ESSAYS ON FINANCE, TECHNOLOGY, AND LABOR

A Dissertation
Presented to
The Academic Faculty

By

Linghang Zeng

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
Scheller College of Business

Georgia Institute of Technology

August 2019

ESSAYS ON FINANCE, TECHNOLOGY, AND LABOR

Approved by:

Dr. Sudheer Chava, Advisor
Scheller College of Business
Georgia Institute of Technology

Dr. Jonathan Clarke
Scheller College of Business
Georgia Institute of Technology

Dr. Alexander Oettl
Scheller College of Business
Georgia Institute of Technology

Dr. Manpreet Singh
Scheller College of Business
Georgia Institute of Technology

Dr. Rohan Ganduri
Goizueta Business School
Emory University

Date Approved: May 2, 2019

To my wife, Ting,

and my son, MeMe

ACKNOWLEDGEMENTS

I am extremely fortunate and blessed to have Sudheer Chava as my advisor. I am deeply grateful for his continuous guidance, help, and support for my research and career. I could not hope for a better advisor. I very much appreciate Manpreet Singh, Rohan Ganduri, and Daniel Weagley for their generosity with their time, and providing constant advice and encouragement. I enjoy and learn a lot from working with Jonathan Clarke, Alex Oettl, Alex Hsu, Daniel Bradley, and Nikhil Paradkar. It is truly an honor and pleasure to have been working with all of you.

I want to thank Suzanne Lee, Soohun Kim for guiding me during the first two years. I am thankful to all other faculty members for advice, comments, and building a great research environment for PhD students. I thank Steven Xiao, Teng Zhang, Youngmin Choi, Minho Wang, Peter Simasek, Baridhi Malakar, and other PhD students for help and interesting discussions. I am grateful to Jeff Dugger, Michael McBurnett, Piyush Patel, Chris Yasko, and Equifax Inc. for providing data and assistance in using the data.

I would like to thank my parents, Xiangdi and Aicui, for their faith in me and continuous support. Lastly, I express my deepest gratitude to my wife, Ting, for love and accompanying me on this journey. To my son, MeMe, you keep me motivated at the last mile.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This dissertation consists of three essays on finance, technology, and labor. In the first essay, using an anonymized employer–employee matched payroll dataset of 70 million workers, I examine the effect of venture capital (VC) investment flows on large local incumbent firms. I find that a VC investment of $1,000 per capita nearly doubles the wage growth for high-skilled workers in that region, while there is no effect for low-skilled workers. This wage effect is particularly strong in regions with inelastic short-term labor supply, and regions with lower enforceability of non-compete agreements. To mitigate concerns about time-varying shocks at the industry-region level, I exploit within-industry-state and within-firm variation, conduct occupation-level analysis, and utilize variation in VC funding. Further, I find that incumbent firms experience lower growth of high-skilled employment and a higher departure rate among patent inventors. Incumbent firms also cut R&D and generate lower-quality innovation. Overall, my results highlight that VC investment inflows, through competition for talent, can have distributional effects on workers and adverse consequences for some local incumbent firms.

In the second essay, using an employer-employee payroll dataset for approximately 2.6 million retail workers, we find that the staggered rollout of a major e-commerce retailer's fulfillment centers (FCs) has a negative effect on the income of retail workers in geographically proximate counties. Wages of hourly workers, especially part-time hourly workers, decrease significantly. There is a decrease in credit scores and an increase in delinquency for retail workers with higher prior credit utilization. Evidence from 3.2 million retail stores shows that geographically proximate stores experience a reduction in sales, number of employees and there is a decrease in entry and an increase in exit for small and young stores. Our robustness tests show that prevailing local economic conditions are unlikely to drive our results. Our results highlight the extent to which a dramatic increase in e-commerce retail sales can have adverse consequences for some of the workers at traditional brick-and-

mortar stores.

In the third essay, we show that, between 2009 and 2013, *Theflyonthewall.com (FLY)* leaks 58% of recommendation revisions with a median delay of 27 minutes relative to the I/B/E/S announcement time. *FLY* improves price discovery, but leaked recommendations hamper the ability of brokers to offer price improvement on trades routed through them. Three major brokers sued *FLY* and using key court dates, we show significant wealth and real effects to the brokerage industry. Overall, the speed with which analyst recommendations are disseminated has led to more rapid price discovery at the expense of a decline in the scope of the sell-side research industry.

CHAPTER 1

IMPACT OF VENTURE CAPITAL FLOWS ON INCUMBENT FIRMS: EVIDENCE FROM 70 MILLION WORKERS

1.1 Introduction

The venture capital (VC) industry in the U.S. experienced massive growth since the 1980s, with annual investment increasing from $610 million in 1980 to $84 billion in 2017.[1] During this time, VC emerged as the dominant form of equity financing for young, privately held, high-tech companies ([76]). An extensive literature documents that VC creates enormous benefits for startups (e.g., spurring innovation, commercializing products, building human capital).[2] These startups compete with large incumbent firms in the product and labor markets. But, there is sparse evidence on how VC flows affect incumbent firms. In this paper, I examine how VC flows into a region affect wages and employment at local incumbent firms.

While VC-backed startups and incumbent firms may compete in multiple markets (e.g., product market, labor market), I focus on competition for talent because human capital is increasingly becoming a crucial asset for firms ([109]). As [86] points out, the potential loss of specific human capital has important implications for firms. Venture capitalists play a vital role in building the human capital of startups ([60, 40]). The expansion of VC-backed startups' workforce creates additional demand for talent. In a frictionless labor market, worker supply and quality can fully adjust to demand shocks. Therefore, the increased demand for talent caused by VC flows may not have a significant impact on wages at local

[1] https://nvca.org/pressreleases/record-unicorn-financings-drove-2017-total-venture-capital-investments-84-billion-largest-amount-since-dot-com-era/. This $84 billion is about 0.43% of the U.S. GDP ($19.49 trillion according to the Bureau of Economic Analysis), and it is 240% more than the U.S. aggregate IPO proceeds ($24.53 billion according to Jay Ritter's website) in 2017.

[2] See [39] for a complete survey.

incumbent firms. However, in the presence of labor market frictions, VC flows may pose a labor turnover risk for incumbent firms and affect their wages. First, if the new local labor supply is limited, the additional labor demand creates more pressure for talent at incumbent firms. Second, in a world where labor mobility is imperfect, larger VC investments can result in higher wage growth ([90]). Third, workers have more bargaining power due to the increased outside opportunities only if they can move freely across firms.

There are two primary challenges in identifying the effect of VC flows into a region on wages and employment at local incumbent firms. The first challenge is that publicly available wages and employment data are at the aggregate level and do not differentiate startups from incumbent firms. To overcome this challenge, I obtain an anonymized employer–employee matched payroll dataset for approximately 70 million workers of incumbent firms. This dataset includes over 70% of the Fortune 500 companies and covers a significant portion of the U.S. workforce. It provides detailed information on workers' wages, education level, occupation, location, and tenure. Also, this dataset allows me to observe establishments belonging to the same firm but in different regions.

The second challenge is that VC investment flows are unlikely to be random. The estimated effect can be biased if unobservable factors at the local area correlate with both VC investment flows into the region and the wage growth at local incumbent establishments. For example, venture capitalists may choose to invest in startups in a good industry and in a good region at a particular time. But, the granularity of the payroll data allows me to compare establishments in the same industry and in the same state in a given year (i.e., industry $\times$ state $\times$ year fixed effects) to control for time-varying shocks at the industry and region level. I can also compare establishments belonging to the same firm but having different exposures to VC flows (i.e., firm $\times$ year fixed effects) to control for a time-varying, firm-specific productivity increase or a technology adoption. However, it is still possible that time-varying omitted variables that are specific to a particular industry and a particular CZ drive the results. Because the inclusion of industry $\times$ CZ $\times$ year fixed effects would completely absorb the variation in VC flows, I exploit heterogeneous wage effects within an establishment based on workers' occupation and exploit plausibly exogenous variation in VC funding.

Following [89], in fixed effects models, I include establishment, industry $\times$ year, state $\times$ year, and industry $\times$ state $\times$ year fixed effects. The inclusion of granular fixed effects should absorb most of the unobserved heterogeneity. I find that VC investments in a commuting zone (CZ) have distributional effects on wages. There is a positive effect on the wage growth of high-skilled workers who have at least a bachelor's degree. This effect results in a level increase in wages. The effect is insignificant for low-skilled workers who have only a high school diploma. Among high-skilled workers, higher-wage workers benefit more. In particular, a VC investment of \$1,000 per capita in a CZ is associated with an increase of \$849–\$1,218 in the 90^{th} percentile of high-skilled workers' wages, nearly doubling the unconditional average wage growth. The results hold after including firm $\times$ year fixed effects.

The main heterogeneity that is not potentially controlled for in the fixed effects models that I have estimated so far is the possibility of time-varying omitted variables that are specific to an industry in a particular CZ. One example would be a technological shock in the IT industry of Atlanta in 2014 that boosts the wages in the local IT industry and, at the same time, attracts more VC investments in the local IT industry. I leverage the rich occupation data to examine the impact of one industry's VC investments on high-skilled workers in other industries. The cross-industry analysis mitigates some of the concerns that a time-varying industry $\times$ CZ level unobservable drives the results that I document so far. Specifically, I analyze the effect of VC investments in the information technology (IT) industry on talented computer workers (90^{th} percentile) in non-IT industries. I use skilled non-computer workers as a control group, and I include establishment $\times$ year and occupation $\times$ year fixed effects to absorb any time-varying establishment-level and occupation-level shocks. I find that, compared to non-computer workers in the same incumbent establishments, computer workers have higher wage growth when VC investments in local IT startups are greater.

Also, I utilize plausibly exogenous variation in VC funding to further establish causality. VC investments are determined by the availability of VC funding and the decisions of venture capitalists to invest. While the former is less of a concern, the latter is endogenous. VC investments in 2014 are almost double the investments in 2013. One dominant explanation is the low-interest-rate environment, which is likely uncorrelated with investment opportunities. Similar to [16], I exploit this funding shock and construct ex ante exposure to national VC investments for each CZ. A difference-in-differences estimation indicates that high-skilled workers in CZs with higher ex ante exposure have higher wage growth after 2014. Also, following [96], and [53], I use endowment returns and inflows into buyout funds as instrumental variables for VC investment flows. The results hold in these instrumental variable (IV) estimations. Another potential endogeneity concern is reverse causality. Namely, startups seek VC investments when local competition for talent is high. An analysis of the dynamic effects of VC flows reveals that wage growth is uncorrelated with future VC investments. Therefore, reverse causality does not appear to be a major issue.

My results can potentially be explained by several alternative hypotheses that can be non-mutually exclusive. I attempt to rule out these alternative hypotheses. First, I show that my results are not driven by other financial activities that have been shown to have an impact on labor outcomes and may be correlated with VC investments, including initial public offerings (IPOs), private equity investments, and horizontal mergers and acquisitions. Second, it is likely that VC investments are correlated with entrepreneurship activities. Therefore, I control for entrepreneurship activities by including CZ-level establishment growth. Third, the results are also robust to controlling for general local economic conditions, such as wage growth and employment growth. Fourth, because the employment growth of incumbent establishments decreases, on average, my results are unlikely to be driven by competition among incumbent establishments. Also, the results are not driven by any VC-intensive region (e.g., California, Massachusetts).

While it is always challenging to establish causality, the totality of evidence from (a) the inclusion of fine-grained fixed effects; (b) the cross-industry, occupation-level analysis; (c) the exploitation of plausibly exogenous variation in VC funding; (d) the lack of reverse causality; and (e) ruling out multiple alternative hypotheses gives me confidence that VC flows have a causal effect on wages at incumbent establishments.

Next, I examine the impact of VC investments on employment at incumbent establishments. It appears that, for incumbent establishments, the employment growth of high-skilled workers decreases, while the employment growth of low-skilled workers stays mostly unaffected when local VC investments are large. Further, I demonstrate that the establishment-level employment growth of high-skilled workers also responds to VC investments in other regions where the firm operates. Specifically, larger VC investments in other areas where the firm operates results in an increase in the establishment-level employment growth of high-skilled workers. The results suggest that incumbent firms may have made strategic labor decisions when facing competition for talent. The results also highlight the potential bright side of geographic diversification.

I discuss four non-mutually exclusive mechanisms through which VC flows affect wages and employment at incumbent establishments. The first hypothesized mechanism is competition for talent. VC investment flows increase local competition for talent, and therefore incumbent establishments must raise wages to hire or retain talent. One of the critical assumptions for this channel is that startups hire more upon receiving VC funding. To validate this assumption, I match VC deal data to labor condition application (LCA) data and conduct event panel studies. I find that startups increase hiring by 4.7% after receiving VC funding compared to firms in the same CZ and the same industry. The effect is more pronounced, at 15.1%, for deals involving at least $100 million.

Competition for talent is consistent with the findings that both the wage increases and the employment decreases are confined to high-skilled workers. The competition for talent hypothesis also predicts that the effect should be stronger in areas with constrained local talent supply, areas that are less appealing to high-skilled workers, as well as areas with stronger labor mobility. I find supporting evidence that the impact on high-skilled workers' wages is more pronounced in regions that lack top-level engineering schools, regions that have fewer pleasant days according to weather data, and regions that have lower enforceability of non-compete agreements.

The second hypothesized mechanism is voluntary departure for entrepreneurship. In this scenario, rather than joining VC-backed startups, workers at incumbent establishments leave their jobs to become entrepreneurs after sensing the opportunities provided by VC. However, the main results are robust to controlling for establishment growth, which is a proxy for entrepreneurship activities. The third mechanism is the creation of new skills. Following VC investments in startups, incumbent establishments create new jobs that require new skills, resulting in a shift in the composition of workers. But, I find that existing workers also experience higher wage growth. The fourth mechanism is knowledge spillovers. VC investments generate knowledge spillovers, and wage growth reflects productivity growth. However, this mechanism cannot fully explain the reduction in employment growth and lower-quality innovation that I document. Overall, while it is possible that other mechanisms may also play a role, I find that the competition for talent channel best explains all the results that I document in the paper.

What is the impact of competition for talent caused by VC flows on incumbent firms? Hypothetically, if firms can reallocate workers from high-VC CZs to low-VC CZs to minimize the impact of competition, the effect on firm-level outcomes is muted. However, the firm-level analysis shows that firms that are headquartered in CZs that have larger VC investments experience higher labor costs and a reduction in employment growth. The evidence indicates that firms may not be able to reallocate workers to completely avoid competition for talent due to VC-backed startups.

The increased labor costs may negatively affect firms' cash flow. At the same time, competition for talent may also tighten labor constraints for some incumbent firms. Therefore, VC flows could affect firms' investment policies. I find that VC flows have a negative effect on the R&D of incumbent firms, while there is no effect on capital expenditure. These results show that firms cut R&D, the input of innovation (See [30]). Next, I examine the output of innovation. I find that while incumbent firms produce more patents, the total adjusted citations do not increase. Consequently, the average adjusted citations decrease, suggesting incumbent firms generate lower-quality innovation. Moreover, I show that investors are more likely to leave incumbent firms, which could partially explain the decrease in innovation quality of incumbent firms. Overall, these results suggest that competition for talent due to VC investment flows may have a long-term impact on incumbent firms.

This paper contributes to the literature on the real effects of VC. VC creates enormous benefits for VC-backed startups.[3] They generate knowledge spillovers across firms ([54] and [97]). VC investments also boost local economic growth ([96]). I document a new channel–competition for talent, through which regional VC flows have distributional effects on workers at local incumbent establishments. This competition creates additional costs for incumbent firms and has implications for firms' innovation input and output. While the benefits of VC are well documented in the literature, the costs of VC are not as well understood ([39]). The results that I document in this paper shed light on this underexplored area.

My paper also contributes to the recent research on local labor market competition and wages. [19] and [13] document that wages are lower in regions with higher employer concentration. They attribute their findings to the monopsony power of employers. I contribute to this literature by focusing on an increase in labor market competition driven by VC in-

[3]Spurring innovation: [76]; Commercializing products: [59]; Building human capital: [60] find that venture capitalists help startups with professionalization of human capital. [40] document a positive relationship between VC funding and employment growth of startups. [32] show that VC increases firm productivity. [95] find that VC-backed firms account for 4% to 5.5% of total employment from 1980 to 2005. [4] provide evidence that venture capitalists who are on the board help recruit managers and board members.

vestment flows. More importantly, I analyze the differential impact on high-skilled and low-skilled workers.

My paper also broadly relates to the emerging literature on labor and finance (see [86] for a survey). In particular, my paper relates to the research on the distributional effects of financial activities and technological innovation on workers (e.g., [31, 75, 78, 83]) and the implications of labor mobility risk for firms (e.g., [49, 71, 74]).

The rest of this paper proceeds as follows. Section 1.2 discusses the empirical design and identification challenges. Section 1.3 describes the data. Section 1.4 presents the main empirical results. Section 1.5 discusses several potential mechanisms. Section 1.6 shows the implications for firms. Section 1.7 concludes.

