table 4 column (3). Standard errors in parentheses are clustered by NF industry × year. *** p<0.01, ** p<0.05, * p<0.1.

Table A.9: Robustness check – Alternative running variables

VARIABLES	(1) Rank 2-d BHAR	(2) Scores 2-d BHAR
Panel A: Winner, 1st trading day after award		
Win(t)	0.00493*	0.00745**
	(0.00282)	(0.00333)
Baseline mean	0.00474	0.00474
Observations	1,117	1,101
R-squared	0.173	0.225
Number of clusters	146	145
Bandwidth	4.54	1719.12
Panel B: Finalist, Avg. in notification week		
Win(t)	0.00509*	0.00514*
	(0.00272)	(0.00278)
Baseline mean	0.00424	0.00424
Observations	1,117	1,101
R-squared	0.197	0.198
Number of clusters	146	145
IK bandwidth	3.88	1242.487
Panel C: Finalist, 1st trading day after award		
Win(t)	-0.01013**	-0.00978***
	(0.00411)	(0.00296)
Baseline mean	0.00424	0.00424
Observations	1,117	1,101
R-squared	0.184	0.212
Number of clusters	146	145
IK bandwidth	3.88	1242.487

Notes: This table reports the main results estimated using rank (column (1)) and raw scores (column (2)) as running variables. Panel A corresponds to column (3) in Table 4, panel B corresponds to panel A column (3) in Table 8, and panel C corresponds to panel B column (3) in Table 8. Standard errors in parentheses are clustered by NF industry × year. *** p<0.01, ** p<0.05, * p<0.1.

Table A.10: Robustness check - Alternative inference methods

VARIABLES	(1) Winner 1^{st} day after award 2-d BHAR	(2) Finalist Avg. in notif week 2-d BHAR	(3) Finalist 1^{st} day after award 2-d BHAR
Win(t)	0.00659	0.00432	-0.01061
-Brokerage clusters	(0.00227)***	(0.00238)*	(0.00354)***
-NF industry clusters	(0.00272)**	(0.00254)*	(0.00400)**
-Stock industry clusters	(0.00285)**	(0.00207)**	(0.00388)***

Notes: This table reports the main results under various inference methods. The coefficients on $Win(t)$ in columns (1), (2), and (3) correspond to those in Table 4 column (3), Table 8 panel A column (3), and Table 8 panel B column (3), respectively. Standard errors under various inference methods are reported in parenthesises. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.11: Robustness check – Local quadratic regression

	(1)	(2)	(3)
	Winner cutoff	Finalist cutoff	
VARIABLES	1^{st} trading day after announcement 2-d BHAR	Average in notification week 2-d BHAR	1^{st} trading day after announcement 2-d BHAR
Win(t)	0.00548*	0.00778**	-0.00964**
	(0.00319)	(0.00351)	(0.00465)
Baseline mean	0.00474	0.00424	0.00424
Observations	1,535	1,088	1,088
R-squared	0.184	0.233	0.228
Number of clusters	146	143	143
Bandwidth	0.080	0.035	0.035

Notes: This table reports the main results estimated using local quadratic regression. The coefficients on $Win(t)$ in columns (1), (2), and (3) correspond to those in Table 4 column (3), Table 8 panel A column (3), and Table 8 panel B column (3), respectively. Standard errors in parentheses are clustered by NF industry $\times$ year. *** p<0.01, ** p<0.05, * p<0.1.

Table A.12: Robustness check – Samples based on alternative number of last stocks

VARIABLES	(1) Last 3 stocks 2-d BHAR	(2) Last 5 stocks 2-d BHAR	(3) All stocks 2-d BHAR
Panel A: Winner, 1st trading day after award			
Win(t)	0.00339*	0.00240	0.00251
	(0.00198)	(0.00166)	(0.00161)
Baseline mean	0.00474	0.00474	0.00474
Observations	3,559	4,711	6,284
R-squared	0.132	0.105	0.099
Number of clusters	146	146	146
Bandwidth	0.080	0.080	0.080
Panel B: Finalist, Avg. in notification week			
Win(t)	0.00108	-0.00013	0.00009
	(0.00139)	(0.00128)	(0.00109)
Baseline mean	0.00424	0.00424	0.00424
Observations	2,564	3,560	5,095
R-squared	0.127	0.107	0.093
Number of clusters	143	144	144
Bandwidth	0.035	0.035	0.035
Panel C: Finalist, 1st trading day after award			
Win(t)	-0.00581**	-0.00503**	-0.00424*
	(0.00271)	(0.00236)	(0.00210)
Baseline mean	0.00424	0.00424	0.00424
Observations	2,564	3,560	5,095
R-squared	0.143	0.130	0.095
Number of clusters	143	144	144
Bandwidth	0.035	0.035	0.035

Notes: This table reports the main results estimated using alternative samples. The samples in column (1), (2), and (3) consist of the last three, five or all stocks recommended by each analyst within 1-30 days before the award announcement or notification week. Results in panels A, B, and C correspond to those in Table 4 column (3), Table 8 panel A column (3), and Table 8 panel B column (3), respectively. Standard errors in parentheses are clustered by NF industry × year. *** p<0.01, ** p<0.05, * p<0.1.

Table A.13: Robustness check – Samples based on alternative pre-award day range

VARIABLES	(1) 31-60d 2-d BHAR	(2) 61-90d 2-d BHAR	(3) 91-120d 2-d BHAR	(4) 121-150d 2-d BHAR
Panel A: Winner, 1st trading day after award				
Win(t)	-0.00204	0.00300	0.00040	0.00193
	(0.00276)	(0.00243)	(0.00207)	(0.00257)
Baseline mean	0.00490	0.00487	0.00485	0.00501
Observations	3,262	1,608	5,273	1,783
R-squared	0.117	0.181	0.089	0.124
Number of clusters	139	140	145	143
Bandwidth	0.080	0.080	0.080	0.080
Panel B: Finalist, Avg. in notification week				
Win(t)	-0.00071	-0.00203	-0.00382	0.00395
	(0.00262)	(0.00274)	(0.00300)	(0.00682)
Baseline mean	0.00447	0.00486	0.00464	0.00517
Observations	567	842	964	207
R-squared	0.326	0.260	0.222	0.700
Number of clusters	119	132	132	86
Bandwidth	0.035	0.035	0.035	0.035
Panel C: Finalist, 1st trading day after award				
Win(t)	0.00671	-0.00148	-0.00412	-0.00385
	(0.00432)	(0.00315)	(0.00406)	(0.00863)
Baseline mean	0.00447	0.00486	0.00464	0.00517
Observations	567	842	964	207
R-squared	0.364	0.244	0.235	0.763
Number of clusters	119	132	132	86

Notes: This table reports the main results estimated using alternative samples. The samples in column (1), (2), (3), and (4) consist of the last stock recommended by each analyst within 31-60 days, 61-90 days, 91-120 days and 121-150 days before the award announcement or notification week. Results in panels A, B, and C correspond to those in Table 4 column (3), Table 8 panel A column (3), and Table 8 panel B column (3), respectively. Standard errors in parentheses are clustered by NF industry × year. *** p<0.01, ** p<0.05, * p<0.1.

Table A.14: Robustness check – Various sample selection
Winner cutoff

VARIABLES	(1) Drop >1 analysts 2-d BHAR	(2) Keep top rank 2-d BHAR	(3) Drop Sat-Tues 2-d BHAR	(4) Drop firm news 2-d BHAR	(5) Drop notification 2-d BHAR
Win(t)	0.01282***	0.00784***	0.00705***	0.00653*	0.00529*
	(0.00404)	(0.00364)	(0.00285)	(0.00340)	(0.00312)
Baseline mean	0.0048	0.00467	0.00456	0.00469	0.00422
Observations	1,227	1,381	1,347	1,140	1,223
R-squared	0.282	0.203	0.199	0.178	0.215
Number of clusters	143	146	146	146	146
Bandwidth	0.08	0.08	0.08	0.08	0.08

Notes: This table reports the main results estimated using alternative samples. These robustness checks are for results in Table 4 column (3). *Drop by>1 analysts* excludes the stocks from the main RD sample if they are coded as ”last stock” by more than one analysts. *Keep top rank* only keeps the analysts with the best rank for stocks which are coded as ”last stock” by more than one analysts. *Drop Sat-Tues* excludes the stocks if they recommended by any analysts between the Saturday and the Tuesday after the award announcement. *Drop firm news* excludes the stocks if the firms issue any firm news between the Friday in the notification week and the Tuesday after the award announcement. *Drop notification* excludes the stocks recommended by the analysts in the notification week. Standard errors in parentheses are clustered by NF industry × year. *** p<0.01, ** p<0.05, * p<0.1.

Table A.15: Robustness check – Various sample restrictions
Finalist cutoff

VARIABLES	(1) Drop >1 analysts 2-d BHAR	(2) Keep top rank 2-d BHAR	(3) Drop Sat-Tues 2-d BHAR	(4) Drop firm news 2-d BHAR
Panel A: Avg. in notification week				
Win(t)	0.00494*	0.00561*	0.00706**	0.00406
	(0.00241)	(0.00320)	(0.00290)	(0.00315)
Baseline mean	0.00443	0.00437	0.00410	0.00406
Observations	870	979	977	815
R-squared	0.261	0.234	0.253	0.236
Number of clusters	141	143	141	139
Panel B: 1st trading day after award				
Win(t)	-0.01121***	-0.01132***	-0.01128***	-0.00833**
	(0.00386)	(0.00321)	(0.00338)	(0.00379)
Baseline mean	0.00443	0.00437	0.00410	0.00424
Observations	870	979	977	815
R-squared	0.282	0.236	0.236	0.270
Number of clusters	141	143	139	139
Bandwidth	0.035	0.035	0.035	0.035

Notes: This table reports the main results estimated using alternative samples. The robustness checks in panels A and B are for the results in Table 8 column (3) panels A and B, respectively. *Drop by>1 analysts* excludes the stocks from the main RD sample if they are coded as "last stock" by more than one analysts. *Keep top rank* only keeps the analysts with the best rank for stocks which are coded as "last stock" by more than one analysts. *Drop Sat-Tues* excludes the stocks if they recommended by any analysts between the Saturday and the Tuesday before the notification week. *Drop firm news* excludes the stocks if the firms issue any firm news between the Friday in the notification week and the Tuesday after the award announcement. Standard errors in parentheses are clustered by NF industry × year. *** p<0.01, ** p<0.05, * p<0.1.

Table A.16: Cumulative abnormal turnover

	(1)	(2)	(3)
	Winner cutoff	Finalist cutoff	
VARIABLES	1^{st} trading day after announcement 2-d CAT	Average in notification week 2-d CAT	1^{st} trading day after announcement 2-d CAT
Win(t)	0.30177*	0.13437	-0.11242
	(0.16505)	(0.14839)	(0.14910)
Baseline mean	0.39656	0.37357	0.37357
Observations	833	1,035	1,033
R-squared	0.369	0.417	0.405
Year FE	Yes	Yes	Yes
Brokerage FE	Yes	Yes	Yes
NF Industry FE	Yes	Yes	Yes
Stock Industry FE	Yes	Yes	Yes
Number of clusters	146	143	143
IK Bandwidth	0.045	0.035	0.035

Notes: The data are at the analyst × stock × date level consisting of the last stock recommended within 1-30 days before the announcement by an analyst whose margin of vote share is within 0.045 from the winner cutoff (column (1)) and 0.035 from the finalist cutoff (columns (2) and (3)). The dependent variable is two-day cumulative abnormal turnover measured in percentage point. Baseline mean is the average two-day CAT for stock recommendations from the analysts in the year before the award announcement. *Win(t)* equals 1 if the analyst is above the cutoff in the year of the award announcement. All regressions are estimated using a linear RD model and triangular weights. Standard errors in parentheses are clustered by NF industry × year. *** p<0.01, ** p<0.05, * p<0.1.