1.2 Empirical Design and Identification Challenges

This paper aims to estimate the impact of VC investment flows on wages at local incumbent establishments. An ideal experiment would be to assume there are two identical regions and venture capitalists randomly invest in startups in one of the two regions. The effect can be quantified by comparing the wage growth of incumbent establishments in the two regions. However, in practice, identifying the causal effect is challenging because VC investments are not random. The major endogeneity problem is the presence of omitted variables that affect the wages of incumbent establishments and are correlated with VC investment flows. For instance, a technological shock in a region boosts wages and attracts more VC investments. A minor identification concern is reverse causality. In this section, I discuss my baseline empirical model, potential concerns, and how I use multiple approaches to tackle identification challenges: fine-grained fixed effects models, occupation-level analysis, and exploiting plausibly exogenous variation in VC funding.

1.2.1 Fixed Effects Models

To quantify the wage effect of VC investment flows on local incumbent establishments, I start with a baseline regression specification at the establishment-year level,

$$\Delta Y_{e,c,t} = \beta VCPerCapita_{c,t} + \eta_e + \gamma_t + \epsilon_{e,c,t}, \tag{1.1}$$

where $\Delta Y_{e,c,t}$ is the dollar change in a given percentile of high-skilled workers' wages (e.g., $\Delta P90\_HighSkilled$) or the dollar change in a given percentile of low-skilled workers' wages (e.g., $\Delta P10\_LowSkilled$) in establishment e, commuting zone (CZ) c, and year t.[4] The variable $VCPerCapita_{c,t}$ is the dollar value of VC investments (in thousands of dollars) in CZ c, in year y, divided by the CZ population in 2000.[5] I include establishment fixed effects (η_e) and year fixed effects (γ_t) to control for time-invariant establishment-specific characteristics and transitory common shocks, respectively. The error term is $\epsilon_{e,c,t}$. Standard errors are clustered at the CZ level to account for correlations within a CZ. CZs that have received VC investments may differ from CZs that have never received VC investments in terms of, e.g., economic environment, regulatory environment, or industry composition. To mitigate this concern, I include in the main analysis CZs that have received VC investments during the sample period.[6]

One of the challenges is classifying workers into high-skilled and low-skilled workers, because skills are difficult to measure and there is no consensus on the optimal method. [46] argues that a proxy for skills should capture education, training, learning from coworkers, and the ability to adapt to changes. Among these factors, education attainment is directly observed and widely used. Following [73] and [6], I define *high-skilled workers* as those

[4] In this paper, a local area is defined as a CZ, which is proposed by [103]. A CZ is a group of counties with strong commuting ties. They represent the local economy better than a county. U.S. counties are uniquely grouped into one of the 709 CZs. See https://www.ers.usda.gov/data-products/commuting-zones-and-labor-market-areas/ for the mapping between counties and CZs.

[5] Scaling considers the fact that a $1 billion investment in a CZ with a large population generates a smaller effect than the same amount of investment in a CZ with a small population. The results remain similar if I scale VC investments by the lagged employment.

[6] The results remain similar if I include all CZs.

who have at least a bachelor's degree, and I define *low-skilled workers* as those who have only a high school diploma.

Equation 1.1 exploits the geographic and temporal variation in VC investment flows. However, VC may select good CZs or good industries to invest in. The inclusion of establishment fixed effects controls for permanent heterogeneity at the plant, CZ, and industry level. The coefficient of interest β represents the contemporaneous wage effect of VC investments, namely, the change in wages (in dollars) for VC investment of $1,000 per capita. The estimation is unbiased only if $VCPerCapita_{c,t}$ is orthogonal to the error term $\epsilon_{e,c,t}$. A positive correlation between VC investments and omitted unobservables may overestimate β.

The major endogeneity problem with the baseline regression is omitted variables. It is possible that time-varying, skill-based technological shocks at the industry or location level result in higher wage growth because productivity increases, while technological shocks attract more VC investments. The granularity of the establishment-level panel data allows me to include industry $\times$ year, state $\times$ year, and industry $\times$ state $\times$ year fixed effects ($\gamma_{j,s,t}$).[7] The identification is achieved by comparing two establishments that are within the same state and industry but face different VC investments in the same year:

$$\Delta Y_{e,c,t} = \beta VCPerCapita_{c,t} + \eta_e + \gamma_{j,s,t} + \epsilon_{e,c,t}. \tag{1.2}$$

Moreover, I can include firm $\times$ year fixed effects to control for time-varying, firm-specific heterogeneity (e.g., firm-level productivity increase, technology adoption), and I identify the coefficient of interest by comparing wage changes in areas with high VC investments to wage changes in areas with low VC investments within the same firm in the same year. Technological shocks are likely to affect all establishments belonging to the same firm, as it is unlikely that a firm, especially one that operates in a single industry, adopts a new technology only at some of its establishments. The inclusion of firm $\times$ year

[7] Industries are defined based on 3-digit NAICS codes.

fixed effects ($\gamma_{i,t}$) should mostly mitigate the concern that unobserved technological shocks are driving the results:

$$\Delta Y_{e,c,t} = \beta VCPerCapita_{c,t} + \eta_e + \gamma_{i,t} + \epsilon_{e,c,t}. \tag{1.3}$$

Similar to the empirical framework of [89], after controlling for fine-grained fixed effects (industry $\times$ state $\times$ year fixed effects or firm $\times$ year fixed effects), the identifying assumption is that VC investments are uncorrelated with omitted unobservables. It is still likely that VC investments are correlated with time-varying industry $\times$ CZ unobserved factors. However, it is impossible to add industry $\times$ CZ $\times$ year fixed effects in Equation 1.1 because these fixed effects can completely absorb variation in VC investments. I use two alternative approaches to alleviate this concern in Section 1.2.2 and 1.2.3.

A minor endogeneity concern is reverse causality. Specifically, startups seek VC investments when competition for talent is high and wage growth is high. To mitigate the concern of reverse causality, I include the lead and lag VC investments in Equation 1.4:

$$\Delta Y_{e,c,t} = \beta_1 VCPerCapita_{c,t+1} + \beta_2 VCPerCapita_{c,t} + \beta_3 VCPerCapita_{c,t-1} + \eta_e + \gamma_{j,s,t} + \epsilon_{e,c,t}. \tag{1.4}$$

The key coefficient for this test is β_1. An insignificant estimation of β_1 suggests there is no significant relationship between the wage growth in year t and VC investments in year $t+1$, and reverse causality is not an issue. The coefficients β_2 and β_3 estimate the contemporaneous and lagged effects of VC investments.

1.2.2 Occupation-Level Analysis

The main unobserved heterogeneity not controlled for in fixed effects models is omitted variables at the industry $\times$ CZ $\times$ year level. For example, a technological shock in the IT industry of Atlanta in 2014 boosts wages in the local IT industry. At the same time, this shock attracts more VC investments in the local IT industry, resulting in higher total VC investments. Because the shock is industry-CZ specific, by definition, it should not directly affect wages in other industries. However, it can affect wages of related workers in other industries whose skills are demanded by the IT industry (e.g., programmers in the financial industry), through the competition for talent channel.

The richness of the worker-level data allows me to observe the occupations of workers. Motivated by the example discussed above, I analyze the impact of VC investments in the IT industry on talented computer and mathematical workers in non-IT industries.[8] The computer and mathematical occupation is one of the high-skilled occupations, with a mean annual wage of \$89,810 in 2017, [9] and the occupation is highly demanded by IT startups.

For each major occupation in an establishment, I compute the 90^{th} percentile of wages. There are 22 major occupations defined by OES. To quantify the effect of VC investments in IT industries on talented computer workers in non-IT industries, I estimate the following regression,

$$\Delta Y_{e,o,c,t} = \beta IT\_VCPerCapita_{c,t} \times Computer_o + \eta_{e,o} + \gamma_{e,t} + \theta_{o,t} + \epsilon_{e,o,c,t}, \qquad (1.5)$$

where $\Delta Y_{e,o,c,t}$ is the dollar change in the 90^{th} percentile of wages ($\Delta P90\_EstOcc$) for occupation o in establishment e, in CZ c, and year t. The variable $IT\_VCPerCapita_{c,t}$ is the dollar value of VC investments in the IT industry of CZ c in year y divided by the CZ population in 2000. The variable $Computer_o$ is equal to 1 if occupation o is computer and mathematical (BLS OES Occupation Code: 15-0000) and 0 otherwise. The inclusion of establishment $\times$ year ($\gamma_{e,t}$) and occupation $\times$ year ($\theta_{o,t}$) fixed effects absorbs any time-varying unobservables at the establishment and occupation level. Talented non-computer workers within the same establishment serve as a control group. A positive β would suggest that VC investments in the IT industry have a positive effect on the wages of talented com-

[8]Top two 2-digit NAICS industries based on the employment of computer and mathematical workers are removed from the analysis, including *information* (51) and *professional, scientific, and technical services* (54).

[9]See https://www.bls.gov/oes/current/oes150000.htm.

puter workers in the non-IT industries. This both mitigates the concern about time-varying, industry-CZ specific shocks and also supports the competition for talent hypothesis.

1.2.3 Exploiting Variation in VC Funding

VC investments are determined by both VC supply and venture capitalists' decision to invest. The former may be less of a concern, but the latter is subject to a selection issue. That's why I exploit a VC supply boom around 2014. Also, I construct two instrumental variables for VC investments.

As shown in Figure 1.3, VC investments jump from \$30 billion in 2013 to \$51 billion in 2014, then remain at the same level afterward. One explanation is that large institutional investors (e.g., mutual funds, hedge funds, private equities, corporations) invest in the VC market to reach for yields as a result of the low-interest rate environment.[10] This explanation is consistent with the finding in [34] that mutual fund holdings of unicorns increase from less than \$1 billion in 2013 to about \$8 billion 2015.

I exploit different exposures to the VC boom. Similar to the idea of the Bartik instrument ([16]), I compute the share of national VC investments in the pre-boom period (1990-2006) for all CZs.[11] I argue that CZs with a higher pre-boom VC share are more likely to be affected by the VC supply shock than CZs with a lower VC share. One concern with this approach is that the pre-boom VC share may not fully capture the actual exposure to the VC supply shock. If anything, the estimation based on the VC share should be the lower bound of the effect.

Specifically, I apply a difference-in-differences estimation to quantify the wage effect of the VC supply shock:

$$\Delta Y_{e,c,t} = \beta Post_t * VCShare\_Boom_c + \eta_e + \gamma_t + \epsilon_{e,c,t}. \tag{1.6}$$

[10] See https://pitchbook.com/news/articles/the-supply-demand-economics-behind-the-current-vc-boom-and-crunch.

[11] To avoid the financial crisis period, the pre-boom period ends in 2006.

I include 4-year data from 2012 to 2015 for CZs with nonzero VC share. The variable $Post_t$ is equal to 1 for 2014 and 2015, and 0 for other years. The variable $VCShare\_Boom_c$ is equal to 1 if CZ c is among the top 20% CZs based on VC share, and 0 otherwise.

I also construct two instrumental variables for VC investments, which are used in the literature. The first instrumental variable represents the distance-weighted portfolio returns of limited partners ([96]). The second instrumental variable represents the distance-weighted inflows into buyout funds ([53]). As argued in these papers, these two instrumental variables are uncorrelated with the investment opportunity sets of venture capitalists. Therefore, they should not be directly correlated with wage growth in the local area.

Although VC investments are not random, the inclusion of establishment, industry $\times$ state $\times$ year, and firm $\times$ year fixed effects absorbs most of the unobserved heterogeneity. Moreover, the evidence from occupation-level analysis and from exploiting plausibly exogenous variation in VC funding gives credence to the argument that there is a causal effect of VC investments on the wages of local incumbent establishments.

1.3 Data

I use five main datasets for the empirical analysis: a) anonymized employer–employee matched payroll data and demographics data from Equifax Inc., b) venture capital investments data from Thomson Financial's VentureXpert database, c) disclosure data of the labor condition application (LCA) program from the Department of Labor (DOL), d) Metropolitan statistical area (MSA) occupation-level employment and wage estimates from the Occupational Employment Statistics (OES) survey conducted by the Bureau of Labor Statistics (BLS), e) firm-level financial information from Standard & Poor's Compustat database. I describe each dataset and data construction in this section.

1.3.1 Worker Data

I obtain a comprehensive anonymized consumer dataset from Equifax Inc., one of the three major credit bureaus. The data contain detailed employment information, including company name, 3-digit NAICS code, the date an employee was most recently hired for the current position, and job title. I standardize approximately one million raw job titles in the data and match each raw job title to one of the 22 major occupations defined by the OES using crosswalks provided by O*NET.[12] for occupation classifications. The data also provide rich payroll information that includes the payment structure by which payments are made to the employee (e.g., hourly, weekly, bi-weekly, semi-monthly, monthly, annually) and the wage rate, which are used to compute annual wages. All wage data are converted to December 2017 dollars using the seasonally adjusted consumer price index (CPI) for all urban consumers from the BLS. I match income and employment data to demographics data that include the ZIP code of residence and education level. I link a worker's ZIP code of residence to a CZ, and I use the ZIP code as a proxy for the CZ of the worker's workplace. [106] computes that 92.5% of U.S. workers live in the CZ where they work. Therefore, the proxy is acceptable.

The dataset contains 285 million anonymized employer–employee matched annual records composed of 70 million workers of 2,402 employers from 2008 to 2017. It is one of the first comprehensive dataset in the U.S. that matches employers, employees, education level, and occupation. Figure 1.1 compares the wage distribution of the sample to the wage distribution of the BLS OES data. The 90^{th} percentile in the sample is $94,351, which is close to the 90^{th} percentile in the OES data, $96,150. It appears that wage data for high-skilled workers, the focus of this paper, are representative. For other percentiles, wages are slightly lower in the dataset compared with the OES data. Figure A.1 and A.2 plot the occupation distribution and industry distribution of the sample.

The dataset includes education data for 42 million workers out of a total of 70 million

[12] See https://www.bls.gov/oes/current/oes_stru.htm

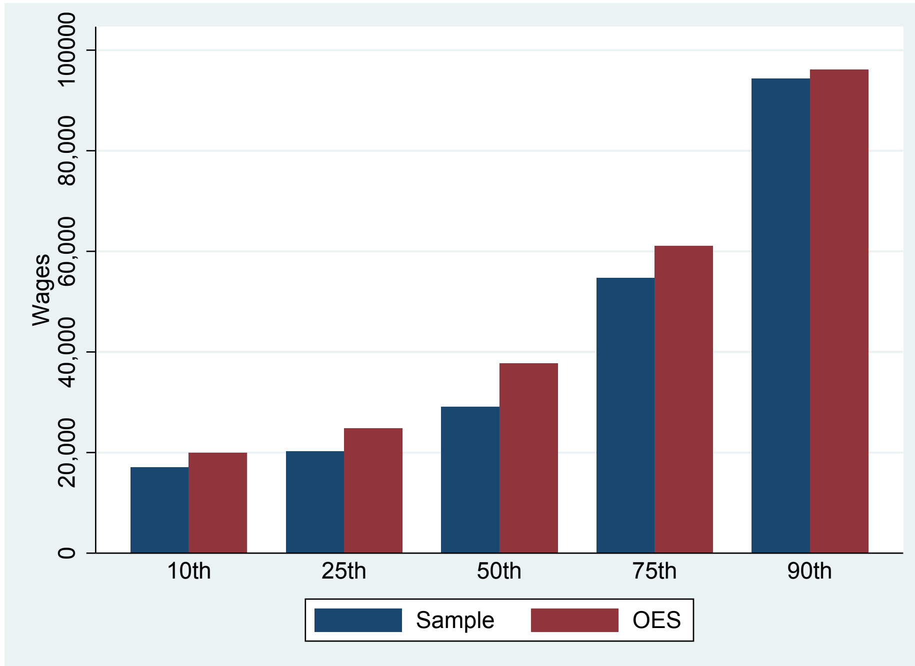


Figure 1.1: Wage Distribution
The figure presents the wage distribution of the work-level data in the sample and the wage distribution of all occupations in the 2017 BLS OES data.

workers. Education data indicate the highest level of education attained by any member of the household (e.g., some high school or less, high school, some college, college, graduate school). In this dataset, 40 million workers have at least a high school diploma. Figure 1.2 compares the educational attainment in the sample and the 2017 Census survey for people with at least a high school diploma. The distribution appears to be similar. One caveat for the education data is that they record highest education attained by any member of a household, so a worker's education may be overstated. In this paper, workers with at least a bachelor's degree are defined as high-skilled workers, while workers with only a high school diploma are defined as low-skilled workers.

I define an *establishment* as an employer in a CZ. To have enough observations to compute wage percentiles, I focus on establishments that have at least 50 employees in a year and have both high-skilled and low-skilled workers. I calculate percentiles for high-skilled workers' wages and low-skilled workers' wages. Further, I compute the changes

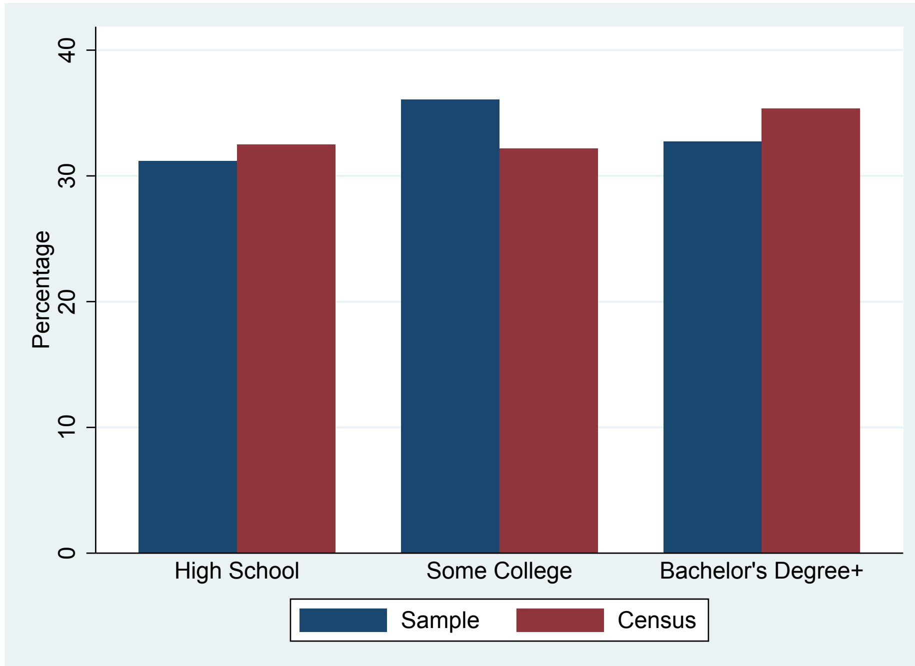


Figure 1.2: Education Distribution
The figure presents the education distribution in the sample and the education distribution in the 2017 Census survey for people who have at least a high school diploma.

in wage percentiles. For any major occupation in an establishment, I also calculate the percentiles of wages and the changes in these percentiles.

The granularity of the employer–employee–education–occupation matched dataset allows me to measure skills in multiple dimensions. I can define skills based on education and exploit heterogeneous effects based on wage percentiles within the same level of education. Also, I can exploit the worker heterogeneity in occupations within the same establishment. Finally, the data allow me to include fine-grained fixed effects to better control for unobservables at the industry, state, firm, and CZ level, and better estimate the coefficient of interest in Equation 1.1.

1.3.2 VC Data

I obtain VC investments data from Thomson Financial's VentureXpert database, one of the most commonly used databases for VC research. I focus on VC deals that are associated

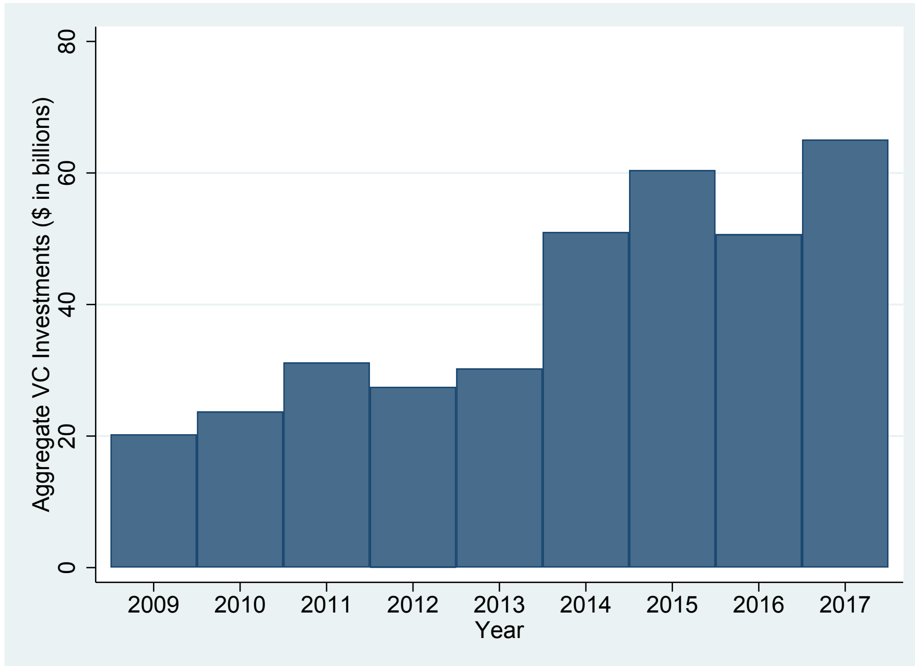


Figure 1.3: Aggregate VC Investments from 2009 to 2017
The figure presents aggregate VC investments (in billions of dollars) in the U.S. from 2009 to 2017.

with U.S. portfolio companies labeled as *Startup/Seed*, *Early Stage*, *Expansion*, or *Later Stage*. I do not restrict investors to VC funds because non-VC funds also invested as venture capitalists.[13] In each deal, I observe round date, round number, round amount disclosed, company name, company address, company SIC code, and information on VC firm and VC fund.

Figure 1.3 plots the aggregate VC investments in the U.S. from 2009 to 2017. The total investments in 2009 are about $20 billion and peak in 2017, approaching $65 billion. The correlations between the time series of annual total VC investments in my sample and the time series in industry reports (e.g., National Venture Capital Association, PwC/CB Insights quarterly MoneyTree Report) are more than 0.99, indicating that my dataset fully captures the aggregate patterns.

The main goal is to quantify the effect VC investments in a CZ on local incumbent

[13]As the most valuable startups in the U.S., Uber's investors include not only VC funds (e.g., Benchmark, Sequoia Capital) but also non-VC funds (e.g., BlackRock, Fidelity Investments).

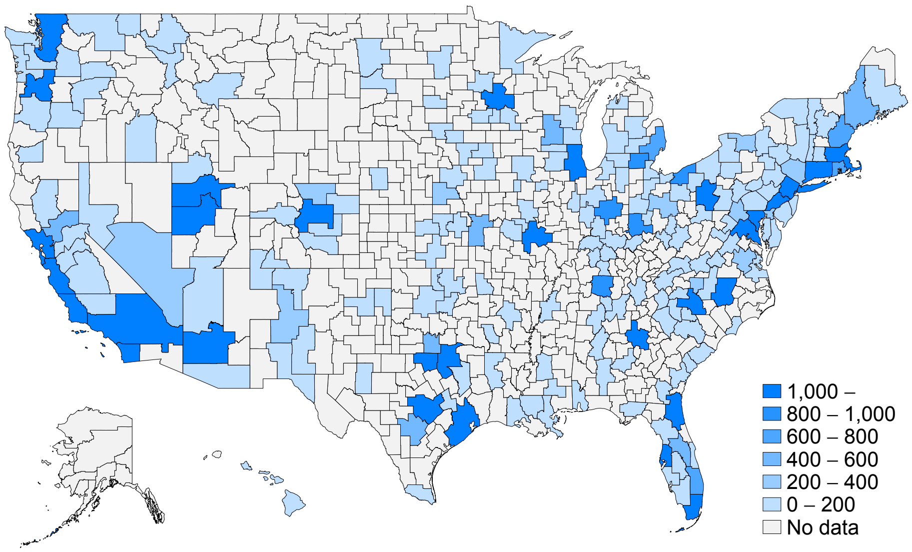


Figure 1.4: CZ-Level VC Investments from 2009 to 2017
The figure presents CZ-level total VC investments (in millions of dollars) from 2009 to 2017.

establishments; therefore, I match startup companies' ZIP codes to a CZ. 32,955 deals that provide information on the round amount disclosed can be matched to a CZ. For each CZ in each year, I compute the dollar value of total VC investments, which is converted to December 2017 dollars using CPI. Out of 709 CZs, 226 received VC investments over the sample period. These CZs account for approximately 82% of the U.S. population in 2000. Figure 1.4 plots the total CZ-level VC investments from 2009 to 2017 on a map. I scale the total annual investments by CZ population in 2000 to construct the main independent variable $VCPerCapita$ in Equation 1.1. CZs that received any VC investments from 2009 to 2017 are used for the main analysis.

Also, I acquire information on initial public offerings (IPOs), private equity (PE) investments, and horizontal mergers and acquisitions from Thomson Financial. These data are used to construct control variables.

1.3.3 LCA Data

To validate whether startups hire more and pay more upon receiving VC funding, I obtain labor condition application (LCA) data from 2010 to 2017 from the Department of Labor (DOL).[14] An LCA is a form employers must file with the DOL on behalf of employees who apply for an immigration H-1B work visa. Workers who apply for an H-1B work visa are mostly high-skilled workers. According to the DOL, the H-1B program allows employers to temporarily employ foreign workers in the U.S. on a nonimmigrant basis in specialty occupations.

Each LCA case contains detailed information on the submission date, employer name, employer address, employer NAICS code, number of foreign workers requested in the case, job title, and wage.[15] I drop employers who file LCAs for fewer than 10 worker from 2010 to 2017. For each employer in each month, I compute the total number of workers requested and the average wage.

1.3.4 Other Data

The Occupation Employment Statistics (OES) program is a semiannual survey conducted by the Bureau of Labor Statistics (BLS).[16] The OES provides annual employment and wage estimates for more than 800 occupations at the national, geographic, and industry level. More than 800 occupations are also grouped into 22 major occupations. I obtain data at the MSA-major occupation level. Each year, for each major occupation in a MSA, the data provide information on total employment and wage percentiles (i.e., 10^{th} percentile, 25^{th} percentile, 50^{th} percentile, 75^{th} percentile, 90^{th} percentile, the average wage). I obtain firms' financial data from Standard & Poor's Compustat database, and I compute relevant firm-level outcomes. Firms' ZIP codes are used to match firms to a CZ. Table 1.1 reports

[14] The sample starts in 2010 because the DOL changed the filing system. In the old system, no information on industry code is provided, which is important for creating the control group.

[15] Data can be downloaded from the DOL website: https://www.foreignlaborcert.doleta.gov/performancedata.cfm.

[16] Data can be downloaded from the BLS website: https://www.bls.gov/oes/.

Table 1.1: Summary Statistics

This table presents summary statistics for key variables. Panel A presents summary statistics for variables at the establishment level. Panel B presents summary statistics for variables at the establishment-occupation level. Panel C presents summary statistics for variables at the CZ level.

	N	Mean	Std Dev
Panel A: Establishment-Level Data			
P90_HighSkilled	337,032	92,537	67,626
ΔP90_HighSkilled	337,032	1,055	31,184
P10_LowSkilled	337,032	25,553	14,583
ΔP10_LowSkilled	337,032	628	6,039
ΔEmp_HighSkilled	337,032	-1.05	28.9
ΔEmp_LowSkilled	337,032	-0.5	27.6
Panel B: Establishment-Occupation-Level Data			
P90_EscOCC	2,310,136	79,233	56,048
ΔP90_EscOCC	2,310,136	616	21,718
Panel C: CZ-Level Data			
VCPerCapita	2,034	0.060	0.265

the summary statistics for key variables used in this paper.

1.4 Main Results

I present the main empirical results in this section. First, I demonstrate the effect of VC investments on wages at local incumbent establishments. I use three different approaches to establish causality. Second, I test several alternative hypotheses and conduct multiple robustness checks. Third, I exploit heterogeneity. Fourth, I show the effect of VC investments on establishment-level employment.

1.4.1 Do VC Investments Affect Wages of Incumbent Establishments?

Baseline Results

The main empirical objective is to quantify the effect of VC investments on the wages of incumbent establishments. It is likely that VC investments create more competition for the best high-skilled workers than competition for other high-skilled workers. Therefore, I focus on the top-end: the 90^{th} of high-skilled workers' wages. Table 1.2, Column (1) reports results estimating the effect on dollar change in the 90^{th} percentile of high-skilled workers' wages (ΔP90_HighSkilled) using Equation 1.1. VC investment of \$1,000 per capita is associated with \$1,218 (t-stat: 6.66) increase in the 90^{th} percentile of high-skilled workers' wages. Given that the unconditional mean of ΔP90_HighSkilled in the data is \$1,055, VC investment of \$1,000 per capita increases wage growth by 115% of its unconditional average.

As discussed in Section 1.2.1, to control for time-varying unobservables at the industry and state level, I include industry $\times$ year, state $\times$ year, and industry $\times$ state $\times$ year fixed effects in Column (2) to Column (4), respectively. The results are robust to the inclusion of these fixed effects. In Column (4), where industry $\times$ state $\times$ year fixed effects are included, the estimated effect is \$849 (t-stat: 4.27). It appears that industry-specific and state-specific shocks do not drive the results.

The establishment-level data also allow me to include firm $\times$ year fixed effects, which control for firm-level productivity increases or technological adoptions. The identification is achieved by comparing an establishment in a high-VC CZ to an establishment of the same firm in a low-VC CZ. The result in Column (5) suggests that VC investment of \$1,000 per capita implies a \$605 (t-stat: 3.06) increase in the 90^{th} percentile of high-skilled workers' wages. Note that the inclusion of firm $\times$ year fixed effects may underestimate the actual effect due to the wage convergence within firms documented by [99].

Further, I examine the effects of VC investments on other high-skilled workers by es-

Table 1.2: Effect of VC Investments on Wages of High-Skilled Workers at Incumbent Establishments

This table presents results of establishment-level panel regressions assessing the effect of VC investments on wages of high-skilled workers at incumbent establishments using Equations 1.1, 1.2, and 1.3. The dependent variable is the dollar change in the 90^{th} percentile of high-skilled workers' wages. VCPerCapita (in thousands of dollars) is the dollar value of VC investments in a CZ scaled by the population in 2000. The regressions include establishments in CZs that received any VC investments from 2009 to 2017. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

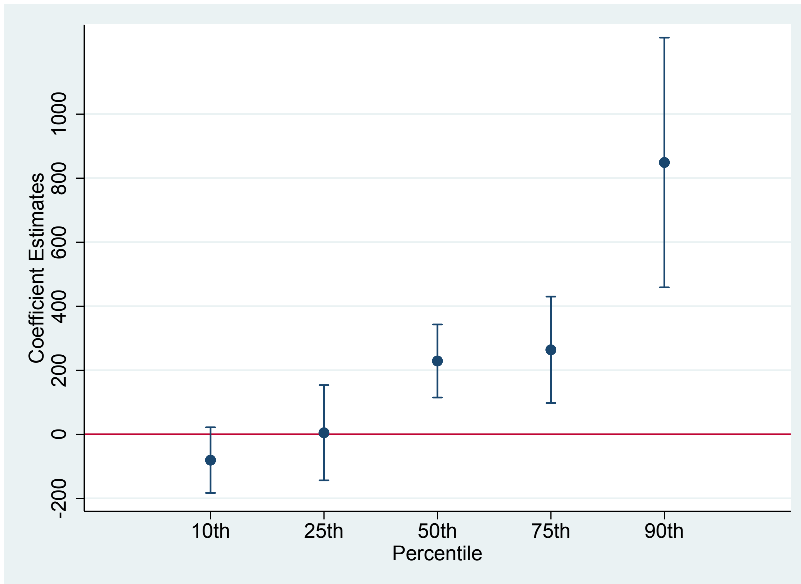
	ΔP90_HighSkilled				
	(1)	(2)	(3)	(4)	(5)
VCPerCapita	1,218***	929***	1,086***	849***	605***
	(183)	(199)	(239)	(199)	(198)
Establishment FE	✓	✓	✓	✓	✓
Year FE	✓				
Industry × Year FE		✓			
State × Year FE			✓		
Industry × State × Year FE				✓	
Firm × Year FE					✓
Observations	337,032	337,032	337,032	337,032	337,032
R-squared	0.107	0.151	0.108	0.217	0.507

timating Equation 1.2 using the dollar change in the 10^{th}, 25^{th}, 50^{th}, and 75^{th} percentile of high-skilled workers' wages as dependent variables. Figure 1.5, Panel A plots the coefficient estimates along with the 95% confidence intervals from five separate regressions. The figure indicates distributional effects among high-skilled workers. Namely, there is no effect on the 10^{th} and 25^{th} wage percentile, a modest positive effect on the 50^{th} and 75^{th} wage percentile, and a strong positive effect on the 90^{th} wage percentile. These distributional effects are consistent with the hypothesis that VC investments create more demand for the best high-skilled workers. Figure A.3 shows similar distribution effects when using log changes in wage percentiles as dependent variables. Distributional effects, to some extent, suggest that the results are less likely to be driven by time-varying shocks at the CZ level.

Ex ante, it is not clear how VC investments affect the wages of low-skilled workers. On the one hand, VC investments may not create significant direct demand for low-skilled workers, but on the other hand, the wages of low-skilled workers may increase because VC investments can generate positive spillovers to the local economy. Table 1.3 presents results assessing the effect of VC investments on the dollar change in the 10^{th} percentile of low-skilled workers' wages (ΔP10_LowSkilled). In Column (1), the coefficient estimate is 196 (*t*-stat: 3.10). In Column (4), where industry × state × year fixed effects are included, the coefficient estimate becomes -61.7 (*t*-stat: -1.06). The results suggest that there is no significant effect on the 10^{th} percentile of low-skilled workers' wages at incumbent establishments. I also examine distributional effects among low-skilled workers. Figure 1.5, Panel B shows that there is a marginally positive effect on the 90th percentile of low-skilled workers' wages, while there is no impact on other percentiles of low-skilled workers' wages.