Table A.17: Placebo test - Post-award effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Market reaction	Earnings forecast			Resource allocation		
VARIABLES	2-d BHAR	Error	Bias	Boldness	Teammate error	Size	Lead author
Panel A: Winner Cutoff							
Win(t)	-0.00070	0.00046	0.00059	0.00031	0.00179	0.09160	0.04340
	(0.00062)	(0.00131)	(0.00172)	(0.00185)	(0.00089)	(0.12162)	(0.04078)
Outcome mean	0.00425	0.03115	0.02150	0.01757	0.02979	1.88767	0.32058
Observations	32,643	10,994	10,994	10,994	28,070	28,166	28,166
R-squared	0.035	0.325	0.305	0.256	0.648	0.467	0.135
Number of clusters	146	140	140	140	140	145	145
Bandwidth	0.042	0.020	0.020	0.020	0.020	0.020	0.020
Panel B: Finalist Cutoff							
Win(t)	0.00015	0.00010	0.00018	0.00127	-0.00109	0.10692	0.03362
	(0.00060)	(0.00176)	(0.00195)	(0.00094)	(0.00288)	(0.09675)	(0.04136)
Outcome mean	0.00411	0.03173	0.02168	0.01825	0.03011	1.77869	0.29060
Observations	41,489	11,076	11,076	10,649	28,164	28,293	28,293
R-squared	0.396	0.327	0.306	0.231	0.555	0.457	0.117
Number of clusters	146	140	140	140	140	146	146
Bandwidth	0.045	0.018	0.018	0.018	0.018	0.018	0.018

Notes: This table reports the placebo tests for the post-award effects. Samples, dependent variables and specifications correspond to the ones in Table 9. The only difference is that the variables are now measured in the year *before* the award announcement. *Win(t)* equals 1 if the analyst is above the winner cutoff (panel A) or the finalist cutoff (panel B) in the year of the award announcement. Standard errors in parentheses are clustered by NF industry × year. *** p<0.01, ** p<0.05, * p<0.1.

Appendix B

Chapter 2 Appendix

B.1 Additional Figures

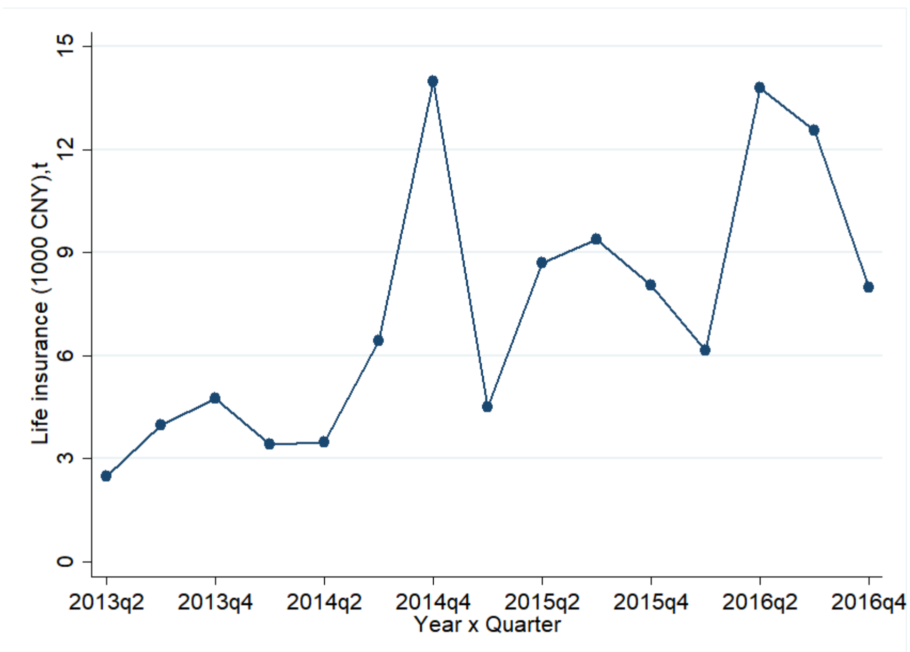
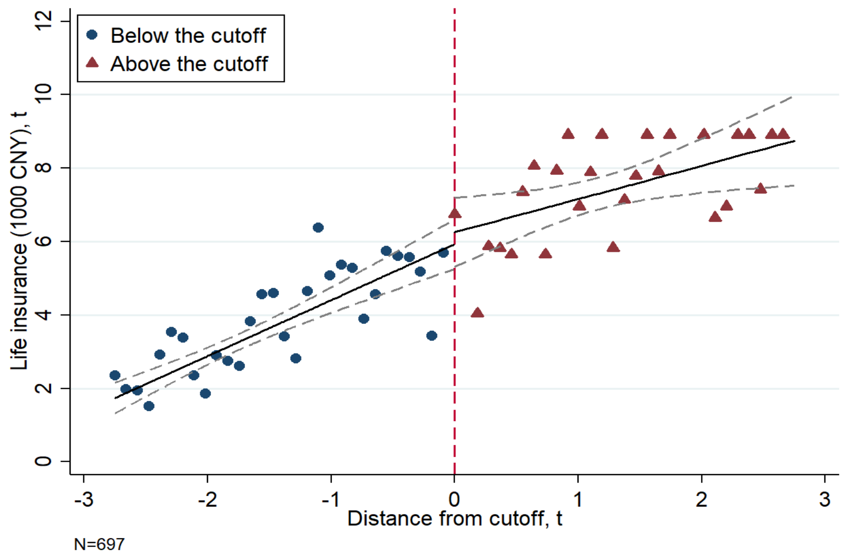
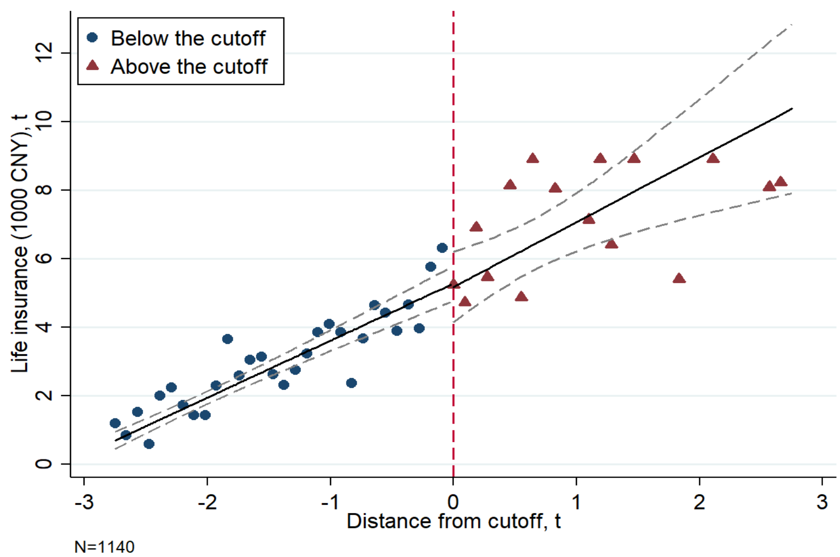


Figure B.1: "Best Rookie Award" cutoff commission by quarter

Notes: This figure plots the cutoff life insurance commission (in 1,000 CNY) for each quarterly "Best Rookie" award during the sample period, i.e., the life insurance commission value at rank tenth among all rookies in each quarter each year.



A: Rank=1-3



B: Rank=4+

Figure B.2: Placebo test for peer sabotage - Performance in quarter t by within-team rank in quarter t

Notes: Each observation is the average first quarter life insurance commission of rookies (main RD sample) who rank 1st to 3rd (panel A) and 4th and worse (panel B) within a team in their first quarter in a 0.09 bin based on their standardized first quarter life insurance commission. Dashed vertical lines denote the 10th rank standardized first quarter life insurance commission in each quarter (normalized to 0). The solid lines are estimated using a linear regression based on individual-level data using triangular weights. The dashed lines denote the 95% confidence interval based on the heteroscedasticity-consistent standard errors.

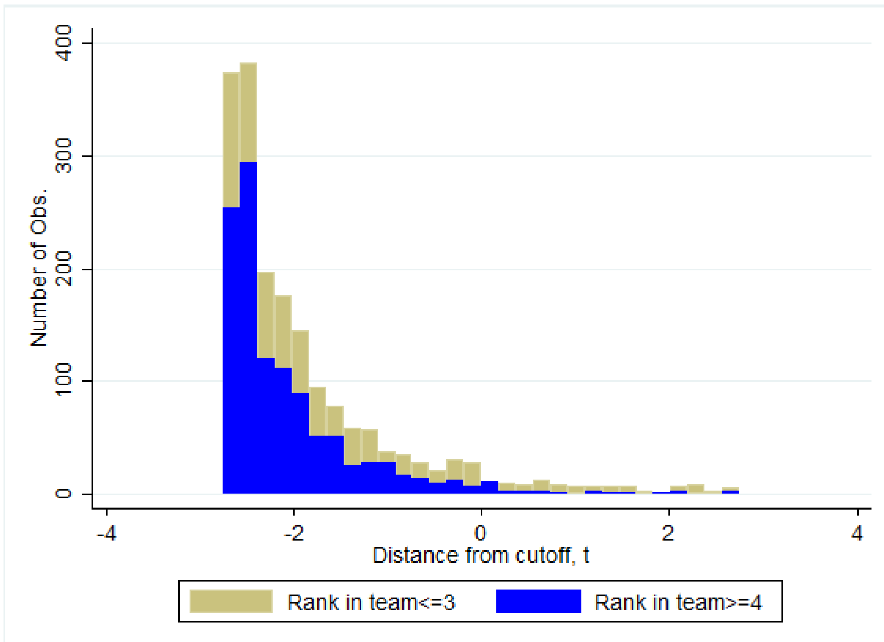
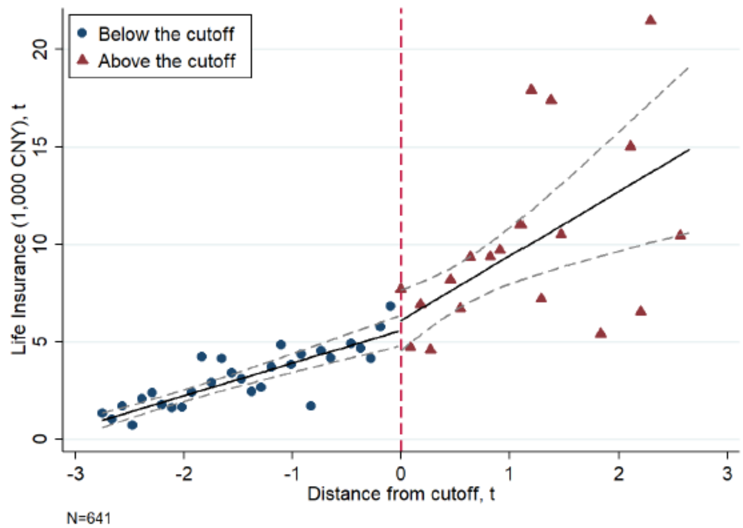
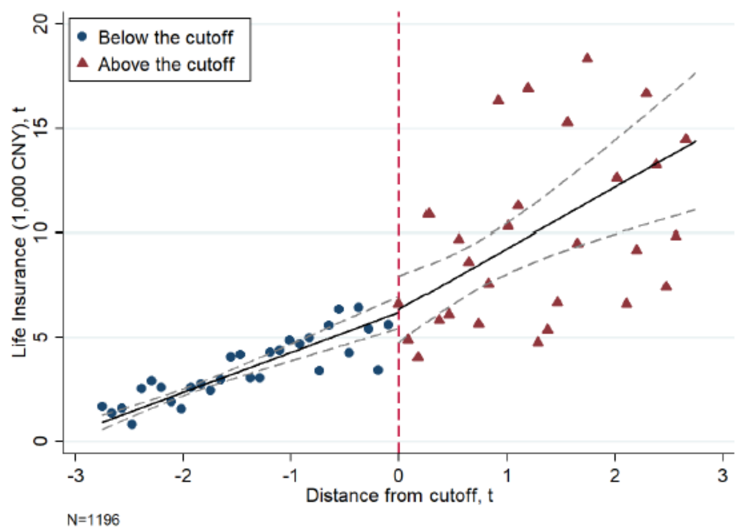


Figure B.3: Distribution of running variable by within-team rank in quarter t

Notes: This figure plots the distribution of standardized first quarter life insurance commission (the award threshold is normalized to 0) among rookies in the main RD sample by their within-team rank in quarter t. The observations are grouped into 30 bins (bin width=0.18). The height of the yellow segment represents the number of observations whose within-team rank is ≤ 3 and the height of the blue segment represents the number of observations whose within-team rank is ≥ 4.



A: With Competitors in teams



B: Without Competitors in teams

Figure B.4: Placebo test for peer sabotage - Performance in quarter t by level of competition in quarter t+1

Notes: Each observation is the average first quarter life insurance commission of rookies (main RD sample) who have at least one competitor (panel A) and no competitors (panel B) within a team in the quarter t+1 in a 0.09 bin based on their standardized first quarter life insurance commission. "Competitors" are qualified senior teammates who are at job level 3 and have the same number of referrals as the competing rookie in the beginning of quarter t+1 (either zero or one referral). Dashed vertical lines denote the 10th rank standardized first quarter life insurance commission in each quarter (normalized to 0). The solid lines are estimated using a linear regression based on individual-level data using triangular weights. The dashed lines denote the 95% confidence interval based on the heteroscedasticity-consistent standard errors.

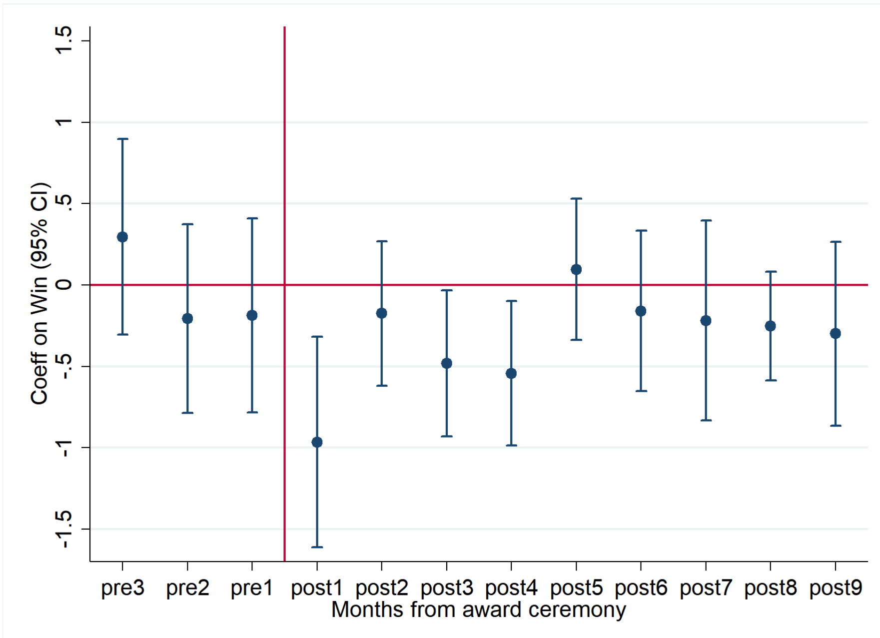
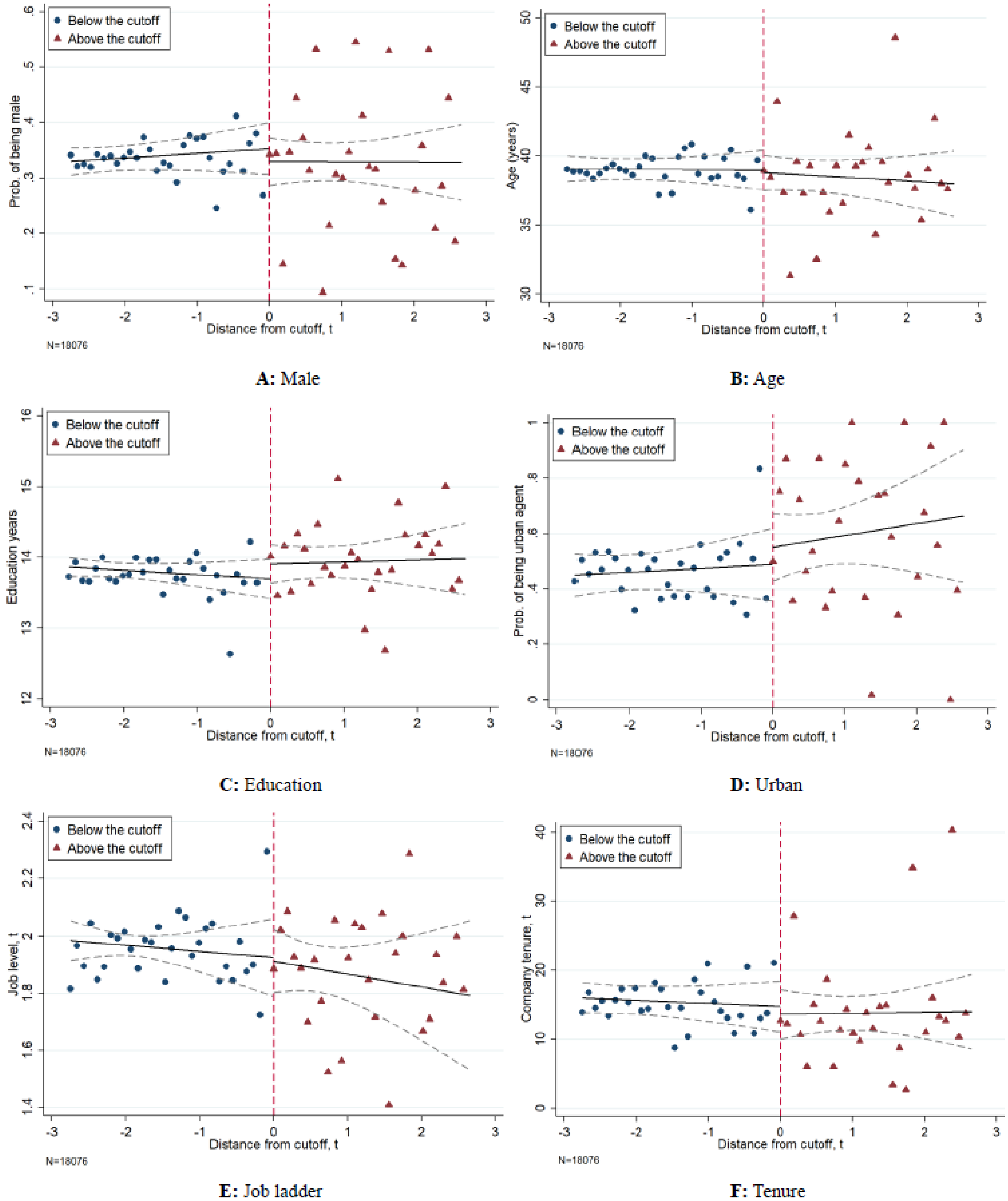


Figure B.5: Monthly dynamics of life insurance commission

Notes: This figure plots coefficients and 95% confidence interval for *Win* dummy in RD regressions using rookies' monthly life insurance commission from three months before the award ceremony to nine months after the award ceremony as outcomes. X-axis denotes the number of months after the award ceremony. For instance, "pre1" and "post1" represent one month before and after the award ceremony, respectively. Specifications mirror the one in Table 3 column (3). The vertical red line refers to the timing of the award ceremony.



A: Male

B: Age

C: Education

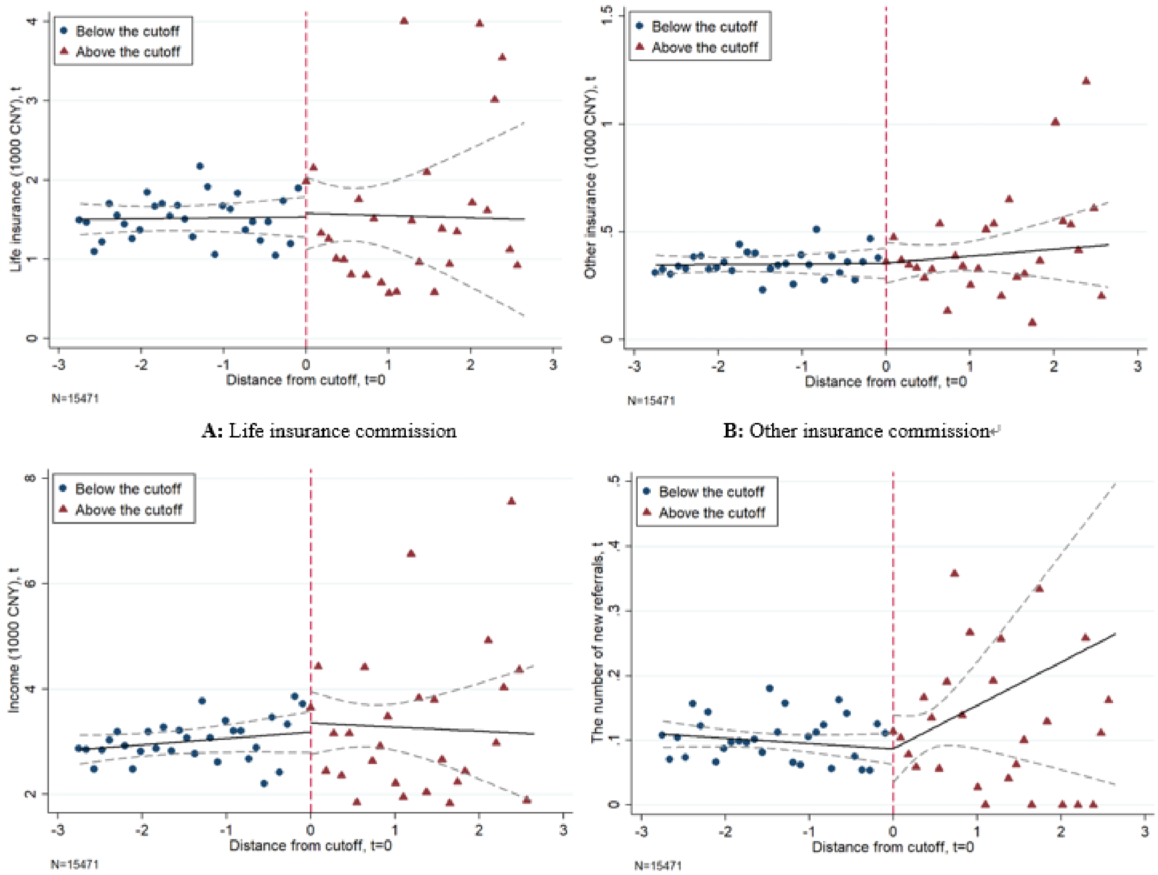
D: Urban

E: Job ladder

F: Tenure

Figure B.6: Validity of RD for spillover effects - Teammates' characteristics quarter t

Notes: Each observation is the average characteristics of participants' non-rookie *teammates* in a 0.09 bin based on *participants'* standardized first quarter life insurance. Variables are described in the notes to Table 1. Dashed vertical lines denote the 10th standardized first quarter life insurance commission in each quarter (normalized to 0). The solid lines are estimated using a linear regression based on individual-level data using triangular weights. The dashed lines denote the 95% confidence interval based on the standard errors clustered by team.



A: Life insurance commission

B: Other insurance commission

C: Income

D: The number of referrals

Figure B.7: Placebo test for spillover effects - Teammates' performance in quarter t

Notes: Each observation is the average first quarter performance of participants' non-rookie *teammates* in a 0.09 bin based on *participants'* standardized first quarter life insurance. Variables are described in the notes to Table 1. Dashed vertical lines denote the 10th standardized first quarter life insurance commission in each quarter (normalized to 0). The solid lines are estimated using a linear regression based on individual-level data using triangular weights. The dashed lines denote the 95% confidence interval based on the standard errors clustered by team.

B.2 Additional Tables

Table B.1: Robustness check - Main results under Heckman selection

VARIABLES	(1) Rookie Sample	(2) Peer Sample
Win	-1.720***	-0.063
	(0.0651)	(0.177)
Observations	1,851	17,409
First Stage Wald Chi-square	184.21	938.03
First Stage Prob > Chi-square	0.00	0.00

Notes: The samples in columns (1)-(2) consists of rookies and their peer, respectively. The dependent variable is the life insurance commission earned in the quarter following an award designation (measured in 1,000 CNY). We use each salesperson's contract start day of the month as the exogenous predictor for the first stage of Heckman selection model. Heteroscedasticity-consistent standard errors are reported in parentheses in column (1). Standard errors in parentheses in column (2) are clustered by team. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.2: Distribution of rank and observations by year×quarter

Year×Quarter	**Total obs.**	**Best rank**	**Worst rank**
2013×Q1	268	3	120
2013×Q2	116	2	117
2013×Q3	117	2	118
2013×Q4	171	4	176
2014×Q1	235	3	166
2014×Q2	100	2	96
2014×Q3	57	2	58
2014×Q4	150	5	154
2015×Q1	304	5	309
2015×Q2	24	5	28
2015×Q3	82	3	84
2015×Q4	101	3	104
2016×Q1	29	2	31
2016×Q2	50	4	53
2016×Q3	33	6	38

Notes: This table presents the sample distribution for the main RD sample. $Total\ obs.$ is the total number of observations in a given year-quarter. *Best rank* and *Worst rank* refer to the rank of the best rookie and the worst rookie in the main RD sample each year and quarter. In principle, $Total\ obs. =$ *Worst rank* $-$ *Best rank* $+1$. But this equation does not hold when there are tied ranks. For instance, in $2013 \times Q1$ there are 153 tied-ranks in the rank 120th among rookies.