As discussed in Section 1.2.1, reverse causality is another endogeneity concern. Namely, startups in regions with high wage growth seek VC funding. To mitigate this concern, I estimate the dynamic effect using Equation 1.4. Figure 1.6 plots the coefficient estimates along

Panel A: High-Skilled Workers



Panel B: Low-Skilled Wokrers

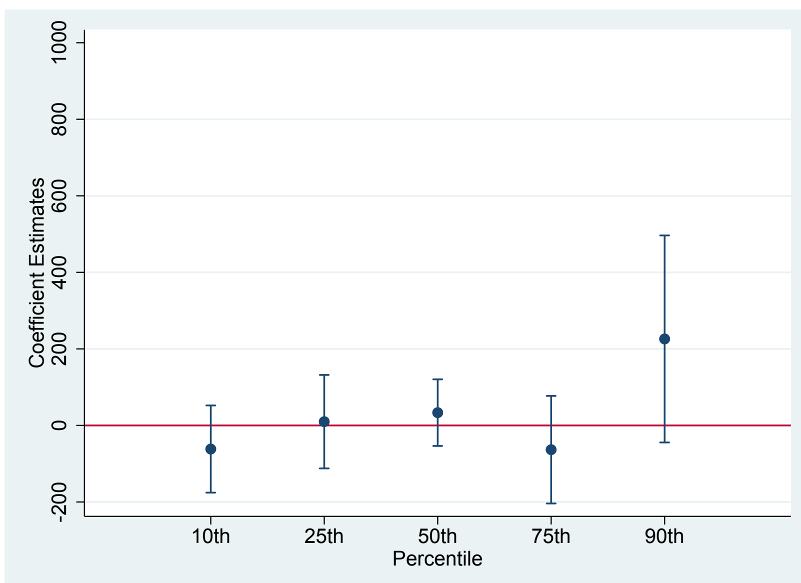


Figure 1.5: Distributional Effects of VC Investments on Wages of Incumbent Establishments
The figure presents distribution effects of VC investments on wages of incumbent establishments using Equation 1.2.

Table 1.3: Effect of VC Investments on Wages of Low-Skilled Workers at Incumbent Establishments

This table presents results of establishment-level panel regressions assessing the effect of VC investments on wages of low-skilled workers at incumbent establishments using Equations 1.1, 1.2, and 1.3. The dependent variable is the dollar change in the 10^{th} percentile of low-skilled workers' wages. VCPerCapita (in thousands of dollars) is the dollar value of VC investments in a CZ scaled by the population in 2000. The regressions include establishments in CZs that received any VC investments from 2009 to 2017. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	ΔP10_LowSkilled				
	(1)	(2)	(3)	(4)	(5)
VCPerCapita	196***	127**	-22.4	-61.7	56.4
	(63.2)	(54.6)	(56.3)	(58.1)	(47.1)
Establishment FE	✓	✓	✓	✓	✓
Year FE	✓				
Industry × Year FE		✓			
State × Year FE			✓		
Industry × State × Year FE				✓	
Firm × Year FE					✓
Observations	337,032	337,032	337,032	337,032	337,032
R-squared	0.164	0.194	0.165	0.260	0.437

with the 95% confidence intervals for two separate regressions for ΔP90_HighSkilled and ΔP10_LowSkilled. Panel A shows a significant contemporaneous effect of VC investments on the wages of high-skilled workers. The coefficient estimate on the VC investments in the next year is not statistically different from zero, suggesting that reverse causality is not a concern. There is a small reversal in wage growth one year after large VC investments. Panel B confirms this muted effect on the wages of low-skilled workers.

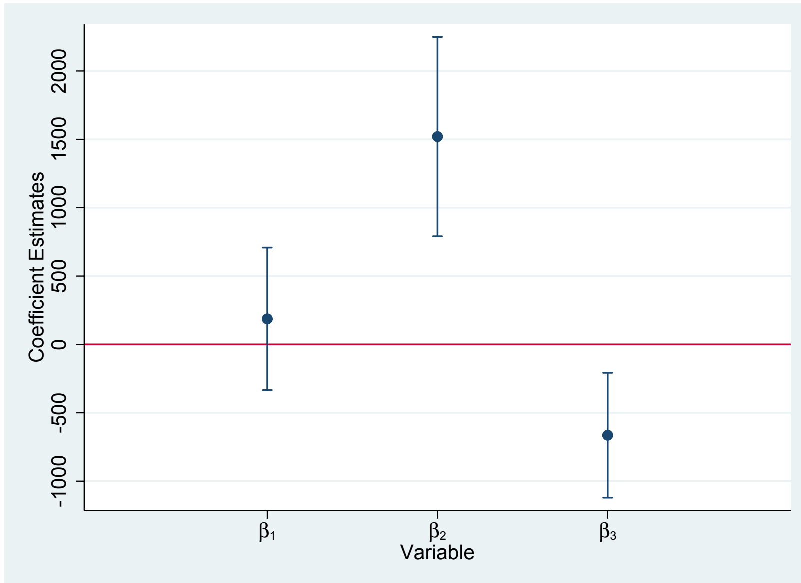
Table 1.2 and Table 1.3 indicate that VC investments have positive effects on the wages of high-skilled workers, but no impact on the wages of low-skilled workers. This asymmetric effect implies that the wage dispersion within an establishment would increase. I find evidence in Table A.1 supporting this hypothesis. VC investment of $1,000 per capita increases the gap between the 90^{th} percentile of high-skilled workers' wages and the 10^{th} percentile of low-skilled workers' wages by $492–$859.

Do Time-Varying Unobservables at the Industry-CZ Level Drive the Results: Evidence from Occupation-Level Analysis

As discussed in Section 1.2.2, the major unobservables that have not been controlled for in fixed effects models of Section 1.4.1 are the time-varying unobservables at the industry $\times$ CZ level. To alleviate this concern, I examine the impact of VC investments in an industry on related talented workers in other industries. Specifically, I analyze the effect of VC investments in the IT industry on talented computer workers in non-IT industries.

Table 1.4, Column (1) reports results from estimating Equation 1.5. It implies that VC investment of $1,000 per capita in the IT industry results in an $857 increase in the wages of talented computer workers compared to other talented workers in the same establishment of non-IT industries, supporting the competition for talent hypothesis. In unreported results, I find similar evidence that VC investments in the IT industry have an impact on the wages of computer workers using BLS OES data. Column (2) suggests that VC investments in non-IT industries do no generate the same effect. The explanation is that VC investments

Panel A: ΔP90_HighSkilled



Panel B: ΔP10_LowSkilled

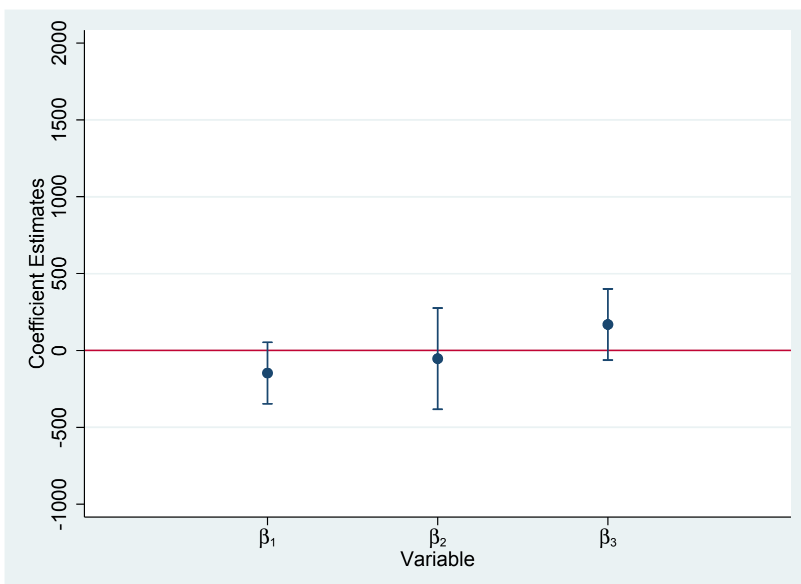


Figure 1.6: Dynamic Effect of VC Investments on Wages of Incumbent Establishments
The figure presents the dynamic effect of VC investments on wages of incumbent establishments using Equation 1.4.

Table 1.4: Effect of VC Investments in the IT Industry on Wages of Talented Computer Workers in Non-IT Industries

This table presents results of establishment-occupation-level panel regressions assessing the effect of VC investments in the IT industry on wages of talented computer workers in non-IT industries using Equation 1.5. The dependent variable is the dollar change in the 90^{th} percentile of wages for an occupation in an establishment. IT_VCPerCapita (Other_VCPerCapita) is the dollar value of VC investments in the IT industry (non-IT industries) in a CZ scaled by the population in 2000. Computer is equal to 1 if the occupation is computer and mathematical, and 0 otherwise. The regressions include establishments in CZs that received any VC investments from 2009 to 2017. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	ΔP90_EstOcc			
	All	All	High Tech	Low Tech
	(1)	(2)	(3)	(4)
IT_VCPerCapita × Computer	857**		1,596**	324
	(413)		(737)	(697)
Other_VCPerCapita × Computer		159		
		(319)		
Establishment × Occupation FE	✓	✓	✓	✓
Establishment × Year FE	✓	✓	✓	✓
Occupation × Year FE	✓	✓	✓	✓
Observations	2,310,136	2,310,136	696,039	1,614,097
R-squared	0.375	0.375	0.393	0.367

in non-IT industries may create demand for talented computer workers as well as talented workers in other occupations. The effect due to competition for talent should be stronger for talented computer workers whose skills are more transferable to IT startups. It is likely that talented computer workers in high-tech industries have more transferable skills than talented computer workers in low-tech industries. Therefore, the effect should be stronger for talented workers in non-IT high-tech industries.[17] Columns (3) and (4) find consistent results.

Do Time-Varying Unobservables at the Industry-CZ Level Drive the Results: Evidence from Exploiting Variation in VC Funding

VC investments are determined by a) VC supply; b) venture capitalists' decision to invest. The latter is subject to a selection issue, whereas the former is less of an issue. I exploit plausibly exogenous variation in VC funding as discussed in Section 1.2.3.

Table 1.5 presents results from estimating Equation 1.6. Column (1) indicates that high-skilled workers in boom counties, which are ex ante more exposed to VC investments, have \$2,352 higher wage growth. The results are robust to the inclusion of industry $\times$ year, state $\times$ year, and industry $\times$ state $\times$ year fixed effects. I plot the dynamic effect in Figure A.4. The difference in wage growth between boom CZs and non-boom CZs in 2012 is not significantly different from the difference in 2013. It appears that there is no pre-trend in the data, and the parallel trends assumption is valid. Compared to the difference in 2013, the difference increases to more than \$3,000 in 2014 and remains at a similar level in 2015.

Table A.2 reports results from IV regressions. Columns (1) and (2) use the distance-weighted portfolio returns of limited partners as an instrumental variable. Columns (3) and (4) use the distance-weighted inflows into buyout funds as an instrumental variable. Both results are consistent with the OLS results that VC investments result in higher wage

[17] The classification of high-tech industries follows [95]. I match 4-digit SIC codes used in their paper to 3-digit NAICS codes.

Table 1.5: Difference-in-Differences Estimation around 2014 VC Boom

This table presents results of a difference-in-differences estimation assessing the effect of the 2014 VC boom on wages of high-skilled workers at incumbent establishments using Equation 1.6. The dependent variable is the dollar change in the 90^{th} percentile of high-skilled workers' wages. Post is equal to 1 for 2014 and 2015, and 0 otherwise. VCShare_Boom is equal to 1 if the CZ is among the top 20% CZs based on VC share. Establishments in CZs that have positive VC share are included in the regressions. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	ΔP90_HighSkilled			
	(1)	(2)	(3)	(4)
Post × VCShare_Boom	2,352***	1,043***	2,178***	917**
	(382)	(365)	(412)	(387)
Establishment FE	✓	✓	✓	✓
Year FE	✓			
Industry × Year FE		✓		
State × Year FE			✓	
Industry × State × Year FE				✓
Observations	174,684	174,684	174,684	174,684
R-squared	0.171	0.216	0.172	0.275

growth for high-skilled workers at local incumbent establishments.

1.4.2 Alternative Hypotheses and Robustness Checks

Are VC Investments Related to Other Financial Activities that Affect Wages?

CZs that receive more VC investments may also have more IPOs. [22] show that firms have higher employment growth after IPOs. [14] find that IPOs have a positive effect on new firm creation. Further, [29] and [37] document that IPOs have an impact on local labor markets. To control for this potential IPO effect, I compute the per capita IPO proceeds (IPOPerCapita) for each CZ in each year. I include IPOPerCapita in Equation 1.2.

VC is one type of private equity (PE). Another primary PE are buyouts. VC investments may be correlated with buyout investments. [2] find that workers of firms acquired by PE investors have higher wages after the investment. Also, firms in my sample may be subject to mergers and acquisitions (M&A). [83] and [78] document that M&A has distributional effects on the wages of workers at target firms. To control for potential PE and M&A effects, I compute PE investments per capita (PEPerCapita) and M&A transaction values per capita (MAPerCapita) in each CZ each year, and I include these variables in Equation 1.2.

Conceptually, VC investments are different from these financial activities for two reasons. First, VC investments are concentrated on high-tech, high-growth startups. VC-backed startups are more likely to hire more high-skilled workers and create competition for talent. However, it is less likely that these financial activities affect labor outcomes through competition for talent. Second, the real effect of M&A and PE is likely to be a slow process, while the VC impact is contemporaneous. Empirically, Table 1.6 shows that coefficient estimates for VCPerCapita remain stable after controlling for IPOPerCapita, PEPerCapita, and MAPerCapita.

Table 1.6: Testing Alternative Hypotheses

This table presents results of establishment-level panel regressions assessing the effect of VC investments on wages of high-skilled workers at incumbent establishments to test alternative hypotheses using Equation 1.2. The dependent variable is the dollar change in the 90^{th} percentile of high-skilled workers' wages. VCPerCapita, IPOPerCapita, PEPerCapita, and MAPerCapita (in thousands of dollars) are the dollar value of VC investments, IPO proceeds, PE investments, and M&A transactions, respectively, in a CZ scaled by the population in 2000. EmpGrowth, WageGrowth, and EstGrowth are the CZ-level employment growth, wage growth, and establishment growth, respectively. The regressions include establishments in CZs that received any VC investments from 2009 to 2017. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	ΔP90_HighSkilled				
	(1)	(2)	(3)	(4)	(5)
VCPerCapita	769***	838***	853***	788***	847***
	(209)	(206)	(196)	(195)	(208)
IPOPerCapita	-349				
	(294)				
PEPerCapita		169*			
		(90.0)			
MAPerCapita			13.3		
			(11.5)		
EmpGrowth				-2,762	
				(5,488)	
WageGrowth				-3,544	
				(2,504)	
EstGrowth					-563
					(6,359)
Observations	337,032	337,032	337,032	337,032	337,032
R-squared	0.217	0.217	0.217	0.217	0.217
Establishment FE	✓	✓	✓	✓	✓
Industry × State × Year FE	✓	✓	✓	✓	✓

Are Results Driven by General Local Economic Conditions?

Are the results simply driven by good local labor markets? The asymmetric effect between high-skilled workers and low-skilled workers suggests that this may be not the case. To further alleviate this concern, I control for CZ-level wage growth and employment growth computed from the Quarterly Census of Employment and Wages (QCEW) data. Table 1.6, Column (4) shows that the coefficient estimate changes very little.

Further, it is likely that VC selects startups. Even without VC, startups can still grow and hire. The effect could be driven by the presence of more startups in the local area rather than VC investment flows. To control for the impact of local entrepreneurship activities, I construct CZ-level establishment growth from QECW. The coefficient estimate remains similar in Column (5). One caveat with this approach is that CZ-level establishment growth may not fully capture startup activities in high-tech industries.

Are Results Driven by Competition among Incumbent Establishments?

It is possible that incumbent establishments compete with each other for talent. Workers move from one incumbent establishment to another establishment and receive higher compensation. If VC investments are correlated with the intensity of competition for talent among incumbent establishments, the main results may be overestimated. But, the employment growth of high-skilled workers in incumbents should not change if workers move within incumbent establishments. Table 1.8 and Table 1.13 show that the employment growth of high-skilled workers decreases for establishments and firms headquartered in high-VC CZs. These results suggest it is unlikely that competition for talent among incumbent establishments explains the main results.

Additional Robustness Checks

To ensure that the effect of VC investments on the wages of high-skilled workers at incumbent establishments is robust, I conduct multiple robustness checks by varying the dependent variable, independent variable, clustering, and the sample, and I report results in Table 1.7. The results hold when I use the 90^{th} percentile, the log change in the 90^{th} percentile, or the dollar change in the average of high-skilled workers' wages as a dependent variable. The results are robust to adjusting VC investments by lagged CZ-level employment, clustering standard errors at the state–year level, including establishments from all CZs, and removing establishments from non-tradable sectors.[18]

As shown in Figure 1.4, some CZs receive more VC investments than other CZs. To mitigate the concern that a particular VC-heavy region drives the results in Table 1.2, I reestimate Equation 1.2 by removing one of the nine census divisions from the sample, and I plot the coefficient estimates along with the 95% confidence intervals from the nine separate regressions. Figure 1.7 shows that all coefficient estimates are statistically significant and the magnitudes are stable. The figure suggests that no particular region, including California, drives my results.

Also, I reestimate Equation 1.2 by interacting VCPerCapita with three size dummies (*Small, Medium, Large*). *Small* establishments are those with fewer than 250 employees. *Medium* establishments are those with at least 250 but fewer than 1,000 employees. *Large* establishments are those with at least 1,000 employees. Figure 1.8 shows that the coefficient estimates are positive and significant across all three establishment size groups.

Although VC investments are not random, it is reasonable to conclude that VC flows likely have a causal effect on wages at incumbent establishments, given (a) the inclusion of fine-grained fixed effects; (b) the cross-industry, occupation-level analysis; (c) the exploitation of plausibly exogenous variation in VC funding; (d) the lack of reverse causality; and (e) ruling out multiple alternative hypotheses.