Table B.3: Summary statistics - Full sample

Variable	N	Mean	Std. Dev.	Min.	Max.
Panel A: Full Rookie Sample					
Life insurance commission (t+1)	10,996	1.12	2.35	0.00	14.88
Other insurance commission (t+1)	10,996	0.15	0.30	0.00	1.85
Number of referrals (t+1)	10,996	0.11	0.58	0.00	28.00
Income (t+1)	10,996	2.48	3.66	0.00	20.20
Exit (t+1)	10,996	0.09	0.29	0.00	1.00
Life insurance commission (t)	10,996	1.03	1.51	0.00	9.01
Other insurance commission (t)	10,996	0.15	0.30	0.00	1.85
Number of referrals (t)	10,996	0.04	0.51	0.00	46.00
Income (t)	10,996	1.87	2.38	0.00	12.06
Male	10,996	0.36	0.48	0.00	1.00
Age	10,996	34.34	7.81	18.00	57.00
Education	10,996	14.26	1.29	9.00	21.00
Urban	10,996	0.48	0.50	0.00	1.00
Duration (t)	10,996	29.42	18.36	1.00	64.00
Panel B: Peer Sample					
Life insurance commission (t+1)	44,254	1.24	3.94	0.00	37.59
Other insurance commission (t+1)	44,254	0.30	0.61	0.00	4.52
Number of referrals (t+1)	44,254	0.07	0.41	0.00	14.00
Income (t+1)	44,254	2.69	6.10	0.00	64.56
Promotion (t+1)	44,254	0.31	0.46	0.00	1.00
Exit (t+1)	44,254	0.11	0.32	0.00	1.00
Life insurance commission (t)	44,254	1.45	4.10	0.00	37.84
Other insurance commission (t)	44,254	0.32	0.59	0.00	3.46
Number of referrals (t)	44,254	0.11	0.52	0.00	29.00
Income (t)	44,254	2.82	4.77	0.00	28.39
Promotion (t)	44,254	0.17	0.38	0.00	1.00
Male	44,254	0.32	0.47	0.00	1.00
Age	44,254	39.20	9.67	18.00	75.00
Education	44,254	13.91	1.55	9.00	19.00
Urban	44,254	0.48	0.50	0.00	1.00
job level (t)	44,254	2.05	0.77	1.00	3.00
Tenure (t)	44,254	16.33	18.57	2.00	82.00

Notes: The full rookie sample is defined as all the rookies during our sample period. The peer sample is defined as the non-rookie teammates of the participants in the full rookie sample. Observation is at the salesperson $\times$ quarter level. All variables are described in the notes to Table 1.

Table B.4: Robustness check - Main result under different inference methods

VARIABLES	(1) Life insurance	(2) Life insurance	(3) Life insurance
Win	-1.857	-1.803	-1.720
-Heteroscedasticity-consistent	(0.696)***	(0.646)***	(0.655)***
-Team clusters	(0.688)***	(0.638)***	(0.649)***
-Year-quarter clusters (wild bootstrap)	(0.676)***	(0.650)***	(0.654)***
-Two-way clusters (team and yq)	(0.676)***	(0.650)***	(0.654)***
Observations	1,837	1,837	1,837
R-squared	0.078	0.214	0.229
Top 20 baseline mean	6.209	6.209	6.209

Notes: This table investigates the robustness of our inferences in the main results (Table 3 column (3)). The entries after row 1 present different levels of clustering for standard errors. Note that when clustering by year-and-quarter, we use *wild bootstrap* method (1,000 times) to obtain robust clustered standard errors (Cameron et al., 2008), because we only have 15 year-and-quarter cells which are too few to obtain correct inference. The dependent variable is the life insurance commission in quarter t+1. Specifications mirror the one reported in Table 3 column (3). $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.5: Placebo test - Performance in quarter t

VARIABLES	(1) Life insurance	(2) Other insurance	(3) Referral	(4) Income
Win	-0.079	-0.001	-0.061	-0.907
	(0.148)	(0.048)	(0.063)	(0.577)
Observations	1,837	1,837	1,837	1,837
R-squared	0.951	0.209	0.079	0.590
Top 20 baseline mean	6.209	0.405	0.491	6.693

Notes: The dependent variables in columns (1)-(4) are life insurance commission, other insurance commission, the number of referrals, and the total income in the first quarter of the rookies, respectively. Monetary values are measured in 1,000 CNY. Specifications mirror the one in Table 3 column (3). Heteroscedasticity-consistent standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.6: Robustness check - Main result under various rank restrictions

VARIABLES	(1) 1st≤Rank≤20th	(2) 5th≤Rank≤15th
Win	-1.917**	-1.950*
	(0.882)	(1.051)
Observations	269	164
R-squared	0.278	0.297
Baseline mean	7.295	7.617
No. of winners	115	90

Notes: The dependent variables in columns (1)-(2) is the life insurance commission in the quarter following an award designation. Specifications mirror the one reported in Table 3 column (3). Samples in columns (1)-(2) include top 20 rookies and 5th-15th rookies, respectively. Heteroscedasticity-consistent standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.7: Robustness check - Main result under various bandwidths

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Win	-2.058***	-2.022***	-1.938***	-1.720***	-1.609**	-1.544**	-1.495**
	(0.796)	(0.740)	(0.691)	(0.655)	(0.630)	(0.610)	(0.590)
Observations	671	918	1,383	1,837	2,507	3,169	3,755
R-squared	0.242	0.236	0.234	0.229	0.225	0.220	0.211
Baseline mean	6.039	6.103	6.158	6.209	6.219	6.255	6.283
Bandwidth	2	2.25	2.5	2.75	3	3.25	3.5
No. of winners	93	103	109	115	117	121	124

Notes: The dependent variables in columns (1)-(7) is the life insurance commission in the quarter following an award designation. Specifications mirror the one reported in Table 3 column (3). Columns (1)-(7) show the estimates with the bandwidth from 2 to 3.5. Note that column (4) is the same as column (3) in Table 3. Heteroscedasticity-consistent standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.8: Robustness check - Main result estimated with local quadratic

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Win	-1.825	-1.819*	-1.909*	-2.100**	-2.006**	-1.889**	-1.793**
	(1.146)	(1.051)	(0.974)	(0.922)	(0.880)	(0.839)	(0.809)
Observations	671	918	1,383	1,837	2,507	3,169	3,755
R-squared	0.243	0.236	0.233	0.229	0.225	0.221	0.211
Top 20 baseline mean	6.039	6.103	6.158	6.209	6.219	6.255	6.283
Bandwidth	2	2.25	2.5	2.75	3	3.25	3.5
No. of winners	93	103	109	115	117	121	124

Notes: The dependent variables in columns (1)-(7) is the life insurance commission in the quarter following an award designation. All specifications are local quadratic regressions with triangular weights. For this specification, the IK bandwidth is around 3. Columns (1)-(7) show the estimates with the bandwidth from 2 to 3.5. Heteroscedasticity-consistent standard errors in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.9: Strategic reallocation across salespeople - Teammates' monthly life insurance commission

VARIABLES	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4	+5	+6	+7
							Panel A: Referrers						
Win	-1.606	0.406	-0.356	1.298	0.225	-0.299	-1.122	1.392	-0.613	-0.571	-0.077	-1.508**	-0.680
	(1.355)	(1.121)	(0.788)	(2.522)	(1.047)	(0.623)	(1.322)	(0.885)	(0.855)	(1.437)	(0.739)	(0.756)	(1.072)
Observations	1,412	1,466	1,513	714	1,195	1,831	1,831	1,812	1,784	1,708	1,664	1,618	1,512
R-squared	0.269	0.210	0.195	0.299	0.189	0.173	0.326	0.161	0.178	0.160	0.136	0.175	0.203
Top 20 mean	7.065	3.453	3.903	7.065	3.453	3.903	6.95	3.389	4.014	4.851	3.833	3.658	4.872
							Panel B: Managers						
Win	0.633	-0.444	-1.411	0.129	1.325	-0.631	0.129	1.325	-0.631	-0.097	2.450	-0.621	-0.926
	(2.301)	(1.137)	(0.958)	(1.749)	(0.975)	(0.962)	(1.749)	(0.975)	(0.962)	(2.273)	(1.836)	(1.044)	(1.559)
Observations	1,479	1,479	1,480	1,729	1,712	1,684	1,729	1,712	1,684	1,611	1,572	1,527	1,425
R-squared	0.193	0.170	0.113	0.220	0.118	0.133	0.220	0.118	0.133	0.209	0.087	0.126	0.096
Top 20 mean	7.174	3.252	4.573	7.174	3.252	4.573	8.137	3.692	4.861	6.617	4.18	3.865	5.649
							Panel C: Senior teammates						
Win	0.097	0.012	-0.016	0.053	0.029	0.232	-0.022	0.195	-0.167	-0.090	0.137	-0.149	-0.117
	(0.285)	(0.176)	(0.181)	(0.234)	(0.154)	(0.175)	(0.229)	(0.174)	(0.167)	(0.258)	(0.210)	(0.217)	(0.239)
Observations	3,610	3,632	3,645	4,305	4,306	4,306	4,306	4,257	4,207	4,051	3,996	3,959	3,803
R-squared	0.056	0.068	0.050	0.111	0.057	0.054	0.079	0.071	0.060	0.089	0.053	0.048	0.063
Top 20 mean	1.357	0.956	1.058	1.339	0.736	0.813	1.089	0.829	0.866	1.215	0.874	0.934	1.286

Notes: The dependent variables are the monthly life insurance commission in the sixth month (-6) before the award to the ninth month after the award (+9) for rookies' referrers (panel A), managers (panel B), and senior teammates (panel C). Senior teammates are defined as those whose job level at t=0 were equal to 3. "Top 20 mean" is the outcome mean among the referrers, managers, or senior teammates of top 20 rookies (based on first quarter ranking). The number of observations decreases due to the exit of rookies, managers, referrers, or senior teammates from the company. Specifications mirror the one in Table 3 column (3) and also control for the job level of the referrers, managers, or senior teammates and an indicator of whether the referrer is the same as the manager. Heteroscedasticity-consistent standard errors are reported in parentheses. parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.10: Peer sabotage - Impact of award by within-team percentile in quarter t

	(1)	(2)	(3)
		Panel A	
VARIABLES	Percentile<Top 15%	Percentile≥Top 15%	Prob.>*chi2*
Win	-2.905	-0.904	0.106
	(0.894)	(0.882)	
Observations	1,040	797	
R-squared	0.213	0.263	
No. of winners	31	84	
Baseline mean	1.922	3.553	
		Panel B	
VARIABLES	Percentile<Top 10%	Percentile≥Top 10%	Prob.>*chi2*
Win	-2.187	-0.615	0.217
	(0.974)	(0.847)	
Observations	1,155	682	
R-squared	0.236	0.278	
No. of winners	43	72	
Baseline mean	2.177	3.956	
		Panel B	
VARIABLES	Percentile<Top 5%	Percentile≥Top 5%	Prob.>*chi2*
Win	-2.511	0.193	0.037
	(0.797)	(1.053)	
Observations	1,281	556	
R-squared	0.219	0.333	
No. of winners	66	49	
Baseline mean	2.464	3.011	

Notes: This table splits the main RD sample by a rookie's percentile of average monthly life insurance commission among her teammates' in the quarter before an award designation. Panels A-C split the sample by whether the rookie is above or below top 15%, 10%, or 5% in their teams, respectively. The dependent variable is the life insurance commission in the quarter following the award designation. "No. of winners" refers to the number of award winners in each subsample. Specifications in columns (1) and (2) mirror the one in Table 3 column (3). Column (3) reports the p-value for Chow test on null hypothesis that the coefficients in columns (1) and (2) are statistically indistinguishable. Heteroscedasticity-consistent standard errors are reported in parentheses.