[18] A non-tradable sector is defined based on [88]. It includes 3-digit NAICS industries 441, 442, 443, 445, 446, 447, 448, 451, 452, 453, and 722.

Table 1.7: Robustness Checks

This table presents robustness checks for establishment-level panel regressions assessing the effect of VC investments on wages of high-skilled workers at incumbent establishments using Equation 1.2 by varying the dependence variable, independent variable, clustering standard errors, and the sample. Standard errors are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Specifications	Coefficient Estimates
Baseline specification	849***
	(199)
90^{th} percentile	1,300***
	(445)
Log change in the 90^{th} percentile	0.006***
	(0.002)
Dollar change in the average	265***
	(62.9)
Adjust VC investments by lagged employment	431***
	(96.3)
Cluster standard errors by state-year	849**
	(420)
Include all CZs	886***
	(195)
Remove non-tradable sector	977***
	(168)
Establishment FE	✓
Industry $\times$ State $\times$ Year FE	✓

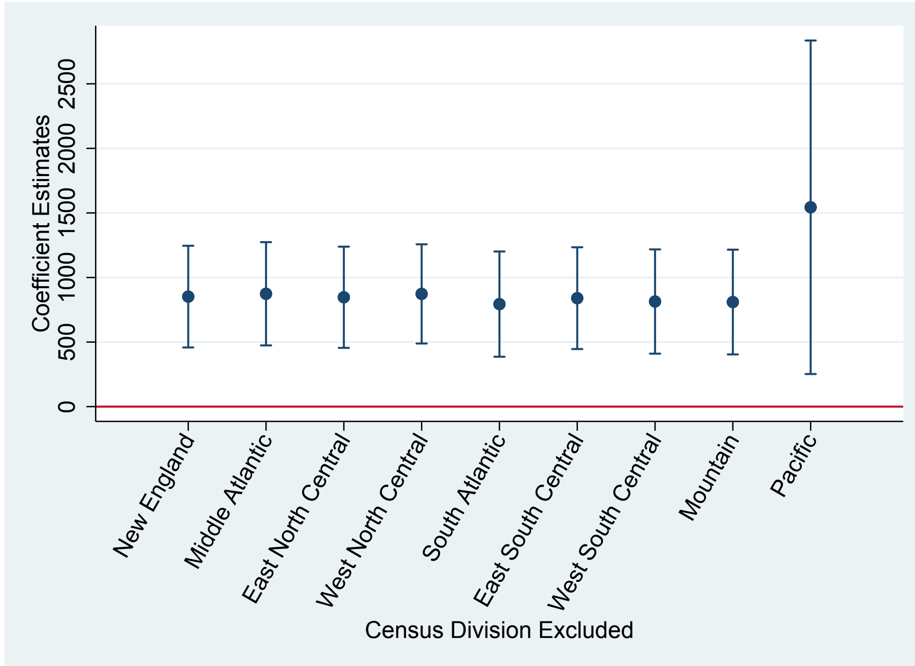


Figure 1.7: Coefficient Estimates after Excluding a Census Division
The figure presents coefficient estimates after excluding one of the nine census divisions using Equation 1.2.

1.4.3 Heterogeneity in Industry and VC

I first exploit heterogeneous effects based on the industry of incumbent establishments. High-tech industries receive almost two-thirds of VC financing ([95]). Also, high-skilled workers from high-tech incumbent establishments are more likely to join startups than high-skilled workers from low-tech incumbents because of skill transferability. Therefore, VC investments may have a larger impact on the wages of high-skilled workers in high-tech industries. To test this hypothesis, I interact VCPerCapita with a HighTech dummy and a LowTech dummy separately. HighTech (LowTech) is equal to 1 if the incumbent establishment is in high-tech (low-tech) industry based on its 3-digit NAICS code, and zero otherwise. Table A.3 shows results consistent with the hypothesis. Across all four specifications, VC investments have a larger impact on the wages of high-tech incumbent establishments.

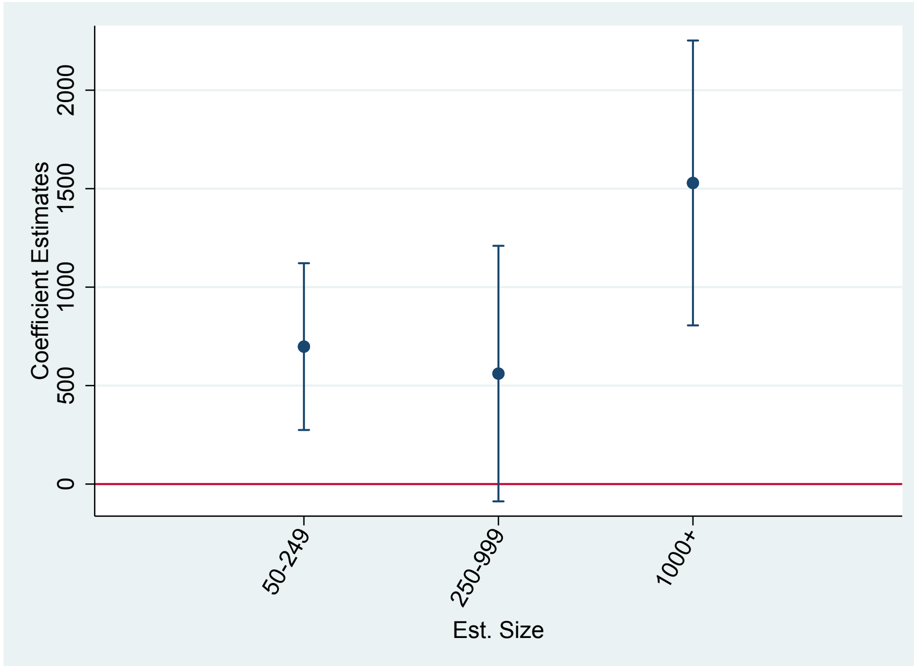


Figure 1.8: Coefficient Estimates by Establishment Size
The figure presents the coefficient estimates by establishment size (employment) using Equation 1.2.

In the main analysis, I aggregate all VC deals in the same CZ and the same year to compute VCPerCapita. It would be interesting to see which type of VC investments matters more. It is expected that VC investments in high-tech startups create more demand for talent. Therefore, I aggregate VC deals in high-tech industries and compute VC investments per capita (Hightech_VCPerCapita). Similarly, I construct Lowtech_VCPerCapita. It is possible that late-stage startups are more likely to expand the workforce because early-stage startups are more likely to focus on developing products. Also, late-stage startups are larger regarding the workforce; therefore, they should create a larger demand for talent upon receiving VC funding. I calculate per capita VC investments in early-stage startups (Early_VCPerCapita) and late-stage startups (Late_VCPerCapita). Consistent with the hypothesis, Table A.4 shows that VC investments in high-tech industries and late-stage startups have a more significant impact.

1.4.4 Do VC Investments Affect Establishment-Level Employment?

So far, I show that VC investments have positive effects on wages at incumbent establishments. The increase in wage prices due to VC investments could have a negative impact on establishment-level employment for two reasons. First, firms strategically reduce their employment of high-skilled workers in areas where competition for talent is high. Second, workers leave companies for VC-backed startups due to higher compensation, and incumbent establishments struggle to find workers to fill these vacant positions.

To test this hypothesis, for each establishment, I compute the annual change in the number of high-skilled workers (ΔEmp_HighSkilled) and the annual change in the number of low-skilled workers (ΔEmp_LowSkilled). I estimate Equation 1.1, and 1.2 by using ΔEmp_HighSkilled or ΔEmp_LowSkilled as a dependent variable. Table 1.8, Panel A reports results on ΔEmp_HighSkilled. In all four models, with various fixed effects, the coefficient estimates are negative and significant. VC investment of \$1,000 implies a decrease in high-skilled employment growth by 2.39–3.03 employees per establishment. Panel B reports results on ΔEmp_LowSkilled. None of the coefficient estimates are significant. These results are consistent with the results in Table 1.2 and Table 1.3, which show that VC investments create competition for high-skilled workers rather than for low-skilled workers.

Table 1.8 provides evidence that VC investments have a direct effect on the employment growth of high-skilled workers. For an establishment belonging to a multi-establishment firm, there might be an indirect effect in addition to the direct impact. Specifically, the employment growth of high-skilled workers may also be related to VC investments in other CZs where the firm operates. If firms make strategic decisions on hiring to avoid competition for talent from VC-backed startups, establishment-level employment growth of high-skilled workers may increase when VC investments are large in other CZs where the firm operates.

I construct an indirect exposure to VC investments to test the indirect effect for each es-

Table 1.8: Effect of VC Investments on Employment at Incumbent Establishments

This table presents results of establishment-level panel regressions assessing the effect of VC investments on the employment of incumbent establishments using Equations 1.1 and 1.2. The dependent variable in Panel A is the change in the employment of high-skilled workers. The dependent variable in Panel B is the change in the employment of low-skilled workers. VCPerCapita (in thousands of dollars) is the dollar value of VC investments in a CZ scaled by the population in 2000. The regressions include establishments in CZs that received any VC investments from 2009 to 2017. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Panel A: ΔEmp_HighSkilled			
	(1)	(2)	(3)	(4)
VCPerCapita	-3.030***	-2.630***	-2.690***	-2.390***
	(0.807)	(0.727)	(0.758)	(0.797)
Observations	337,032	337,032	337,032	337,032
R-squared	0.214	0.261	0.216	0.309
	Panel B: ΔEmp_LowSkilled			
	(1)	(2)	(3)	(4)
VCPerCapita	-0.262	-0.255	-0.514	-0.614
	(0.338)	(0.338)	(0.501)	(0.572)
Observations	337,032	337,032	337,032	337,032
R-squared	0.180	0.238	0.182	0.294
Establishment FE	✓	✓	✓	✓
Year FE	✓			
Industry × Year FE		✓		
State × Year FE			✓	
Industry × State × Year FE				✓

tablishment, following [51]. The measure is employment-weighted VCPerCapita (VCPerCapita_Other), using the employment of all other establishments belonging to the same firm and the VC investments in the corresponding CZs.

The results in Table 1.9 support the hypothesis. Column (1) indicates that the establishment-level employment growth of high-skilled workers is negatively related to its direct exposure to VC investments, but positively related to its indirect exposure to VC investments through the firm's internal network. In Column (2), I include industry $\times$ CZ $\times$ year fixed effects, which absorb VCPerCapita. These fixed effects also completely absorb any time-varying unobserved heterogeneity at the industry $\times$ CZ level. The identification comes by comparing establishments in the same CZ and in the same industry, but they have different indirect exposures to VC investments due to differences in firms' operating networks. The result still holds. For the non-tradable sector, VCPerCapita_Other may not capture competition for talent in other establishments because most workers are low-skilled workers. Columns (3) and (4) show that the indirect effect is concentrated on incumbent establishments in the tradable sector.

1.5 Discussion of the Mechanism

Section 1.4 establishes causal effects of VC investment flows on the wages of incumbent establishments. Specifically, the wages of high-skilled workers at incumbent establishments experience higher wage growth when VC investments in the CZ are higher. I also document that the employment growth of high-skilled workers at incumbent establishments decreases. This section discusses four non-mutually exclusive mechanisms through which VC investments affect the wages and employment of high-skilled workers at incumbent establishments.

Table 1.9: Firm's Internal Network and Effect of VC Investments on Employment
This table presents results of establishment-level panel regressions assessing the indirect effect of VC investments on the employment of high-skilled workers at incumbent establishments spread through a firm's internal network. The dependent variable is the change in the employment of high-skilled workers. VCPerCapita (in thousands of dollars) is the dollar value of VC investments in a CZ scaled by the population in 2000. For an establishment, VCPerCapita_Other is the employment-weighted VCPerCapita using the employment of all other establishments belonging to the same firm and the VCPerCapita in the corresponding CZs. The regressions include establishments in CZs that received any VC investments from 2009 to 2017. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	ΔEmp_HighSkilled			
	All	All	Tradable	Non-tradable
	(1)	(2)	(3)	(4)
VCPerCapita	-3.040***			
	(0.806)			
VCPerCapita_Other	4.410***	10.600***	12.700***	-0.314
	(1.570)	(2.650)	(3.050)	(3.200)
Establishment FE	✓	✓	✓	✓
Year FE	✓			
Industry × CZ × Year FE		✓	✓	✓
Observations	335,868	335,868	216,776	119,092
R-squared	0.215	0.369	0.388	0.319

1.5.1 Competition for Talent

The first potential mechanism is competition for talent. Especially, startups hire more and pay more upon receiving VC funding, which creates local competition for talent. If there is a shortage of talent supply in the local area, incumbent establishments must increase wages to hire or retain talent. In this section, I first validate the critical assumption that startups hire more and pay more upon receiving VC funding. Then, I exploit several labor market frictions which amplify the competition effect.

It is challenging to validate this assumption because hiring data for VC-backed startups are not publicly available. [40] document that VC funding is positively associated with employment growth using a proprietary headcount dataset of startups that are mostly in Silicon Valley. I complement their study by using publicly available LCA data to shed more light on the hiring of VC-backed startups. Any employer in the U.S. that seeks to hire foreign high-skilled workers using H-1B visas must file an LCA with the DOL. H-1B data has been used to measure firms' dependence on high-skilled workers and show the importance of high-skilled workers on innovation, firm value, and investments ([5], [98], [105]). One caveat of this dataset is that it does not provide information on the hiring of domestic high-skilled workers.

I identify all VC deals for startups that have received at least $100 million in VC funding from 2010 to 2017. A VC deal is an event. I conduct event studies by including windows 3 (6) months before and 3 (6) months after the events. To form a control group for a startup in an event, I identify all firms in the LCA dataset that are in the same 6-digit NAICS industry and the same CZ as the startup receiving the VC funding.[19] To quantify the impact of VC funding on startups' hiring, I estimate the following regression specification:

$$Y_{e,i,t} = \beta Post_{e,t} * Treated_{e,i} + \eta_{e,i} + \gamma_{e,t} + \epsilon_{e,i,t}, \tag{1.7}$$

[19] To reduce noise, firms must have filed an LCA for at least 10 workers from 2010 to 2017.

where $Y_{e,i,t}$ is the logarithm of 1 plus the total number of foreign workers being requested by LCA from firm i in month t. The variable e stands for an event. The variable $Post$ is equal to 1 after the VC deal, and 0 otherwise. The variable $Treated$ is equal to 1 if the startup receives funding in event e, and 0 otherwise. I include event $\times$ firm fixed effects ($\eta_{e,i}$) and event $\times$ month fixed effects ($\gamma_{e,t}$). Standard errors are clustered at the CZ level. The formation of the control group and the inclusion of event $\times$ month fixed effects control for time-varying unobservables at the CZ–industry level, which are correlated with employment growth.

Table 1.10 reports the results. Column (1) shows the results on hiring using a 6-month window around the events. VC-backed startups hire 2.6% more employees after receiving VC funding. The effect is stronger using the 12-month window around the events. The magnitude increases to 4.7%, as shown in Column (2). The effect is more pronounced for large VC deals in which the amount invested is at least \$100 million. Large deals result in a 15.1% increase in hiring. I estimate the wage effect in Table A.5. Conditional on hiring, startups increase the average wage for new employees by 0.9% after receiving VC funding. Given that the average wage in the data is approximately \$90,000, a 0.9% increase is equivalent to \$810. These results suggest that startups hire more after receiving VC funding, which could lead to an increase in local competition for talent.

The increase in wage growth and the decrease in employment growth of high-skilled workers are consistent with the competition for talent channel. Several labor market frictions may play an important role in the wage effect. First, if there is enough local talent supply (e.g., new graduates), which absorbs the demand by VC-backed startups, the wage effect on incumbent establishments should be negligible. Therefore, the effect should be stronger in regions with limited talent supply. I test this hypothesis by analyzing the differential effect on states with top-ten U.S. engineering universities and states without.[20]

[20]States with top-ten engineering universities include Massachusetts, California, Michigan, Pennsylvania, Indiana, Georgia, Illinois. See https://www.usnews.com/best-graduate-schools/top-engineering-schools/eng-rankings.

Table 1.10: Effect of VC Funding on Startup Hiring

This table presents results of event studies assessing the effect of VC funding on startup hiring using Equation 1.7. A VC deal is an event. The dependent variable is the logarithm of 1 plus the total number of foreign workers being requested by LCA. Post is equal to 1 after the VC deal, and 0 otherwise. Treated is equal to 1 for startups receiving funding in the VC deal, and 0 otherwise. Column (1) includes windows 3 months before and 3 months after the events. Columns (2) to (4) include windows 6 months before and 6 months after the events. Columns (1) and (2) include all VC deals. Column (3) includes VC deals with amount invested at least \$100 million. Column (4) includes VC deals with amount invested less than \$100 million. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	ln(1 + Hiring)			
Window	[-3,3]	[-6,6]	[-6,6]	[-6,6]
VC Deals	All	All	>= 100 M	< 100 M
	(1)	(2)	(3)	(4)
Post × Treated	0.026***	0.047***	0.151***	0.033***
	(0.010)	(0.010)	(0.027)	(0.008)
Event × Firm FE	✓	✓	✓	✓
Event × Month FE	✓	✓	✓	✓
Observations	1,055,332	2,087,905	293,956	1,793,949
R-squared	0.669	0.617	0.615	0.617

Table 1.11, Column (1) shows that VC investment of $1,000 per capita results in $727 increase for high-skilled workers in states with top engineering schools, while the effect is $1,823 in states without top engineering schools.

Even though there is limited local talent supply, VC investment flows should not generate regional differences in wage growth if cross-region labor mobility is perfect according to the spatial equilibrium model of [90]. However, cross-region labor mobility may be imperfect, and specific regions are less attractive to high-skilled workers. It is likely that the same amount of VC investment flows generate more wage pressure for incumbent establishments in CZs that are less appealing to high-skilled workers. Following [42], I construct a proxy for a city's appeal to high-skilled workers using weather days. I count the number of pleasant days for each CZ and assign CZs to two groups based on the median value. CZs with more pleasant days are more attractive to high-skilled workers. A day is considered a pleasant day if: a) a maximum temperature no higher than 85 degrees F; b) a minimum temperature no lower than 45 degrees F; c) a mean temperature between 55 and 75 degrees F; d) no measurable precipitation.[21] Table 1.11, Column (2) indicates that the wage effect of VC investment flows in less attractive CZs nearly doubles the wage effect in more attractive CZs.

Moreover, the demand for talent by VC-backed startups increases the bargaining power of skilled workers. The bargaining power is especially strong when talented workers can move across companies freely. [49] and [71] document that stronger enforceability of non-compete agreements (NC) results in lower labor mobility. Therefore, workers in states that have low enforceability of NC should have more bargaining power and higher wage growth when competition for talent is higher. I obtain measures for the strength of enforcing NC from [21]. I group states into low-enforceability and high-enforceability states based on the median value. Table 1.11, Column (3) finds evidence supporting the hypothesis. The wage effect is only significant in states with lower enforceability of non-compete agreements.

[21] I acknowledge Kelly Norton for sharing the code in GitHub. See https://kellegous.com/j/2014/02/03/pleasant-places/ for visualization of the measure.

Table 1.11: Labor Market Frictions and Wage Effect

This table presents results of establishment-level panel regressions assessing the role of labor market frictions in the effect of VC investments on wages of high-skilled workers at incumbent establishments using Equation 1.2. The dependent variable is the dollar change in the 90^{th} percentile of high-skilled workers' wages. VCPerCapita (in thousands of dollars) is the dollar value of VC investments in a CZ scaled by the population in 2000. Top10Eng (NonTop10Eng) is a dummy variable indicating the establishment is in a state with (without) top-ten engineering universities. LowPleasant (HighPleasant) is a dummy variable indicating the establishment is in a CZ with a low (high) number of pleasant days based on the median value. LowNC (HighNC) is a dummy variable indicating the establishment is in a state with low (high) enforceability of non-compete agreement based on the median value. The regressions include establishments in CZs that received any VC investments from 2009 to 2017. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	ΔP90_HighSkilled		
	(1)	(2)	(3)
VCPerCapita × NonTop10Eng	1,823**		
	(748)		
VCPerCapita × Top10Eng	727***		
	(242)		
VCPerCapita × LowPleasant		1,741**	
		(871)	
VCPerCapita × HighPleasant		785***	
		(224)	
VCPerCapita × LowNC			866***
			(202)
VCPerCapita × HighNC			479
			(968)
Observations	337,032	304,075	337,032
R-squared	0.217	0.214	0.217
Establishment FE	✓	✓	✓
Industry × State × Year FE	✓	✓	✓

Overall, these results highlight that labor market frictions play an important role in the wage effect and further support competition for talent hypothesis.

1.5.2 Voluntary Departure for Entrepreneurship

It is plausible that some workers at incumbent establishments observe the opportunities provided by VC and decide to leave their jobs and become entrepreneurs. Voluntary departure can result in a reduction in the employment growth of established firms. Firms must increase wages to retain existing workers and hire new workers. However, in Table 1.6, the results are robust to the inclusion of CZ-level establishment growth. CZ-level establishment growth captures entrepreneurship activities, including new startups founded by workers who voluntarily leave incumbent establishments. One caveat for this approach is CZ-level establishment growth may not accurately capture the entrepreneurship activities of high-skilled workers. It also reflects the growth in low-tech industries.

1.5.3 Creation of New Skills

Many VC-backed startups are innovative and create technologies that require new skills. Local incumbent establishments may follow this innovation and create new jobs that require new skills. Since these skills are new and scarce in the labor market, new hires for these jobs are of better quality and are well paid. On the other hand, it is hard for existing workers to adopt these new skills in the short run. The creation of new skills channel predicts that there is a change in the composition of workers, and VC investments should have an impact on the wages of new high-skilled workers rather than existing workers.

To test this hypothesis, I separate workers into *new workers* and *existing workers*. New workers are those who join the establishment in the current calendar year. Existing workers are those who have joined the establishment at least one calendar year earlier. I compute the dollar changes in the 90^{th} percentile of wages of existing high-skilled workers (ΔP90_HighSkilled_Exisiting) and the dollar changes in the 90^{th} percentile of wages of

Table 1.12: Effect of VC on Wages of Existing and New High-Skilled Workers at Incumbent Establishments

This table presents results of establishment-level panel regressions assessing the effect of VC investments on the wages of existing high-skilled workers and the wages of new high-skilled workers at incumbent establishments using Equation 1.1. Δ90_HighSkilled_Existing is the dollar change in the 90^{th} percentile of existing high-skilled workers' wages. ΔP90_LowSkilled_New is the dollar change in the 90^{th} percentile of new high-skilled workers' wages. VCPerCapita (in thousand $) is the dollar value of VC investments in a CZ scaled by the population in 2000. The regressions include establishments in CZs that received any VC investments from 2009 to 2017. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	ΔP90_HighSkilled_Existing	ΔP90_HighSkilled_New
	(1)	(2)
VCPerCapita	823**	1,456***
	(417)	(300)
Observations	128,697	125,118
R-squared	0.307	0.283
Establishment FE	✓	✓
Industry $\times$ State $\times$ Year FE	✓	✓

new high-skilled workers (ΔP90_HighSkilled_New). To ensure there are enough observations to compute ΔP90_HighSkilled_New, in this study, I focus on medium and large establishments, which have no fewer than 250 workers. I estimate Equation 1.2 to analyze the possible differential effect. Table 1.12 shows that $1,000 per capita VC investments increase the 90^{th} percentile of existing high-skilled workers' wages by $823 and increase the 90^{th} percentile of new high-skilled workers' wages by $1,456. The results indicate that VC investments have an impact on both existing and new high-skilled workers. Therefore, the creation of new skills channel cannot completely explain the findings.

1.5.4 Knowledge Spillovers

[97] document that VC-backed startups generate knowledge spillovers. Higher wage growth for high-skilled workers in CZ that have more VC investments may represent productivity increases due to knowledge spillovers. However, the prediction of the knowledge spillovers channel on employment is mixed. On the one hand, a wage increase reflects higher productivity. Firms have no incentive to reduce employment. On the other hand, since productivity increases, firms demand less labor to maintain their output. Further, this channel cannot explain the small reversal in wage growth, the reductions in R&D and patent quality documented in Section 1.6.

Overall, these four mechanisms are not mutually exclusive, and each one may play a role in explaining the results. However, the competition for talent channel can best explain the results. To further disentangle the channels, data on productivity and entrepreneurship are required.

1.6 Implications for Firms

1.6.1 Firm-Level Labor Costs and Employment

Section 1.4.4 suggests that firms may strategically make hiring decisions. However, it is unclear whether firms can reallocate workers from high-VC CZs to low-VC CZs to completely avoid the competition for talent by VC-backed startups, so that the impact on firms would be muted. If they cannot completely avoid this, competition for talent by VC-backed startups may result in higher labor costs and lower employment at incumbent firms. I use Compustat data to test this hypothesis. Ideally, for each firm, I can compute the exposure to VC investments using the locations of all the establishments. However, these data are not available for all Compustat firms. Instead, I use VC investments in the CZ where the firm is headquartered as a proxy for the firm's exposure to VC investments. I estimate the

Table 1.13: Effect of VC Investments on Firm-Level Outcomes

This table presents results of firm-level panel regressions assessing the effect of VC investments on firm-level outcomes. SGA/Sale is selling, general and administrative expense (XSGA) divided by sales (SALE). Emp_G is employment growth. Inv/AT is capital expenditure (CAPX) divided by lagged total assets (AT). RD/AT is research and development expense (XRD) divided by lagged total assets (AT). VCPerCapita (in thousands of dollars) is the dollar value of VC investments in a CZ scaled by the population in 2000. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	SGA/Sale	Emp_G	Inv/AT	RD/AT
	(1)	(2)	(3)	(4)
VCPerCapita	0.016**	-0.014***	-0.000	-0.008***
	(2.08)	(-2.89)	(-0.11)	(-3.26)
Firm FE	✓	✓	✓	✓
Industry × Year FE	✓	✓	✓	✓
State × Year FE	✓	✓	✓	✓
Observations	29,864	27,743	29,137	15,051
R-squared	0.853	0.375	0.802	0.871

following firm-level panel regression:

$$Y_{i,j,c,s,t} = \beta VCPerCapita_{c,t} + \eta_i + \gamma_{j,t} + \theta_{s,t} + \epsilon_{i,j,c,s,t}, \tag{1.8}$$

where $Y_{i,j,c,s,t}$ are firm-level outcomes of firm i in industry j, headquartered in CZ c and state s, in year t. Labor cost (SGA/Sale) is selling, general and administrative expense (XSGA) divided by sales (SALE). Employment growth (Emp_G) is the current employment (EMP) divided by the lagged employment minus 1. I include firm fixed effects (η_i), industry $\times$ year fixed effects ($\gamma_{j,t}$), and state $\times$ year fixed effects ($\theta_{s,t}$) to control for time-invariant, firm-specific characteristics and time-varying industry-level and state-level shocks. Standard errors are clustered at the CZ level.

Table 1.13, Columns (1) and (2) show that VC investments in CZs where incumbent firms are headquartered imply an increase in labor costs and a decrease in employment growth for incumbent firms. These results are consistent with the establishment-level results, and they suggest that firms cannot completely avoid competition from VC-backed startups.

1.6.2 Innovation

Next, I examine the implications for firms' investment policies. Competition for talent could affect incumbents' investments through two channels. First, the increased labor costs negatively affect firms' cash flows. Second, competition for talent tighten labor constraints for some incumbent firms. I examine two investment outcomes. Investment (CAPX/AT) is capital expenditure (CAPX) divided by lagged total assets (AT). R&D (RD/AT) represents research and development expenses (XRD) divided by lagged total assets (AT). Columns (3) and (4) suggest that firms reduce their R&D, but there is no change in capital expenditure. These results are broadly consistent with [5] and [105], who demonstrate that the reduction in skilled labor supply negatively affects firm investments in the setting of immigrants.

Last, I examine the innovation output. Following [91], I examine both the quantity and the quality of innovation by incumbent firms.[22] Since innovation is a slow process, I also include lagged VCPerCapita in Equation 1.8. Table 1.14 reports the results. Column (1) shows that incumbent firms file more patents. These results are consistent with [76] and [97], who show that VC spurs innovation. Surprisingly, Column (2) and Column (3) demonstrate that incumbent firms are more likely to file patents that are in the bottom 10% of the citation distribution than patents that are in the top 10% of the citation distribution. This indicates that the quality of new patents is not very high. Column (4) shows that the total adjusted citations do not increase. Column (5) indicates that average adjusted citations decrease. This evidence suggests incumbent firms generate lower-quality innovation after

[22] I greatly appreciate Manpreet Singh for generously sharing their data. The sample period is from 1990 to 2006.

Table 1.14: Effect of VC Investments on Firm Innovation

This table presents results of firm-level panel regressions assessing the effect of VC investments on firm innovation. ln(1 + patents) is the logarithm of 1 plus the number of patents filed by a firm in a year. ln(patents_bot10) is the logarithm of the number of patents that are in the bottom 10% of the citation distribution. ln(patents_top10) is the logarithm of the number of patents that are in the top 10% of the citation distribution. ln(1 + cites) is the logarithm of 1 plus the total adjusted citations. ln(1 + avg_cites) is the logarithm of 1 plus the average adjusted citations. VCPerCapita (in thousands of dollars) is the dollar value of VC investments in a CZ scaled by the population in 2000. Lag_VCPerCapita is lagged VCPerCapita. ln(AT) is the logarithm of lagged total assets. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	ln(1 + patents)	ln(patents_bot10)	ln(patents_top10)	ln(1 + cites)	ln(1 + avg_cites)
	(1)	(2)	(3)	(4)	(5)
VCPerCapita	0.019***	0.032***	0.002	0.030	0.007
	(3.07)	(2.73)	(0.35)	(1.45)	(0.45)
Lag_VCPerCapita	0.027***	0.039***	0.019***	-0.016	-0.037***
	(3.67)	(5.86)	(2.61)	(-1.33)	(-5.44)
ln(AT)	0.106***	0.187***	0.110***	0.118***	0.038**
	(5.61)	(9.29)	(5.07)	(3.84)	(2.31)
Firm FE	✓	✓	✓	✓	✓
Industry × Year FE	✓	✓	✓	✓	✓
State × Year FE	✓	✓	✓	✓	✓
Observations	42,902	14,330	14,330	42,902	42,902
R-squared	0.817	0.826	0.829	0.729	0.616

higher VC investment flows. Both the increase in the patent filing and the decrease in patent quality may be explained by the patent thicket that incumbent firms use to defend against competitors ([108]).

1.6.3 Patent Inventors Turnover

So far, I show both the input and output of innovation at incumbent firms decrease in regions with greater VC flows. Next, I examine the impact of VC flows on patent inventors turnover at incumbent firms, because investors are critical in generating innovation. If VC flows create competition for talent, it is likely that patent inventors leave incumbent firms to join VC-backed startups and it is difficult for incumbent to fill the vacant positions. In Table 1.15, I find that incumbent firms in regions with greater VC flows have higher departure rate of patent inventors. Although hey also hire at a higher rate, the magnitude is smaller than the departure rate. It results in a net leave of patent inventors. These results are consistent with the finding that employment growth declines, and they can also partly explain the decline in innovation quality.

1.7 Conclusion

In this paper, I study the impact of VC investment flows in a region on wages, employment, and innovation at large local incumbent firms that do not receive VC funding. I document that VC investments affect local competition for talent, resulting in higher (insignificant) wage growth for high-skilled (low-skilled) workers at incumbent establishments. Among high-skilled workers, there is a distributional effect, with higher-wage workers benefiting more. It is always challenging to establish causality. However, multiple tests, including exploiting within-industry-state variation and within-firm variation, occupation-level analysis, difference-in-differences estimation using the 2014 VC boom, and IV estimation all suggest that the effect is likely to be causal. I also find that labor market frictions, including imperfect labor mobility, amplify the wage effect. This result highlights the importance of reducing labor market frictions.

Further, I show that employment growth at incumbent firms declines, and patent inventors are more likely to leave. Subsequently, incumbent firms cut R&D and generate low-quality innovation. The evidence suggests that VC may relax labor and finance constraints for startups while tightening labor constraints for incumbent firms. This result highlights that labor constraints may cause some adverse consequences for some of the local incumbent firms.

My results should not be interpreted as suggesting that VC investments do not benefit the recipients or the local area. Indeed, as shown extensively in the literature, VC

Table 1.15: Effect of VC Investments on Patent Inventors Turnover

This table presents results of firm-level panel regressions assessing the effect of VC investments on patent inventors turnover. Leavers is the number of inventors who have produced a patent at the sample firm within one past year or a year before but produce at least one patent at a different firm, including inventors who produce their last patent in the sample firm. Hires is the number of inventors who produce at least one patent at the sample firm after producing a patent at a different firm within 1 year and a year after, including inventors who file their first patent with the sample firm. Net leavers is the difference between leavers and hires. VCPerCapita (in thousands of dollars) is the dollar value of VC investments in a CZ scaled by the population in 2000. Lag_VCPerCapita is lagged VCPerCapita. ln(AT) is the logarithm of lagged total assets. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	ln(1 + leavers)	ln(1 + hires)	ln(1 + net leavers)
	(1)	(2)	(3)
VCPerCapita	0.021***	0.006*	0.016***
	(6.96)	(1.88)	(7.26)
Lag_VCPerCapita	0.023***	0.021***	0.015***
	(5.20)	(4.29)	(5.05)
ln(AT)	0.057***	0.030***	0.043***
	(5.06)	(4.42)	(4.76)
Firm FE	✓	✓	✓
Industry × Year FE	✓	✓	✓
State × Year FE	✓	✓	✓
Observations	42,902	42,902	42,211
R-squared	0.668	0.619	0.605

investments create massive benefits for society. However, the costs of VC remain relatively unknown. This paper sheds light on the costs borne by some of local incumbent firms in the short run. In doing so, I am able to provide some evidence to quantify the aggregate effect. The results may have implications for regions that are aggressively trying to attract VC flows for startups ([76]).

CHAPTER 2

THE DARK SIDE OF TECHNOLOGICAL PROGRESS? IMPACT OF E-COMMERCE ON EMPLOYEES AT BRICK-AND-MORTAR RETAILERS

2.1 Introduction

Technological advances can create enormous economic benefits for society. But, technological progress and automation can also reshape and transform some labor markets. They can change the way some tasks are conducted, and these changes can augment the productivity of some workers but replace other workers entirely.[1] In this paper, we study the impact of e-commerce, which is one manifestation of technological advances in the retail sector. In particular, we study the effect of e-commerce on the employees of traditional brick-and-mortar retail stores.