Table B.11: Validity of RD - Rookies' characteristics in quarter t by rookies' within-team rank in quarter t

Within-team rank≤3	Within-team rank≥4									
VARIABLES	Male	Age	Edu	Urban	Duration	Male	Age	Edu	Urban	Duration
Win	0.205	0.071	-0.298	-0.114	-1.600	0.114	-2.557	-0.285	0.073	-0.180
	(0.129)	(2.141)	(0.320)	(0.137)	(6.707)	(0.104)	(1.637)	(0.286)	(0.112)	(4.347)
Observations	1,140	1,140	1,140	1,140	1,140	697	697	697	697	697
R-squared	0.060	0.083	0.121	0.036	0.207	0.032	0.092	0.172	0.043	0.140
Number of winners	30	30	30	30	30	85	85	85	85	85
Top 20 baseline mean	0.450	37.500	13.740	0.480	2.480	0.290	36.024	14.053	0.562	2.331

Notes: We split the main RD sample by rookies' within-team rank in their first quarter in the company. In column (1), the dependent variable is a dummy indicating whether a rookie's within-team rank is worse or equal to 4th place in quarter t. Columns (2)-(6) present the estimates in the low-rank sample and columns (6)-(10) present the estimates in high-rank sample. "No. of winners" refers to the number of award winners in each subsample. Specifications mirror the one in Table 3 column (3). Heteroscedasticity-consistent standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.12: Validity of RD - Teammates' characteristics and performance in quarter t by rookies' within-team rank in quarter t

	Panel A: Teammates' characteristics											
	Within-team rank$\geq$4						Within-team rank$\leq$3					
VARIABLES	Male	Age	Edu	Urban	Level	Tenure	Male	Age	Edu	Urban	Level	Tenure
Win	-0.043	-1.035	0.045	0.084	0.030	-0.825	-0.006	0.976	0.375	0.147	0.019	0.425
	(0.038)	(1.025)	(0.223)	(0.102)	(0.051)	(2.631)	(0.045)	(1.340)	(0.228)	(0.130)	(0.094)	(3.659)
Observations	11,205	11,205	11,205	11,205	11,205	11,205	4,346	4,346	4,346	4,346	4,346	4,346
R-squared	0.003	0.009	0.024	0.025	0.040	0.024	0.006	0.015	0.055	0.054	0.059	0.052
Num of winners	30	30	30	30	30	30	85	85	85	85	85	85
Baseline mean	0.333	40.563	13.737	0.483	2.134	19.962	0.317	37.419	13.814	0.503	1.991	11.559

	Panel B: Teammates' performance							
	Within-team rank$\geq$4				Within-team rank$\leq$3			
VARIABLES	Life insurance	Other insurance	Referral	Income	Life insurance	Other insurance	Referral	Income
Win	-0.835	0.058	-0.003	0.531	-0.222	0.028	-0.008	-0.210
	(1.089)	(0.063)	(0.028)	(0.419)	(0.220)	(0.041)	(0.018)	(0.426)
Observations	11,205	11,205	11,205	11,205	4,346	4,346	4,346	4,346
R-squared	0.123	0.303	0.024	0.276	0.300	0.378	0.050	0.328
Num of winners	30	30	30	30	85	85	85	85
Baseline mean	1.660	0.388	0.059	3.338	1.063	0.278	0.039	2.593

Notes: We split the main RD sample by rookies' within-team rank in their first quarter in the company. Panels A and B present balance- and placebo-test results, respectively, for all rookies' teammates. In panel A, columns (1)-(5) present the estimates in the low-rank sample and columns (6)-(11) present the estimates in high-rank sample. In panel B, columns (1)-(4) present the estimates in the low-rank sample and columns (5)-(8) present the estimates in high-rank sample. "Num of winners" refers to the number of award winners in each subsample. Specifications mirror the one in Table 3 column (3). Standard errors reported in parentheses are clustered by team. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.13: Validity of RD - Rookies' characteristics in quarter t by Level of competition in quarter t+1

	With Competitors					Without Competitors				
VARIABLES	Male	Age	Edu	Urban	Duration (t)	Male	Age	Edu	Urban	Duration (t)
Win	-0.015	-0.994	-0.310	0.016	2.270	0.198*	-2.216	-0.234	0.036	-1.817
	(0.133)	(2.164)	(0.336)	(0.137)	(7.189)	(0.113)	(1.675)	(0.270)	(0.111)	(4.449)
Observations	641	641	641	641	641	1,196	1,196	1,196	1,196	1,196
R-squared	0.075	0.076	0.133	0.050	0.175	0.029	0.088	0.153	0.031	0.172
Num of winners	40	40	40	40	40	75	75	75	75	75
Baseline mean	0.432	36.547	14.000	0.537	47.911	0.305	36.586	13.902	0.575	44.440

Notes: This table splits the main RD sample by whether rookies have at least one competitor who competes with them for internal resources in the quarter after an award designation. "Competitors" are qualified senior teammates who are at job level 3 and have the same number of referrals as the competing rookie in the beginning of quarter t+1 (either zero or one referral). In column (1), the dependent variable is a dummy indicating whether a rookie has at least one such competitor. "No. of winners" refers to the number of award winners in each subsample. Specifications mirror the one in Table 3 column (3). Heteroscedasticity-consistent standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.14: Peer sabotage - Impact of award by whether rookie is manager's direct referral

VARIABLES	(1) Not referral	(2) Direct referral	(3) Prob.¿*chi2*
Win	-2.060**	-1.247	0.524
	(0.888)	(0.894)	
Observations	887	950	
R-squared	0.263	0.251	
No. of winner	48	67	
Top 20 baseline mean	6.222	6.191	

Notes: "Direct referral" represents the sample in which a rookie's manager in quarter t directly referred rookie to the firm in quarter t-2; "Not referral" represents the remaining rookie sample. The dependent variable is the life insurance commission in the quarter following an award designation. "No. of Winners" refers to the number of award winners in each subsample. Specifications in columns (1) and (2) mirror the one in Table 3 column (3). Column (3) reports the p-value for Chow test on null hypothesis that the coefficients in columns (1) and (2) are statistically indistinguishable. Heteroscedasticity-consistent standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.15: Quarterly dynamics of main result - Life insurance commission in subsequent quarters

VARIABLES	(1) t	(2) t+1	(3) t+2	(4) t+3
Win	-0.079	-1.720***	-0.531	-0.737
	(0.148)	(0.655)	(0.616)	(0.714)
Observations	1,837	1,837	1,716	1,526
R-squared	0.951	0.229	0.137	0.121
Top 20 baseline mean	6.209	6.209	6.209	6.209

Notes: The dependent variables in columns (1)-(4) are the life insurance commission in the corresponding quarter. The number of observations decreases due to the exit of salespeople from the company. Specifications mirror the one in Table 3 column (3). Heteroscedasticity-consistent standard errors are reported in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix C

Chapter 3 Appendix

C.1 Additional Figures

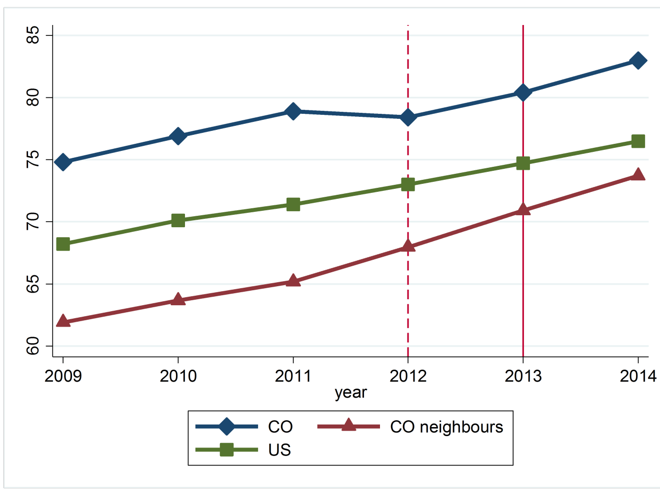
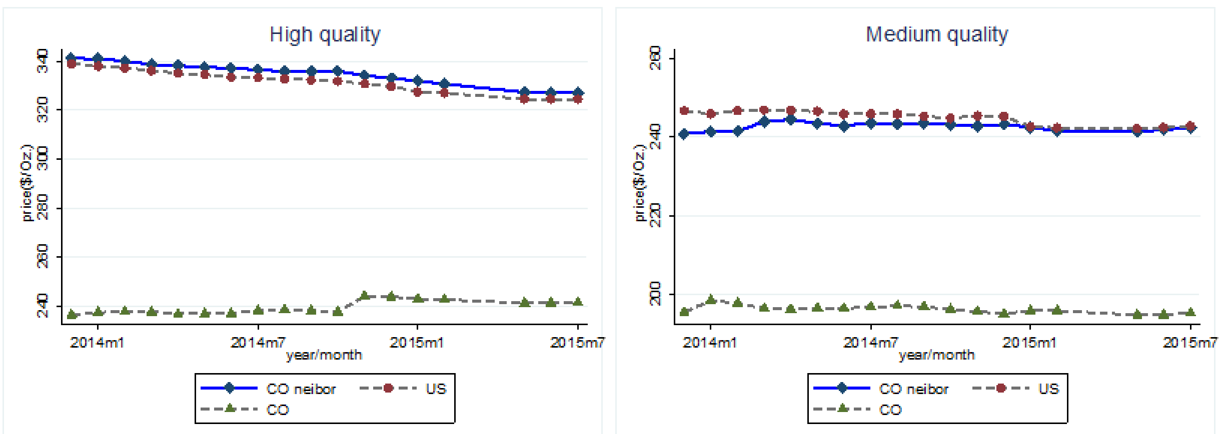


Figure C.1: Perception of marijuana among juveniles

Notes: The figure plots the two-year moving average of the percentage of "persons aged 12-17 who perceived no great risk from smoking marijuana once a month". Year t in the figure refers to the average of t-1 and t. Data is from State level Behavioral Health Barometer hosted by SAMSHA (https://www.samhsa.gov/data/behavioral-health-barometers). Dashed vertical line marks the time when Colorado passed RML (Dec 2012), and solid vertical line marks the time when the effect of RML is supposed to appear (i.e. one year after the passage of RML). "CO" is the state level marijuana perception in Colorado; "CO neighbours" is the average of state level perception in Utah, Kansas, Nebraska, Oklahoma, and Texas; "US" is the average of all states with BHB.



Panel A Panel B

Figure C.2: Price of marijuana

Notes: The figure plots the state-level month-end price of weed in the unit of dollar/ounce during 2013/12-2015/07. No data is available for 2015/03-2015/04. Price data is scraped by Frank Bi (retrieved June 16th 2017 from https://github.com/frankbi/price-of-weed) from an online anonymous price submission website (http://www.priceofweed.com). "CO" is the state level marijuana price; "CO neibor" is the average of state level marijuana price in Utah, Kansas, Nebraska, Oklahoma, and Texa"; "US" is the average of all NIBRS reporting states in the U.S.

C.2 Additional Tables

Table C.1: Supplement Result - Effect of RML on stand-alone adult male marijuana sale/manufacture arrests

VARIABLES	(1) Sale	(2) Manufacture	(3) Sale	(4) Manufacture
Distance to CO≤150mi*Post	14.293	16.168		
	(11.873)	(12.356)		
Distance to CO*Post			-3.417	-7.329
			(3.686)	(5.193)
Baseline mean	17.52	17.52	17.52	17.52
Observations	1,490	1,490	1,490	1,490
R-squared	0.031	0.094	0.027	0.094
Agency FE	Yes	Yes	Yes	Yes
Year FE	Yes	No	Yes	No
State*Year FE	No	Yes	No	Yes
Number of Clusters	148	148	148	148

Notes: Outcome is the number of marijuana sale and manufacture arrests per 100k residents covered by a NIBRS agency. Baseline mean is the mean of y during 2009-2012 across all agencies in the regression sample. Distance dummy is one if an agency is within 150mi from Colorado. Distances are measured in the unit of 100 miles. Post equals 1 for year 2013-2015. All specifications include the county unemployment rate and the county population ratio of black male aged 20+. Standard errors (in parentheses) are clustered at county level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table C.2: Event Study - Effect of RML on stand-alone adult male marijuana sale/manufacture arrests

VARIABLES	Distance to CO≤150mi Mj sale/100k	Distance to CO Mj sale/100k
Treat*2009	-4.032	7.652*
	(7.220)	(4.083)
Treat*2010	3.656	3.026
	(6.695)	(3.317)
Treat*2011	2.097	-0.576
	(5.840)	(3.122)
Treat*2013	6.921	-0.183
	(6.769)	(2.974)
Treat*2014	25.354	-8.588
	(17.359)	(7.234)
Treat*2015	17.911	-7.691
	(16.839)	(7.140)
Baseline mean		
Observations	1,490	1,490
R-squared	0.098	0.103
Agency FE	Yes	Yes
State*Year FE	Yes	Yes
Number of clusters	148	148

Notes: Year 2012 is the omitted year. Outcome is the number of marijuana sale and manufacture per 100k residents covered by an agency. Baseline mean is the mean of outcome during 2009-2012 across all agencies in the regression sample. Distance dummy is one if an agency is within 150mi from Colorado. Distances are measured in the unit of 100 miles. Post equals 1 for year 2013-2015. All specifications include the county unemployment rate and the county population ratio of black male aged 20+. Standard errors (in parentheses) are clustered at county level. *** p<0.01, ** p<0.05, * p<0.1.

Table C.3: Robustness check - Controlling for medical marijuana patients in CO counties

	Treatment: distance to CO≤150mi			Treatment: distance to CO		
VARIABLES	NIBRS offense	Stand-Alone arrest	UCR arrest	NIBRS offense	Stand-Alone arrest	UCR arrest
Treat*Post	72.308**	52.263**	22.997	-45.282***	-31.620***	-13.076*
	(31.347)	(22.619)	(15.002)	(13.758)	(10.430)	(7.044)
Patient/100k	-0.001	-0.001	-0.002	-0.004	-0.003	-0.003
	(0.015)	(0.013)	(0.009)	(0.015)	(0.012)	(0.008)
Baseline	226.05	163.54	115.70	226.05	163.54	115.70
Observations	1,490	1,490	1,468	1,490	1,490	1,468
R-squared	0.083	0.091	0.129	0.092	0.097	0.130
Agency FE	Yes	Yes	Yes	Yes	Yes	Yes
State*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of clusters	148	148	148	148	148	148

Notes: The number of medical marijuana patients is the number in December each year from https://www.colorado.gov/pacific/cdphe/medical-marijuana-statistics-and-data. Marijuana possession rates are the counts of marijuana possession per 100k residents covered by an agency. Baseline mean is the mean during 2009-2012 across all agencies in the main result sample. Distance dummy equals to 1 if an agency is within 150 miles from Colorado. Distances are measured in the unit of 100 miles. Post equals to 1 for year 2013-2015. All specifications include the unemployment rate and the population ratio of black male aged 20+ at county level. Standard errors (in parentheses) are clustered at county level. *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Table C.4: Robustness check - Various sample selections and model specifications

	Treatment: distance to CO≤150mi				Treatment: distance to CO			
VARIABLES	Mj poss/100k	Drop 0	Drop TX	Balanced	Mj poss/100k	Drop 0	Drop TX	Balanced
Panel A: NIBRS offense								
Treat*Post	48.083	81.184**	74.207**	62.595*	-19.906	-46.523***	-45.991***	-48.796***
	(31.618)	(36.709)	(32.417)	(32.040)	(16.390)	(15.256)	(14.078)	(14.248)
Baseline mean	226.05	226.05	226.05	226.05	226.05	226.05	226.05	226.05
Observations	1,490	1,385	1,471	1,134	1,490	1,385	1,471	1,134
R-squared	0.460	0.085	0.082	0.098	0.460	0.093	0.092	0.112
Panel B: Stand-Alone arrest								
Treat*Post	34.145	53.288*	54.262**	44.102*	-12.254	-31.045***	-32.328***	-33.876***
	(24.836)	(27.918)	(23.396)	(22.269)	(13.964)	(11.593)	(10.676)	(10.852)
Baseline mean	163.54	163.54	163.54	163.54	163.54	163.54	163.54	163.54
Observations	1,490	1,336	1,471	1,134	1,490	1,351	1,471	1,134
R-squared	0.474	0.093	0.090	0.109	0.474	0.098	0.096	0.118
Panel C: UCR arrest								
Treat*Post	14.906	26.473	22.802	17.785	1.586	-13.833	-12.953*	-15.408*
	(16.250)	(18.110)	(15.550)	(15.746)	(10.518)	(8.426)	(7.244)	(8.062)
Baseline mean	115.70	115.70	115.70	115.70	115.70	115.70	115.70	115.70
Observations	1,468	1,286	1,449	1,099	1,468	1,286	1,449	1,099
R-squared	0.471	0.136	0.127	0.160	0.471	0.137	0.129	0.164
Agency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State*Year FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Agency linear trend	Yes	No	No	No	Yes	No	No	No

Notes: Marijuana possession rates are the counts of marijuana possession per 100k residents covered by a police agency near highway. Distance dummy equals to 1 if an agency is within 150 miles from Colorado. Distances are measured in the unit of 100 miles. "Drop 0" means replacing 0 as missing. "drop TX" means dropping observations from TX. "Balanced" means keeping only agencies that appear in the regression for all seven years. Baseline mean is the mean during 2009-2012 across all agencies in the main result sample. Post equals to 1 for year 2013-2015. All specifications include the unemployment rate and the population ratio of black male aged 20+ at county level. Standard errors (in parentheses) are clustered at county level. *** p<0.01; ** p<0.05; * p<0.1.

Table C.5: Robustness check - Other versions of adult male marijuana possession

VARIABLES	IHS	Log	Mj poss/total	Mj poss/drug poss
Panel A: NIBRS offense				
Distance to CO*Post	-0.247**	-0.161***	-0.008	-0.005
	(0.112)	(0.054)	(0.006)	(0.016)
Baseline mean	5.46	5.17	0.10	0.67
Observations	1,490	1,385	1,490	1,452
R-squared	0.046	0.065	0.050	0.085
Number of clusters	148	144	148	147
Panel B: Stand-Alone arrest				
Distance to CO*Post	-0.289**	-0.144**	-0.001	-0.001
	(0.119)	(0.058)	(0.009)	(0.018)
Baseline mean	4.97	4.87	0.17	0.68
Observations	1,490	1,336	1,474	1,421
R-squared	0.039	0.061	0.057	0.085
Number of clusters	148	143	148	146
Panel C: UCR arrest				
Distance to CO*Post	-0.187	-0.065	0.001	0.016
	(0.117)	(0.054)	(0.002)	(0.015)
Baseline mean	4.47	4.48	0.05	0.63
Observations	1,468	1,286	1,467	1,370
R-squared	0.038	0.065	0.069	0.074
Number of clusters	148	147	147	147
Agency FE	Yes	Yes	Yes	Yes
State*Year FE	Yes	Yes	Yes	Yes

Notes: Marijuana possession rates are the counts of marijuana possession per 100k residents covered by a police agency. "IHS" refers to inverse hyperbolic sine transformed outcome. "Log" refers to natural log transformed outcome. Baseline mean is the mean of y during 2009-2012 across all agencies in regression sample. Distances are measured in the unit of 100 miles. Post equals to 1 for year 2013-2015. All specifications include the unemployment rate and the population ratio of black male aged 20+ at county level. Standard errors (in parentheses) are clustered at county level. *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Table C.6: Robustness check - Various population cutoff, distance dummy as treatment

VARIABLES	10%	20%	30%	40%	50%	60%	70%	80%	90%
Panel A: NIBRS offense									
Distance to CO≤150mi*Post	73.451***	79.477***	83.955***	71.020**	69.247**	83.171**	90.293**	63.552	64.833
	(27.528)	(27.832)	(31.593)	(30.027)	(33.244)	(39.693)	(42.836)	(45.397)	(51.740)
Baseline mean	198.196	204.720	213.880	224.187	225.258	230.394	227.361	209.907	199.939
Observations	2,410	2,141	1,874	1,606	1,340	1,069	801	534	265
R-squared	0.046	0.058	0.078	0.076	0.083	0.086	0.086	0.091	0.195
Number of clusters	189	184	171	155	138	108	82	51	27
Panel B: Stand-Alone arrest									
Distance to CO≤150mi*Post	55.243**	60.584***	62.510**	51.379**	53.146**	62.690**	75.471**	53.508	47.901
	(22.225)	(22.480)	(25.443)	(21.880)	(24.137)	(30.260)	(35.388)	(32.958)	(35.454)
Baseline mean	146.78	151.78	156.39	163.75	162.21	165.43	164.16	152.45	141.32
Observations	2,410	2,141	1,874	1,606	1,340	1,069	801	534	265
R-squared	0.057	0.066	0.087	0.085	0.096	0.092	0.096	0.118	0.175
Number of clusters	189	184	171	155	138	108	82	51	27
Panel C: UCR arrest									
Distance to CO≤150mi*Post	34.329**	34.976**	34.600*	20.817	24.652	33.633*	45.805**	30.609	28.449
	(16.639)	(16.723)	(18.794)	(14.745)	(16.255)	(20.038)	(20.330)	(20.322)	(24.208)
Baseline mean	103.455	107.664	111.375	115.530	115.916	117.645	116.427	107.152	103.963
Observations	2,277	2,066	1,832	1,579	1,324	1,059	796	530	263
R-squared	0.088	0.096	0.116	0.111	0.129	0.114	0.126	0.075	0.157
Number of clusters	188	184	171	155	138	108	82	51	27
Population cutoff	872	1230	1757	2263	3140	4509	6671	11346	25476
Agency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State*Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Marijuana possession rates are the counts of marijuana possession per 100k residents covered by a police agency. Baseline mean is the mean of y during 2009-2012 across all agencies above the specific population cutoff. Distances are measured in the unit of 100 miles. Post equals 1 for year 2013-2015. All specifications include the unemployment rate and the population ratio of black male aged 20+ at county level. Standard errors (in parentheses) are clustered at county level. *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Table C.7: Robustness check - Various population cutoff, continuous distance as treatment

VARIABLES	10%	20%	30%	40%	50%	60%	70%	80%	90%
Panel A: NIBRS offense									
Distance to CO*Post	-50.790***	-54.613***	-52.224***	-44.568***	-40.048***	-44.570**	-57.548***	-46.923**	-41.174
	(11.296)	(12.166)	(13.343)	(13.488)	(14.449)	(17.540)	(21.106)	(21.629)	(32.371)
Baseline mean	198.196	204.720	213.880	224.187	225.258	230.394	227.361	209.907	199.939
Observations	2,410	2,141	1,874	1,606	1,340	1,069	801	534	265
R-squared	0.057	0.070	0.087	0.084	0.090	0.091	0.101	0.110	0.202
Number of clusters	189	184	171	155	138	108	82	51	27
Panel B: Stand-Alone arrest									
Distance to CO*Post	-42.207***	-42.513***	-38.862***	-30.804***	-27.011**	-29.906**	-46.553***	-30.414*	-24.545
	(9.851)	(10.290)	(11.130)	(10.350)	(10.622)	(13.106)	(16.880)	(16.124)	(22.911)
Baseline mean	146.78	151.78	156.39	163.75	162.21	165.43	164.16	152.45	141.32
Observations	2,410	2,141	1,874	1,606	1,340	1,069	801	534	265
R-squared	0.068	0.077	0.094	0.090	0.099	0.093	0.109	0.125	0.169
Number of clusters	189	184	171	155	138	108	82	51	27
Panel C: UCR arrest									
Distance to CO*Post	-24.226***	-24.213***	-21.351**	-12.288*	-10.503	-14.573*	-27.848***	-15.581	-13.656
	(7.279)	(7.679)	(8.439)	(7.062)	(7.063)	(8.393)	(9.826)	(9.762)	(15.025)
Baseline mean	103.455	107.664	111.375	115.530	115.916	117.645	116.427	107.152	103.963
Observations	2,277	2,066	1,832	1,579	1,324	1,059	796	530	263
R-squared	0.094	0.102	0.119	0.112	0.129	0.113	0.134	0.076	0.151
Number of clusters	188	184	171	155	138	108	82	51	27
Population cutoff	872	1230	1757	2263	3140	4509	6671	11346	25476
Agency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State*Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Marijuana possession rates are the counts of marijuana possession per 100k residents covered by a police agency. Baseline mean is the mean of y during 2009-2012 across all agencies above the specific population cutoff. Distances are measured in the unit of 100 miles. Post equals to 1 for year 2013-2015. All specifications include the unemployment rate and the population ratio of black male aged 20+ at county level. Standard errors (in parentheses) are clustered at county level. *** p<0.01; ** p<0.05; * p<0.1.