The retail sector is a major employer in the U.S., employing approximately 16 million workers, or 13% of private sector employment, at the end of 2016 (Bureau of Labor Statistics). The retail industry landscape has changed dramatically in the last few decades. In the earlier decades, the disruption was mainly driven by the expansion of major retail chains such as Walmart and by the rise of discount retailers ([72, 62, 17, 93]). The recent disruption in the retail sector, including closures of thousands of stores, is attributed to the rise of technology led by e-commerce.[2] Yet, e-commerce retail, despite experiencing tremendous growth over the last few years, still accounts for only 9.1% of total U.S. retail sales in 2018Q3.[3] The effect of e-commerce on workers is ambiguous as on the one hand, brick-and-mortar retailers may focus on increased customer service as a result of an increase in

[1] See [92, 1, 12, 26, 27, 7, 9].

[2] See https://on.wsj.com/2poCwtG. Further, some online retailers have highly automated warehouses that use robots to bring items for a retail order from their storage shelves. The world's largest e-commerce retailer employed $45,000$ robots in its fulfillment centers, a 50% increase from the previous year's holiday season. https://bit.ly/2EldAGf

[3] https://fred.stlouisfed.org/series/ECOMPCTSA

competition from e-commerce retailers and consequently rely more heavily on employees. On the other hand, facing lower sales, affected stores may not only cut wages but also adjust both their employment composition and levels. As such, the retail sector provides an economically important setting for studying the impact of technological progress, as represented by e-commerce, on the large labor force of the traditional retail sector.

Identifying the causal impact of e-commerce on the employees of traditional brick-and-mortar retailers is challenging, since we cannot observe the counterfactual, i.e., what would the income and employment of brick-and-mortar retailer employees be in the absence of e-commerce? We address this identification challenge by using the staggered rollout of the fulfillment centers (*FCs*) of a major e-commerce retailer across U.S. from 1997 to 2016. We use this rollout as a proxy for the presence of local e-commerce. Our empirical strategy estimates the causal impact of the establishment of a new FC on the income and employment of retailer workers in that county and in neighboring counties.

We use a matched employer-employee payroll dataset for 57 major retail firms that employ a total of approximately 2.6 million retail workers. This comprises 18% of total U.S. retail employment in the first quarter of 2010. Our rich payroll information from a major credit bureau allows us to group workers into hourly workers and non-hourly workers (referred to as *salaried workers*). These data include total compensation, wage/salary, overtime, bonuses, commissions, and wage/salary rate.

We analyze this matched employer-employee dataset, and we use a difference-in-differences setting to exploit the staggered introduction of the FCs. We find that the labor income of retail workers in counties with FCs, on average, decreases by 2.4% after the establishment of FCs. This negative effect is also significant for workers within 50 or 100 miles of FCs. These results are confined to hourly workers, who experience a decrease in labor income by 2.5%, which is equivalent to an $825 decrease in annual income. Most of the effect derives from a reduction in the number of hours worked. Among hourly workers, we find a particularly strong negative impact on part-time hourly workers and, both young and old workers experiencing a sharper drop in labor income. Further, among the affected workers, those who have a prior higher credit card utilization and those who are otherwise more financially vulnerable experience higher credit card delinquencies and a subsequent decline in their credit score.

One potential concern with our identification strategy and our results may be whether (and why) the e-commerce retailer's FCs matter for traditional brick-and-mortar store sales. At the beginning of 2000, the e-commerce retailer had only three FCs, but the staggered introduction of FCs across different counties resulted in the retailer having more than 90 FCs by the end of 2016 (Figure 2.1 and Figure 2.2). Optimizing and expanding the FC network is an important strategy used by the e-commerce retailer to meet customer demand and save costs. The establishment of FCs allows the e-commerce retailer to optimize and distribute their inventory placement (even for third party sellers). This in turn allows the e-commerce retailer to reduce its shipping costs and shipping times. The e-commerce retailer can offer same-day or 2-day shipping for a longer time during the shopping day, making a purchase from the online retailer attractive.[4]

Moreover, the e-commerce retailer does not collect local sales taxes and has only recently begun to collect state sales tax on sales from its inventory ([18]). However, it does not yet collect state sales tax from most of its third party sellers (that account for more than 60% of the retailer's sales). This could lead to a price advantage over the traditional retailers. The establishment of an FC is meant to avoid long-zone shipping, and therefore is likely to have a more significant effect on geographically nearby areas, a fact that we exploit in our identification strategy.

Consistent with aforementioned arguments, we find that the establishment of the e-commerce retailer's FCs impact the sales of geographically proximate traditional brick-and-mortar retail stores. Using sales data from the National Establishment Time Series (NETS), we find that after the establishment of the e-commerce retailer's FC, the annual

[4][64] analyze the various trade-offs the e-commerce retailer faces in establishing an FC.

sales of stores decrease by \$63,639 per store, i.e., 2.8% of the total annual sales of the average store in our sample. We find that the annual sales of stores in the top tercile, based on sales one year before the FC, decrease by \$200,389.

A second potential concern regarding our identification strategy and our results is that the decision of establishing an FC in a county is probably not random. The decision may in fact be correlated with local economic conditions. As the goal is to better serve customers in surrounding areas, FCs are more likely to be built close to, or in, areas with high retail sales and population density. Our fixed effect methodology can control for these level differences. However, it is possible that the e-commerce retailer may choose to locate its FCs in areas with decreasing competition from brick-and-mortar retail stores, i.e., areas where the retail sales of brick-and-mortar stores are declining.

We formally test this hypothesis using differences in the demographics and retail sales data computed from 2000 and 2010 county-level census data. We find that the change in population density is the only change that is significantly and positively related to an FC establishment, while growth in the unemployment rate, median household income, and age distribution do not correlate with the location choice of FCs. Further, the positive coefficient on retail sales growth gives us confidence that our estimation may not be driven by a downward trend in the traditional retail sector. We also find a lack of pre-trends before the establishment of the FCs, lending credence to our identification strategy.

The inclusion of state-year-quarter fixed effects mitigates potential concerns about unobservable local economic conditions being the driver of our results. We further add more granular county-year-quarter fixed effects to control for local economic conditions by using the data available for workers of non-retail firms as a control group. Moreover, the opening of FCs has no impact on the sales of full-service restaurants in the county, supporting the view that our results are not due to negative local economic conditions.

Another potential concern is that there may be an omitted firm-specific shock to the traditional retailers that is contemporaneous with the establishment of the e-commerce retailer's FC in that county. For example, some firms may have concentrated operations in certain areas, and they may face firm-specific negative shocks (e.g., the failure of a major lender, the bankruptcy of a supplying firm) at the same time the e-commerce retailer builds an FC in that county. To address this concern, we include firm$\times$year-quarter fixed effects in our regressions and our results remain similar.

Due to the increased competition from the e-commerce retailer, the affected stores may focus on customer service and rely more on employees. Alternately, facing lower sales, affected stores may not only cut wages but also adjust their overall employment level. We find that for all stores, employment decreases by 2.4%, equivalent to 41 fewer workers per 100 stores for a store with an average of 22 employees. For large stores, employment decreases by 1 worker per store for a store with an average of 40 employees.

Using NETS data, we find that store closure rate increases by 3%. The average exit rate in our sample is 13.6%. Small stores and young stores are more likely to exit than large stores and older stores. After the establishment of FCs in a county, the entry rate for small stores reduces significantly by 11.8%. This low entry rate is not just limited to counties with FCs, but also exists for counties in 50 or 100 miles of FCs.

Finally, using publicly available Quarterly Census of Employment & Wages (QCEW) provided by BLS, we find that the establishment of an FC has a negative effect on the employment growth in the retail sector. We find a 2.9% decline in employment growth which implies a loss of 938 jobs per county per quarter. While, on the other hand, the establishment of FCs do create 256 jobs in transportation and warehousing sector. Further, we find a mildly significant positive spillover effect in restaurants, which leads to 143 more jobs per county per quarter.[5] Further, we find that counties within 50/100 miles of FCs, the positive effect on transportation and warehousing sector and restaurants disappears while the negative effect on the retail sector diminishes but remains negative and statistically

[5]During 2010-2016, an average county with fulfillment center employs 32,373 workers in the retail sector, about 12,155 workers in transportation and warehousing, and nearly 28,589 workers in restaurants and accommodation.

significant. We find similar results for aggregate wage growth.

Our paper is related to the literature that documents the effect of technological changes on labor market ([77, 11, 1, 12, 7, 9]). We add to the literature on the income and employment effect of competition in the context of expansion of large chain stores ([17, 93]) and import competition ([8, 10]). The rise of e-commerce and the consequent increase in competition for brick-and-mortar retail stores is an important technological change for the retail sector. We analyze the impact on the income and employment of affected retail stores using the establishment of a major e-commerce retailer's FCs as a proxy for an increase in the local competition.

Our paper also relates to research on the causes and consequences of disruption in the retail sector. Existing research focuses on the disruption that results from the rollout of large chains. [62] estimates the benefits and costs for the rollout of Walmart store openings. [72] quantifies the effect of the expansion of retail chain stores on other retailers. [17] and [93] estimate the employment and earnings effect as a result of Walmart store openings. We contribute to this literature by investigating the disruption attributed to e-commerce.

Compared to brick-and-mortar retailers, e-commerce provides consumers a lower price ([28]), and increased product variety ([25, 50]). The introduction of e-commerce by a firm increases its market value and revenue ([102, 94]). We add to this literature by analyzing the labor market consequences of a major e-commerce retailer's expansion of its FC network.

However, our results are limited to documenting only one facet of the impact of the technological innovation in the retail sector. There are many positive benefits of e-commerce for consumers, including potentially lower prices, more choices, convenience in shopping, gains from competition, and lower effort (e.g., no driving). Moreover, the establishment of the FC may have positive spillovers in the local community and could increase employment. However, the scope of our paper is limited to the impact on the workers in the geographically proximate traditional retail stores.

In the absence of labor market frictions (i.e., workers can easily switch jobs and their skills are completely transferable), the short-term displacement of some traditional retail store workers may not matter for the workers or to the local economy. However, in the presence of labor market frictions, the short-term impact can be negative. Moreover, to the extent that the scope of work differs between traditional retail stores and FCs, at least some workers can be worse off. For example, there is no need for cashiers at an FC, and skills may not be completely transferable.

In the long run, some of the affected workers may find alternate employment in their same field, or they may acquire new skills to find employment in another field. The scope of our paper is limited, and we focus only the short-term impact of the establishment of the FCs of the e-commerce retailer. Our results highlight one negative consequence of technological progress in the retail sector: the short-term negative impact on the wages and employment of some traditional brick-and-mortar retail store employees.

The rest of the paper proceeds as follows. We discuss our empirical methodology and identification threats in Section 2.2. In Section 2.3, we describe our data and summary statistics. Our empirical results are presented in Section 2.4. We conclude in Section 2.5.

2.2 Empirical Design and Identification Challenges

2.2.1 Empirical Design

In this study, we seek to examine the impact of technological progress, in the form of the expansion of e-commerce, on traditional brick-and-mortal retail establishments and their workers. In order to isolate the effects of the establishment of fulfillment centers (FCs) from other regional, sectoral, and macro-level shocks, we exploit the staggered rollout of FCs to capture the increase in the presence of local e-commerce. While the location that the major e-commerce retailer chooses for its FCs is certainly not random, we present evidence demonstrating that the timing of an FC establishment is plausibly orthogonal to unobserved factors that may impact local retail establishment performance. Specifically, our empirical strategy estimates the impact of the establishment of an FC in a county on retail workers in the same county and in neighboring counties.[6]

Why do FCs matter? First, the optimization and expansion of the FC network is important to the e-commerce retailer to meet customer demand and reduce costs. The e-commerce retailer built its first FC in 1997 and the number of FCs increased to over 90 by the end of 2016 (Figure 2.1 and Figure 2.2). Second, the establishment of a new FC is meant to avoid long-zone shipping – the FC is established to reduce the costs associated with shipping as well as the time required for customers to receive their packages. This is reflected in FCs being primarily located on the east and west coasts, where population density is highest.[7] Third, as discussed in [64], customers may value the convenience effect due to faster delivery, so that the establishment of an FC would induce customers nearby to be more willing to shop through the major e-commerce retailer rather than shop at a local brick-and-mortar store.

E-commerce retailing in principal allows for the ability to serve all potential customers on a national scale. However, in practice, the establishment of an FC will have a larger effect on areas surrounding the FC. As such, this increase in the value to local consumers of shopping from the major e-commerce retailer may reduce the attractiveness of local brick-and-mortar retail and thus negatively affect the income and employment of the employees in these brick-and-mortar stores.

Our empirical objective is to evaluate the local effect of the establishment of a new FC.[8] We do so by focusing on three definitions of *local*: 1) the focal county (i.e., the county where the FC opened) 2) all counties within 50 miles of the FC (excluding the focal county where the FC is located) and 3) all counties within 100 miles of the FC (again excluding

[6] A complete list of fulfillment centers of the e-commerce retailer is available at http://www.mwpvl.com/.

[7] The 2016 annual report of the e-commerce retailer states "If we do not adequately predict customer demand or otherwise optimize and operate our fulfillment network and data centers successfully, it could result in excess or insufficient fulfillment or data center capacity, or result in increased costs, impairment charges, or both, or harm our business in other ways."..."In addition, a failure to optimize inventory in our fulfillment network will increase our net shipping cost by requiring long-zone or partial shipments.". New FCs would be close to large cities, allowing for the possibility of next-day or same-day delivery and the wider rollout of its grocery business ([101]).

[8] We follow [64] and remove FCs that are established in a county with an existing FC or within about 20 miles of an existing FC, which reduces our sample to 50 FCs.

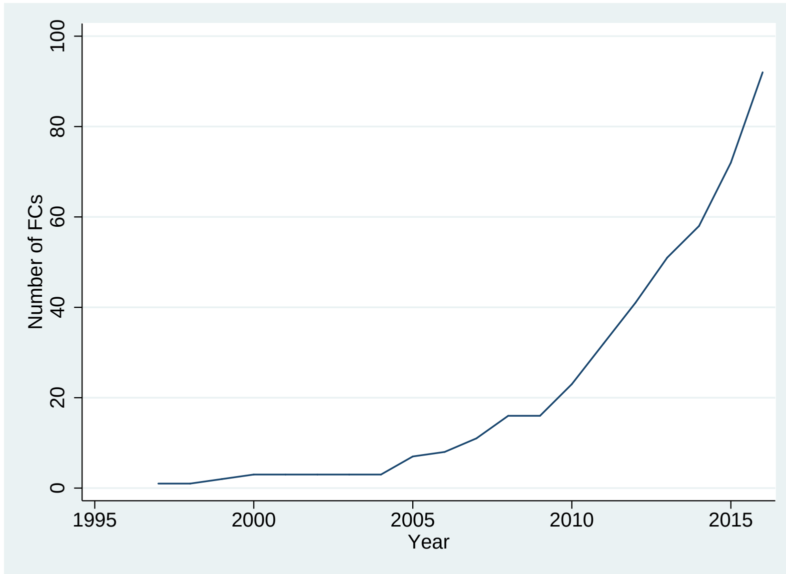


Figure 2.1: Number of E-Commerce Retailer's Fulfillment Centers
The figure plots the number of the major e-commerce retailer's fulfillment centers over time.

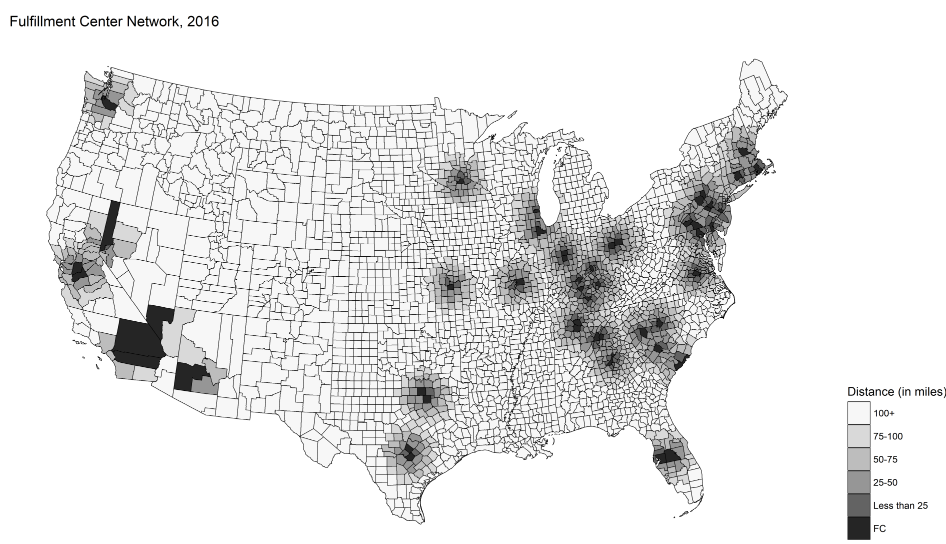


Figure 2.2: Major E-Commerce Retailer's FC Network
The map highlights the locations of the major e-commerce retailer's fulfillment centers. The dark regions highlight the counties with fulfillment centers, while the light regions highlight the neighboring counties.

the focal county where the FC is located).