Table C.8: Robustness check - Effect of RML on juvenile male marijuana possession

VARIABLES	NIBRS offense	Stand-Alone arrest	UCR arrest
Distance to CO*Post	-1.107	-1.943	-0.095
	(2.954)	(2.687)	(2.902)
Baseline mean	51.53	35.53	29.58
Observations	1,490	1,490	1,468
R-squared	0.053	0.035	0.041
Agency FE	Yes	Yes	Yes
State*Year FE	Yes	Yes	Yes
Number of clusters	148	148	148

Notes: Juvenile here refers to anyone aged below 18. Marijuana possession rates are the counts of marijuana possession per 100k residents covered by an agency. Baseline mean is the mean of y during 2009-2012 across all agencies in the regression sample. Distances are measured in the unit of 100 miles. Post equals to 1 for year 2013-2015. All specifications include the unemployment rate and the population ratio of black male aged 20+ at county level. Standard errors (in parentheses) are clustered at county level. *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Table C.9: Robustness check - Effect of RML on other adult male illicit drug possessions

VARIABLES	Cocaine	Heroin	Opium	Morphine	Hashish
Panel A: NIBRS offense					
Distance to CO*Post	-0.755	-2.598	-0.313	0.411	-0.499
	(2.296)	(2.055)	(0.398)	(0.578)	(0.459)
Baseline mean	13.80	6.07	0.87	1.41	0.58
Observations	1,490	1,490	1,490	1,490	1,490
R-squared	0.085	0.111	0.031	0.038	0.039
Panel B: NIBRS stand-alone arrest					
Distance to CO*Post	-0.695	-0.749	-0.077	0.048	-0.553
	(1.280)	(0.702)	(0.195)	(0.381)	(0.403)
Baseline mean	8.92	4.07	0.63	0.71	0.40
Observations	1,490	1,490	1,490	1,490	1,490
R-squared	0.093	0.140	0.035	0.073	0.046
Number of clusters	148	148	148	148	148
Agency FE	Yes	Yes	Yes	Yes	Yes
State*Year FE	Yes	Yes	Yes	Yes	Yes
Number of clusters	148	148	148	148	148

Notes: Drug possession rates are the counts of drug possession per 100k residents covered by an agency. Baseline mean is the mean of y during 2009-2012 across all agencies in the regression sample. Distances are measured in the unit of 100 miles. Post equals to 1 for year 2013-2015. All specifications include the county unemployment rate and the county population ratio of black male aged 20+. Standard errors (in parentheses) are clustered at county level. *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Table C.10: Robustness check - Quantity of possessed marijuana in other locations

	(1)	(2)
VARIABLES	mj poss(kg)	mj arr(kg)
Distance to CO*Post	-0.279	-0.268
	(0.284)	(0.236)
Baseline mean	0.21	0.20
Observations	1,262	1,185
R-squared	0.022	0.028
Agency FE	Yes	Yes
State*Year FE	Yes	Yes
Number of clusters	142	139

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: Other location is defined as any locations that are not code "13" (near highway/road/alley/street/sidewalk). Amount of possessed marijuana is in unit KG. Baseline mean is Baseline mean is the mean of y during 2009-2012 across all agencies in the regression sample. Distances are measured in the unit of 100 miles. Post equals to 1 for year 2013-2015. All specifications include the unemployment rate and the population ratio of black male aged 20+ at county level. Standard errors (in parentheses) are clustered at county level. *** p<0.01; ** p<0.05; * p<0.1.

Bibliography

[1] Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). When should you adjust standard errors for clustering? (No. w24003). National Bureau of Economic Research

[2] Adda, J., McConnell, B., & Rasul, I. (2014). Crime and the depenalization of cannabis possession: Evidence from a policing experiment. Journal of Political Economy, 122, 1130-1202.

[3] Ager, P., Bursztyn, L., & Voth, H. J. (2016). Killer incentives: Status competition and pilot performance during world war II (No. w22992). National Bureau of Economic Research.

[4] Akerlof, G. A. (1997). Social distance and social decisions. Econometrica: Journal of the Econometric Society, 1005-1027.

[5] Akerlof, G. A., & Kranton, R. E. (2005). Identity and the Economics of Organizations. Journal of Economic Perspectives, 19(1), 9-32.

[6] Anderson, D. M., & Rees, D. I. (2014). The role of dispensaries: The devil is in the details. Journal of Policy Analysis and Management, 33, 235-240.

[7] Andiappan, M., & Dufour, L. (2020). Jealousy at work: A tripartite model. Academy of Management Review, 45(1), 205-229.

[8] Andrade, S. C., Bian, J., & Burch, T. R. (2013). Analyst coverage, information, and bubbles. Journal of Financial and Quantitative Analysis, 48(5) 1573-1605.

[9] Ammann, M., Horsch, P., & Oesch, D. (2016). Competing with superstars. Management Science, 62(10) 2842-2858.

[10] Ariely, D., Bracha, A., & Meier, S. (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. American Economic Review, 99(1), 544-55.

[11] Ashraf, N., Bandiera, O., & Lee, S. S. (2014). Awards unbundled: Evidence from a natural field experiment. Journal of Economic Behavior and Organization, 100, 44-63.

[12] Ashraf, A. (2017). Do Rank Incentives Increase Productivity? Evidence from a Field Experiment. Working Paper.

[13] Bachhuber, M. A., Saloner, B., Cunningham, C. O., & Barry, C. L. (2014). Medical cannabis laws and opioid analgesic overdose mortality in the United States, 1999-2010. JAMA internal medicine, 174, 1668-1673.

[14] Barber, B. M., Lehavy, R., McNichols, M., & Trueman, B. (2006). Buys, holds, and sells: The distribution of investment banks' stock ratings and the implications for the profitability of analysts' recommendations. Journal of accounting and Economics, 41(1-2) 87-117.

[15] Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. The Review of Financial Studies, 21(2) 785-818.

[16] Bénabou, R., & Tirole, J. (2002). Self-confidence and personal motivation. The Quarterly Journal of Economics, 117(3), 871-915.

[17] Bernheim, B. D. (1994). A theory of conformity. Journal of Political Economy, 102(5), 841-877.

[18] Bharadwaj, P. (2015). Impact of Changes in Marriage Law Implications for Fertility and School Enrollment. Journal of Human Resources, 50, 614-654.

[19] Blader, S., Gartenberg, C., & Prat, A. (2020). The contingent effect of management practices. The Review of Economic Studies, 87(2), 721-749.

[20] Boni, L., & Womack, K. L. (2006). Analysts, industries, and price momentum. Journal of Financial and Quantitative Analysis, 41(1) 85-109.

[21] Borjas, G. J., & Doran, K. B. (2015). Prizes and productivity how winning the fields medal affects scientific output. Journal of Human Resources, 50(3), 728-758.

[22] Bradler, C., Dur, R., Neckermann, S., & Non, A. (2016). Employee recognition and performance: A field experiment. Management Science, 62(11), 3085-3099.

[23] Bradler, C., Neckermann, S., & Warnke, A. J. (2019). Incentivizing Creativity: A Large-Scale Experiment with Performance Bonuses and Gifts. Journal of Labor Economics, 37(3), 793-851.

[24] Bradshaw, M. T. (2011). Analysts' forecasts: what do we know after decades of work?. Available at SSRN 1880339.

[25] Breathes, W. (2012). The history of cannabis in Colorado... or how the state went to pot. Westword News Web, 1. http://www.westword.com/news/the-history-of-cannabis-in-coloradoor-how-the-state-went-to-pot-5118475

[26] Breza, E., Kaur, S., & Shamdasani, Y. (2017). The morale effects of pay inequality. The Quarterly Journal of Economics, 133(2), 611-663.

[27] Bursztyn, L., & Jensen, R. (2015). How does peer pressure affect educational investments?. The Quarterly Journal of Economics, 130(3), 1329-1367.

[28] Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. Econometrica, 82(6) 2295-2326.

[29] Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. The Review of Economics and Statistics, 90(3):414–427.

[30] Carpenter, J. N., & Whitelaw, R. F. (2017). The development of China's stock market and stakes for the global economy. Annual Review of Financial Economics, 9 233-257.

[31] Caulkins, J. P. (2010). Cost of marijuana prohibition on the California criminal justice system. RAND Drug Policy Research Center Working Paper WR-763-RC. Santa Monica: RAND.

[32] Chan, H. F., Frey, B. S., Gallus, J., & Torgler, B. (2014). Academic honors and performance. Labour Economics, 31, 188-204.

[33] Chan, L. K., Karceski, J., & Lakonishok, J. (2007). Analysts' conflicts of interest and biases in earnings forecasts. Journal of Financial and Quantitative Analysis, 42(4) 893-913.

[34] Chen, S. Y., & Lu, R. (2020). Market Reactions to Awards: Evidence from "Best Financial Analyst" Ranking. Working Paper.

[35] Chu, Y. W. L. (2014). The effects of medical marijuana laws on illegal marijuana use. Journal of Health Economics, 38, 43-61.

[36] Clarke, J., Khorana, A., Patel, A., & Rau, P. R. (2007). The impact of all-star analyst job changes on their coverage choices and investment banking deal flow. Journal of Financial Economics, 84(3), 713-737.

[37] Cohen, L., Frazzini, A., & Malloy, C. (2010). Sell-side school ties. The Journal of Finance, 65(4) 1409-1437.

[38] Cornelissen, T., Dustmann, C., & Schönberg, U. (2017). Peer effects in the workplace. American Economic Review, 107(2), 425-56.

[39] Corwin, S. A., Larocque, S. A., & Stegemoller, M. A. (2017). Investment banking relationships and analyst affiliation bias: The impact of the global settlement on sanctioned and non-sanctioned banks. Journal of Financial Economics, 124(3) 614-631.

[40] Cowen, A., Groysberg, B., & Healy, P. (2006). Which types of analyst firms are more optimistic?. Journal of Accounting and Economics, 41(1-2) 119-146.

[41] Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. The Journal of Finance, 52(3) 1035-1058.

[42] DellaVigna, S., List, J. A., & Malmendier, U. (2012). Testing for altruism and social pressure in charitable giving. The Quarterly Journal of Economics, 127(1), 1-56.

[43] Dewatripont, M., Jewitt, I., & Tirole, J. (1999). The economics of career concerns, part II: Application to missions and accountability of government agencies. The Review of Economic Studies, 66(1), 199-217.

[44] Dontoh, A., Ronen, J., & Sarath, B. (2003). On the rationality of the post-announcement drift. Review of Accounting Studies, 8(1) 69-104.

[45] Ellison, J. M., & Spohn, R. E. (2015). Borders Up in Smoke: Marijuana Enforcement in Nebraska After Colorado's Legalization of Medicinal Marijuana. Criminal Justice Policy Review, 0887403415615649..

[46] Emery, D. R., & Li, X. (2009). Are the Wall Street analyst rankings popularity contests?. Journal of Financial and Quantitative Analysis, 44(2) 411-437.

[47] Engelberg, J. E., & Parsons, C. A. (2011). The causal impact of media in financial markets. The Journal of Finance, 66(1) 67-97

[48] Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. Journal of Financial Economics, 116(1) 1-22.

[49] Fang, L., & Yasuda, A. (2009). The effectiveness of reputation as a disciplinary mechanism in sell-side research. The Review of Financial Studies, 22(9) 3735-3777.

[50] Fedyk, A. (2018). Front page news: The effect of news positioning on financial markets. working paper

[51] Fedyk, A., & Hodson, J. (2019). When can the market identify stale news?. Available at SSRN 2433234

[52] Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, & cooperation. The Quarterly Journal of Economics, 114(3), 817-868.

[53] Feuer, A. (2016, April 14). Taking a Trip, Literally, on Colorado's Pot Trail. Retrieved June 11, 2017, from https://www.nytimes.com/2016/04/17/travel/colorado-weed-marijuana-tour.html?_r=0.

[54] Frey, B. S., & Gallus, J. (2017). Towards an economics of awards. Journal of Economic Surveys, 31(1), 190-200.

[55] Gallus, J. (2016). Fostering public good contributions with symbolic awards: A large-scale natural field experiment at Wikipedia. Management Science, 63(12), 3999-4015.

[56] Gieringer, D. (2009, October 1). Benefits of Marijuana Legalization in California. Retrieved June 11, 2017, from http://www.canorml.org/background/CA_legalization2.html.

[57] Gilbert, T., Kogan, S., Lochstoer, L., & Ozyildirim, A. (2012). Investor inattention and the market impact of summary statistics. Management Science, 58(2) 336-350.

[58] Glassdoor.com (Glassdoor). 2016. Community Guidelines. http://help.glassdoor.com/article/Community-Guidelines/enUS

[59] Goolsbee, A., Lovenheim, M. F., & Slemrod, J. (2010). Playing with fire: Cigarettes, taxes, and competition from the internet. American Economic Journal: Economic Policy, 2, 131-154.

[60] Gubler, T., Larkin, I., & Pierce, L. (2016). Motivational spillovers from awards: Crowding out in a multitasking environment. Organization Science, 27(2), 286-303.

[61] Guryan, J., Kroft, K., & Notowidigdo, M. J. (2009). Peer effects in the workplace: Evidence from random groupings in professional golf tournaments. American Economic Journal: Applied Economics, 1(4), 34-68.

[62] Hansen, B., Miller, K., & Weber, C. (2017). How Extensive is Inter-State Diversion of Recreational Marijuana? (No. w23762). National Bureau of Economic Research. Retrieved Oct 11, 2017, from http://www.nber.org/papers/w23762.pdf

[63] Hao, Z., & Cowan, B. (2017). The Cross-Border Spillover Effects of Recreational Marijuana Legalization (No. w23426). National Bureau of Economic Research.

[64] Hendricks, K. B., & Singhal, V. R. (1996). Quality awards and the market value of the firm: An empirical investigation. Management science, 42(3) 415-436.

[65] Hong, H., & Kubik, J. D. (2003). Analyzing the analysts: Career concerns and biased earnings forecasts. The Journal of Finance, 58(1) 313-351.

[66] Hu, Y., Lin, T. W., & Li, S. (2008). An examination of factors affecting Chinese financial analysts' information comprehension, analyzing ability, and job quality. Review of Quantitative Finance and Accounting, 30(4), 397-417.

[67] Huberman, G., & Regev, T. (2001). Contagious speculation and a cure for cancer: A nonevent that made stock prices soar. The Journal of Finance, 56(1) 387-396.

[68] Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. The Review of Economic Studies, 79(3) 933-959.

[69] Irvine, P., Lipson, M., & Puckett, A. (2007). Tipping. The Review of Financial Studies, 20(3) 741-768.

[70] Jegadeesh, N., & Kim, W. (2006). Value of analyst recommendations: International evidence. Journal of Financial Markets, 9(3) 274-309.

[71] Kandel, E., & Lazear, E. P. (1992). Peer pressure and partnerships. Journal of Political Economy, 100(4), 801-817.

[72] Kaniel, R., & Parham, R. (2017). WSJ Category Kings–The impact of media attention on consumer and mutual fund investment decisions. Journal of Financial Economics 123(2), 337-356.

[73] Kelly, E., & Rasul, I. (2014). Policing cannabis and drug related hospital admissions: Evidence from administrative records. Journal of Public Economics, 112, 89-114.

[74] Khatapoush, S., & Hallfors, D. (2004). "Sending the wrong message": did medical marijuana legalization in California change attitudes about and use of marijuana?. Journal of Drug Issues, 34, 751-770.

[75] Kosfeld, M., & Neckermann, S. (2011). Getting more work for nothing? Symbolic awards and worker performance. American Economic Journal: Microeconomics, 3(3), 86-99.

[76] Kolstad, J. T. (2013). Information and quality when motivation is intrinsic: Evidence from surgeon report cards. American Economic Review, 103(7), 2875-2910.

[77] Kube, S., Maréchal, M. A., & Puppe, C. (2012). The currency of reciprocity: Gift exchange in the workplace. American Economic Review, 102(4), 1644-62.

[78] Lafortune, J., Riutort, J., & Tessada, J. (2018). Role models or individual consulting: The impact of personalizing micro-entrepreneurship training. American Economic Journal: Applied Economics, 10(4), 222-45.

[79] Larkin, I. (2011). Paying $30,000 for a gold star: An empirical investigation into the value of peer recognition to software salespeople. Unpublished working paper.

[80] Latané, B. (1981). The psychology of social impact. American psychologist, 36, 343.

[81] Latané, B., Liu, J. H., Nowak, A., Bonevento, M., & Zheng, L. (1995). Distance matters: Physical space and social impact. Personality and Social Psychology Bulletin, 21, 795-805.

[82] Lawrence, A., Ryans, J., Sun, E., & Laptev, N. (2018). Earnings announcement promotions: A Yahoo Finance field experiment. Journal of Accounting and Economics, 66(2-3) 399-414.

[83] Lazear, E. P. (1989). Pay equality and industrial politics. Journal of Political Economy, 97(3), 561-580.

[84] Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. Journal of Economic Literature, 48(2) 281-355.

[85] Leone, A. J., & Wu, J. S. (2007). What does it take to become a superstar? Evidence from institutional investor rankings of financial analysts. Evidence from Institutional Investor Rankings of Financial Analysts (May 23, 2007). Simon School of Business Working Paper No. FR 02-12.

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

[87] Lo, A. W., & Wang, J. (2000). Trading volume: definitions, data analysis, and implications of portfolio theory. The Review of Financial Studies, 13(2) 257-300.

[88] Loh, R. K., & Stulz, R. M. (2010). When are analyst recommendation changes influential?. The Review of Financial Studies, 24(2) 593-627.

[89] Lovenheim, M. F. (2008). How far to the border?: The extent and impact of cross-border casual cigarette smuggling. National Tax Journal, 7-33.

[90] Lovenheim, M. F., & Slemrod, J. (2010). The fatal toll of driving to drink: The effect of minimum legal drinking age evasion on traffic fatalities. Journal of health economics, 29, 62-77.

[91] Lynch, J. P., & Jarvis, J. P. (2008). Missing data and imputation in the uniform crime reports and the effects on national estimates. Journal of Contemporary Criminal Justice, 24, 69-85.

[92] Malmendier, U., & Tate, G. (2009). Superstar ceos. The Quarterly Journal of Economics, 124(4) 1593-1638.

[93] Mark Anderson, D., Hansen, B., & Rees, D. I. (2013). Medical marijuana laws, traffic fatalities, and alcohol consumption. The Journal of Law and Economics, 56, 333-369.

[94] Mas, A., & Moretti, E. (2009). Peers at work. American Economic Review, 99(1) 112-45.

[95] McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. Journal of Econometrics, 142(2) 698-714.

[96] Medvec, V. H., Madey, S. F., & Gilovich, T. (1995). When less is more: counterfactual thinking and satisfaction among Olympic medalists. Journal of Personality and Social Psychology, 69(4), 603.

[97] Merriman, D. (2010). The micro-geography of tax avoidance: evidence from littered cigarette packs in Chicago. American Economic Journal: Economic Policy, 2, 61-84.

[98] Merton, R. K. (1968). The Matthew effect in science: The reward and communication systems of science are considered. Science, 159(3810) 56-63.

[99] Miller, R. N., & Kuhns, J. B. (2012). Exploring the impact of medical marijuana laws on the validity of self-reported marijuana use among juvenile arrestees over time. Criminal Justice Policy Review, 23, 40-66.

[100] Miron, J. A. (2010). The budgetary implications of drug prohibition. Report from the Criminal Justice Policy Foundation.

[101] Model, K. E. (1993). The effect of marijuana decriminalization on hospital emergency room drug episodes: 1975–1978. Journal of the American Statistical Association, 88, 737-747.

[102] Moreira, D. (2016). Success Spills Over: How Awards Affect Winners' and Peers' Performance in Brazil. Working Paper.

[103] Morgenroth, T., Ryan, M. K., & Peters, K. (2015). The motivational theory of role modeling: How role models influence role aspirants' goals. Review of General Psychology, 19(4), 465-483.

[104] Mullen, E., & Monin, B. (2016). Consistency versus licensing effects of past moral behavior. Annual review of psychology, 67, 363-385.

[105] National Drug Intelligence Center. (2010). Drug Movement Into and Within the United States. Retrieved June 11, 2017, from https://www.justice.gov/archive/ndic/pubs38/38661/movement.htm

[106] Neckermann, S., & Frey, B. S. (2013). And the winner is...? The motivating power of employee awards. The Journal of Socio-Economics, 46, 66-77.

[107] Neckermann, S., Cueni, R., & Frey, B. S. (2014). Awards at work. Labour Economics, 31, 205-217.

[108] Niehaus, G., & Zhang, D. (2010). The impact of sell-side analyst research coverage on an affiliated broker's market share of trading volume. Journal of Banking & Finance, 34(4) 776-787.

[109] Pacula, R. L., Kilmer, B., Grossman, M., & Chaloupka, F. J. (2010). Risks and prices: The role of user sanctions in marijuana markets. The BE Journal of Economic Analysis & Policy, 10.

[110] Peress, J. (2014). The media and the diffusion of information in financial markets: Evidence from newspaper strikes. The Journal of Finance, 69(5), 2007-2043.

[111] Phillips, B., Pukthuanthong, K., & Rau, P. R. (2014). Past performance may be an illusion: Performance, flows, and fees in mutual funds. Critical Finance Review, Forthcoming.

[112] Robinson, C. D., Gallus, J., Lee, M. G., & Rogers, T. (2019). The demotivating effect (and unintended message) of awards. Organizational Behavior and Human Decision Processes.

[113] Sequeira, S., Spinnewijn, J., & Xu, G. (2016). Rewarding schooling success and perceived returns to education: Evidence from India. Journal of Economic Behavior & Organization, 131, 373-392.

[114] Spence, Michael. (1973). Job Market Signaling. Quarterly Journal of Economics, 87(3): 355—374.

[115] Stehr, M. (2007). The Effect of Sunday Sales Bans and Excise Taxes on Drinking and Cross—Border Shopping for Alcoholic Beverages. National Tax Journal, 85-105.

[116] Stein, J. C. (2009). Presidential address: Sophisticated investors and market efficiency. The Journal of Finance, 64(4), 1517-1548.

[117] Stickel, S. E. (1992). Reputation and performance among security analysts. The Journal of Finance, 47(5) 1811-1836.

[118] Tetlock, P. C. (2011). All the news that's fit to reprint: Do investors react to stale information?. The Review of Financial Studies, 24(5) 1481-1512.

[119] United States Department of Justice, Federal Bureau of Investigation. (September 2015). Crime in the United States, 2014. Retrieved June 11, 2017, from https://ucr.fbi.gov/crime-in-the-u.s/2014/crime-in-the-u.s.-2014/resource-pages/about-cius.

[120] Wall, M. M., Poh, E., Cerdá, M., Keyes, K. M., Galea, S., & Hasin, D. S. (2011). Adolescent marijuana use from 2002 to 2008: higher in states with medical marijuana laws, cause still unclear. Annals of epidemiology, 21, 714-716.

[121] Warzynski, F., Smeets, V., & Waldman, M. (2017). Performance, Career Dynamics, & Span of Control. Journal of Labor Economics.

[122] Wen, H., Hockenberry, J. M., & Cummings, J. R. (2015). The effect of medical marijuana laws on adolescent and adult use of marijuana, alcohol, and other substances. Journal of health economics, 42, 64-80.

[123] Wu, J. S., & Zang, A. Y. (2009). What determine financial analysts' career outcomes during mergers?. Journal of Accounting and Economics, 47(1-2) 59-86.

[124] Xu, N., Chan, K. C., Jiang, X., & Yi, Z. (2013). Do star analysts know more firm-specific information? Evidence from China. Journal of Banking & Finance, 37(1) 89-102.