We treat each county as *treated* in the first quarter in which an FC opens in one of the three definitions of a *local* county. For example, for the analysis that focuses on the focal county level, the indicator for Fulton County, GA, turns on in the first quarter of 2015 since an FC was opened in Union City, GA, (Fulton County) in February of 2015. In our 50-mile analysis, Cobb County, GA, (a county that abuts Fulton County, GA) is treated in the first quarter of 2015 from the opening of the same FC in Union City, GA. Yet, in our 100-mile level analysis, Cobb County, GA, is treated in the third quarter of 2011 due to the opening of an FC in Hamilton County, TN, in September of 2011. As our data start in 2010, our study focuses on the 39 FCs established after 2010.[9]

A standard approach for evaluating the impact of the opening of an FC would be to compare differences in brick-and-mortar establishment performance before and after the FC opening in treated and in untreated counties. For this difference-in-differences specification to yield unbiased estimates, parallel trends between the treated and control counties must be present. However, Table B.1 indicates that focal and surrounding counties where FCs open are very different not only in levels but also in trends from the rest of the U.S. in terms of demographics and local economic variables, such as population, population density, retail sales, retail sales per capita, household income, and unemployment rate. Therefore, these untreated counties may not serve as appropriate counterfactuals for our analysis. As a result, to ensure that both treated and control counties are on the common empirical support, we only include counties that were classified as *treated* at any time by the opening of an FC in our analysis, and we exploit the variation in the timing of the establishment of FCs, using FCs that will be treated but are not yet as de facto controls.

In our baseline analysis, we apply a difference-in-differences estimation to quantify the impact of the establishment of an FC on the income of workers in brick-and-mortar stores

[9] Our results remain robust to the inclusion of all 50 FCs (which includes FCs established before 2010) and is reported in Table 2.8.

by estimating the following:

$$\ln(Total\ Income_{i,c,t}) = \alpha + \beta PostFC_{c,t} + \eta_i + \theta_t + \epsilon_{i,c,t}, \tag{2.1}$$

where each quarterly observation is the income of worker i working in county c at time t. *PostFC* is an indicator that equals 1 in the quarter an FC is established in county c or within 50 or 100 miles of county c, and it remains 1 for all subsequent quarters. Time-invariant worker-specific characteristics and year-quarter shocks are controlled for with the inclusion of worker (η_i) and year-quarter (θ_t) fixed effects, respectively. Standard errors are clustered at the county level. The variable β estimates the percentage change in income attributed to the establishment of an FC.

2.2.2 Identification Challenges

In order for β from Equation 2.1 to represent an unbiased estimate of the impact of the establishment on an FC on the income of local brick-and-mortar retail workers, we must assume that *PostFC* is orthogonal to any unobservables contained within ϵ. Yet, because the location of FCs is not randomly decided by the major e-commerce retailer, dealing with this endogenous selection represents our main econometric challenge.

A primary concern in our analysis is that the decision to establish an FC in a specific county will naturally be a function of local economic conditions. Since one of the primary objectives of establishing FCs is to improve the ability to serve local customers, FCs may be more likely to be built in areas with high retail sales and high population density. To test this, on a cross-section consisting of all counties in the U.S., we regress the likelihood of establishing an FC in county c on county-level long differences (between 2000 and 2010) of retail sales, population density, unemployment rate, household income, and the percentage of the population between the ages of 18 and 65 (Table 2.1).

We observe that FCs are more likely to be located in counties with faster growing pop-

Table 2.1: Determinants of FC Locations: OLS

This table presents the results for determinants of fulfillment centers' location. The dependent variable is whether the county has an FC in the post-2010 period. The independent variables are long differences computed from 2000 and 2010 county-level census data.

	FC				
	(1)	(2)	(3)	(4)	(5)
Δ Log(Retail Sales)	0.012***	0.005	0.005	0.005	0.005
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Δ Log(Population Density (per sq mile))		0.114***	0.115***	0.115***	0.116***
		(0.025)	(0.025)	(0.025)	(0.025)
Δ Unemployment rate			-0.000	-0.000	-0.000
			(0.001)	(0.001)	(0.001)
Δ Log Median household income				-0.008	-0.008
				(0.019)	(0.019)
Δ Perc. age b/w 18 and 65					0.000
					(0.001)
Constant	0.013***	0.009***	0.009***	0.011**	0.010**
	(0.002)	(0.002)	(0.002)	(0.004)	(0.005)
Observations	3128	3109	3109	3109	3109
R^2	0.04	0.05	0.05	0.05	0.05

ulation densities; but, importantly, not those that have experienced more growth in retail sales. While our county-level fixed effects will absorb all time-invariant level effects of county-specific characteristics, our estimates will be biased downward insofar as counties that experience high population density increases are also more likely to engage in e-commerce transactions that substitute for local retail purchases. However, if this were true, we could also observe downward trends in retail sales and income of retail workers in the periods before the FC's establishment, which we do not (Figure 2.3).

A second concern surrounds the ability for the major e-commerce retailer to negotiate with the state and local government about the location of the FC in exchange for tax benefits or other incentives. It is plausible that governments may want the FC to be built in an area with weak economic conditions to boost the local employment and economy. As a result, FC county selection may be negatively correlated with local economic conditions; and, in particular, negatively correlated with the economic fortunes of brick-and-mortar retail stores, thus potentially biasing our estimates downward.

However, as shown in Table 2.1, Columns (3) and (4), the establishment of an FC does not relate to the ex ante change in the local unemployment rate and median household income, which lessens this concern. Furthermore, we deepen our analysis by conducting a placebo test on sales in another non-tradeable sector: full-service restaurants. As reported in Section 2.4.2, we do not find any effect on their sales with the establishment of FCs in the local area. Nonetheless, we control for any state-level time-varying unobservables, such as tax incentives, using state-year-quarter fixed effects.

Lastly, we may be concerned that the arrival of an FC in a county changes the composition of firms. In this case, higher performing firms or firms with greater options in managing their geographically-varied portfolio of stores choose to exit those counties. Or conversely, better performing retailers choose not to enter the treated counties. While this would lead to the appearance of a drop in retail performance, the effect would be entirely attributable to a compositional effect whereby the firm-quality distribution experiences a leftward shift. To deal with this, we focus solely on the intensive margin of competition by including only firms that are present before and after the arrival of an FC. In addition, we include establishment-level fixed effects to control for establishment-level quality and other time-invariant covariates of the establishment.

In summary, while establishing a causal relationship between FC openings and brick-and-mortar retail performance is not without its challenges, our confidence in interpreting our *PostFC* coefficient as causal is increased by the inclusion of fine-grained establishment, worker, industry-year, state-year, and year-quarter fixed effects. In addition, our confidence is further amplified by our ability to rule out the strongest threats to identification – reverse causality and selection on trends – by the lack of any discernible pre-trends in our regression and through the implementation of our full-service restaurants placebo test.

2.3 Data

Our empirical analysis makes use of data at three levels of analysis: 1) individual worker-level data that we obtain from a major credit bureau, 2) establishment-level data that we obtain from the National Establishments Time Series Database, and 3) county-level data that we obtain from the Bureau of Labor Statistics. We describe each dataset and its construction in this section.

2.3.1 Worker Data

Our novel comprehensive consumer data are provided by a major credit bureau. The data contain detailed employment information, including company name, 3-digit NAICS, the date an employee was most recently hired for the current position, an indicator of whether an employee is presently active, and rich payroll information that includes the payment structure by which payments are made to the employee, total compensation, wage/salary, overtime, bonuses, commissions, and wage/salary rate. We group workers into hourly workers and non-hourly workers (referred to as *salaried workers*) based on their payment structure.

We obtain income and employment data of active employees at the end of each quarter from the retail firms, which consistently supply data from 2010 to 2016.[10] The data are matched to credit files through tokenized Social Security Numbers (SSNs), which provide demographic information such as the individual's ZIP code of residence, age, and gender. We use the workers' county of residence to determine their location when examining the impact of the arrival of an FC.[11] Our sample consists of all workers employed in the first quarter of 2010, and the sample follows them until they exit. Our sample is thus unbalanced and does not allow for worker entry. In addition, we use the workers' residency in 2010Q1 to determine their location (county) for all empirical analysis. We drop workers who have multiple employers at any time during our sample period. If a worker switches employers during the sample period, we keep our observations for the first job. All dollar values are converted to December 2016 dollars using the seasonally adjusted consumer price index for all urban consumers from the Bureau of Labor Statistics.

Our sample contains 2.6 million workers from 57 retail firms, which accounts for 18% of the 14.42 million total U.S. retail employment in the first quarter of 2010. The median firm has more than 14,000 workers in the sample, which suggests that we cover mostly large firms. Table 2.2, Panel A presents summary statistics for worker-level payroll data. The mean quarterly income of hourly workers is $7,326. Annualized income is $29,304, which is slightly higher than the mean income of 8.79 million retail sales workers ($25,250) and slightly higher than the mean income of 4.53 million retail salespersons ($27,180), as estimated by the BLS in May 2016. The mean number of hours worked is 30.9 per

[10]We identify firms in 3-digit NAICS industries that are most likely to compete with the major e-commerce retailer's product catalog. The 3-digit NAICS codes that we classify as *retail* includes 442 (furniture and home furnishing stores), 443 (electronic and appliance stores), 444 (building material and garden equipment and supplies dealers), 448 (clothing and clothing accessories stores), 451 (sporting goods, hobby, book, and music stores), 452 (general merchandise stores), and 453 (miscellaneous store retailers). In robustness tests, we use workers from all non-retail firms as a control group.

[11]We do not observe the county of the workers' workplace. However, more than 90% of workers in our sample are hourly workers who are less likely to spend time and money on commuting to work. While we believe that a workers' residence does an adequate job of proxying for their workplace, we recognize that this will be measured with some error.

week, with an average wage rate of \$14.9 per hour. For retail workers in our sample, wage income contributes to about 87% of their total income. The remaining income derives from overtime/bonuses/commissions (referred to as a bonus). In our sample, salaried workers earn \$86,016 annually, on average.

The granularity of our worker-level data helps us answer questions that cannot be addressed solely from aggregate data. Given the fine-grained nature of these data, we can examine deeper worker-level heterogeneity and analyze which workers are more vulnerable to the establishment of an e-commerce FC. For example, are full-time workers more affected than part-time workers? Further, does a worker's gender and age insulate or exacerbate these effects? The detailed composition of the worker's compensation also allows us to understand the channels through which workers are affected. Do firms reduce workers wages or bonuses? Do firms cut wage rates or the number of hours worked? Lastly, these granular data allow us to improve the identification of our regression parameters through the inclusion of fine-grained fixed effects within a panel regression environment.

2.3.2 Establishment Data

In addition to worker-level data, we make use of establishment-level data for the retail sector from the National Establishment Time Series (NETS) Database (Walls & Associates, 2014).[12] This database provides an annual record for a large part of the U.S. economy that includes establishment job creation and destruction, sales growth performance, survivability of business startups, mobility patterns, changes in primary markets, corporate affiliations that highlight M&A, and historical D&B credit and payment ratings. At the beginning of our sample year, 2010, the database covers 3,287,183 active establishments employing 27,404,989 workers with total sales of \$2.9 trillion. These data are available until 2014.

Similar to how we defined retail firms with our worker-level dataset, we select estab-

[12]Walls & Associates converts Dun & Bradstreet (D&B) archival establishment data into a time-series database of establishment information.

Table 2.2: Summary Statistics

This table presents the summary statistics for the full sample and for counties with fulfillment centers. Panel A presents statistics for the quarterly worker-level data between 2010 and 2016 for retail workers. Panel B provides statistics for the annual sales and employment data at the establishment level for retail stores between 2010 and 2014. Panel C shows statistics for the quarterly county-industry level (2-digit NAICS) employment data between 2010 and 2016 for all industries. All dollar values are converted to December 2016 dollars using CPI from BLS.

	Full Sample			FC Counties		
	N	Mean	Std Dev	N	Mean	Std Dev
Panel A: Worker-Level Data						
Hourly Workers						
Total Income ($ per quarter)	34,806,676	7,326	4,197	1,881,184	8,248	4,363
Wage Income ($ per quarter)	34,689,666	6,362	3,553	1,875,176	7,162	3,693
Bonus ($ per quarter)	32,510,284	1,030	1,105	1,764,173	1,151	1,195
Hours Worked (per week)	34,489,187	30.9	10.1	1,866,683	32.7	9.65
Wage Rate ($ per hour)	34,489,187	14.9	4.54	1,866,683	16	4.69
Salaried Workers						
Total Income ($ per quarter)	5,438,083	21,504	18,271	293,960	22,710	18,335
Panel B: Establishment-Level Data						
Sales ($ 000s per year)						
All Stores	13,902,516	1,494.2	8,708.3	186,400	2,272.8	10,747.7
Small Stores	4,855,755	205.1	65.6	64,208	260.9	331.3
Medium Stores	4,413,654	461.8	102.9	59,372	617.4	677.8
Large Stores	4,633,087	3,828.8	14,800.0	62,820	5,893.8	17,954.7
Employment (workers per year)						
All Stores	13,902,516	12.3	63.0	186,400	17.0	68.8
Small Stores	4,855,755	3.7	3.8	64,208	4.0	7.5
Medium Stores	4,413,654	5.5	3.2	59,372	6.7	4.3
Large Stores	4,633,087	27.9	107.4	62,820	40.1	114.8
Panel C: County-Industry Level Data						
Employment (Avg. All Sectors)	1,837,920	1,692.6	8724.1	26,000	12,072.1	23,555.3
Retail Trade (NAICS 44-45)	91,896	4,664.297	14,836.41	1,400	32,373.0	37,766.3
Transportation & Warehousing & (NAICS 48-49)	91,896	1,212.8	5,471.0	1,400	12,155.6	14,117.1
Restaurants & Accommodation (NAICS 72)	91,896	3,695.5	13,805.2	1,400	28,589.6	43,897.2

lishments in 6-digit NAICS industries that are more likely to be affected based on the e-commerce retailer's product catalog. Table B.2 provides a complete list of industries selected. To reduce noise from very small retail stores, we keep retail stores with more than two employees before the establishment of the fulfillment center. Table 2.2, Panel B reports summary statistics for our sample. It shows that the average retail store in our sample has annual sales of approximately $1.5 million and 12 employees.

2.3.3 County Employment Data

For county-level analysis, we use publicly available Quarterly Census of Employment & Wages (QCEW) provided by BLS. The data provides county-level data on employment and total wages in each 2-digit NAICS industries for each quarter. Here, we keep all industries and counties to understand the aggregate effect on retail, warehouse/transportation and restaurants. We use quarterly data beginning first quarter of 2010 and ending fourth quarter of 2016 and report summary statistics in Table 2.2, Panel C. Note that in the counties with fulfillment centers, about 32,373 workers have employment in the retail sector while transportation and warehousing account for 12,155 workers per county.

2.4 Results

In this section, we describe our main empirical results. We first describe our baseline results using worker-level data. We then describe the robustness tests that we conduct to rule out competing interpretations of our results and to strengthen the identification of our parameters. Next, we describe results using NETS establishment data that allow us to analyze the impact of FC entry on the entry and exit of establishments in the local retail sector. Finally, we present the impact of FC establishments on the aggregate county-industry level employment using QCEW data from BLS.

2.4.1 How Do FCs Affect Local Brick-and-Mortar Stores? Evidence from Worker-level Data

Baseline Results

In Table 2.3, we report the impact of the establishment of FCs on the income of retail workers in counties with FCs or neighboring counties using the difference-in-differences specification shown in Equation 2.1. We include worker fixed effects and year-quarter fixed effects in all regressions in order to absorb as much variation as possible arising from worker-specific time-invariant characteristics and temporal trends. Since we define the arrival of an FC at the county level, we cluster all standard errors at the county level.

As shown in Panel A, Column (1), the total income of retail workers in counties with FCs decreases by 2.4%, on average, after the establishment of an FC. Moving to workers in counties within 50 or 100 miles of the focal county where an FC was established, we continue to observe a strong negative effect on total income (Panel B). Since the arrival of an FC may differentially impact hourly and salaried workers, we run separate regressions for those two types of workers.

Results in Column (2) show that the income of hourly workers decreases by 2.5%, equivalent to an $825 cut in annual income. As shown in Column (3), salaried workers mostly have muted responses to the establishment of FCs. These muted responses may be attributed partly to the infrequent adjustment of salaries or the inflexibility of firms in adjusting the incomes of salaried employees in the short term. We focus on hourly workers thoughout the rest of our analysis as they account for more than 90% of our sample and are the ones who experience the largest negative effects.

Our identification strategy, which relies on the staggered temporal rollout (shocks) of FCs across different counties, assumes that workers in counties that have yet to be treated by the establishment of an FC serve as an appropriate control group. This assumption would be violated if FCs are established in counties or regions that are experiencing upward

Table 2.3: Effect of FCs on Income of Retail Workers

This table presents results of worker-level panel regressions assessing the effect of FCs on income using Equation 2.1. Panel A includes retail workers in counties with FCs. Panel B (C) includes retail workers in counties within 50 (100) miles of FCs but not in counties with FCs. Column (1), (2), (3) includes all workers, hourly workers, and salaried workers, respectively. All regressions include worker and year-quarter fixed effects. Standard errors clustered by county are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% level.

	Log(Total Income)		
	All Workers	Hourly Workers	Salaried Workers
	(1)	(2)	(3)
Panel A: Counties with FCs			
PostFC	-0.024^{***}	-0.025^{***}	-0.010
	(0.005)	(0.005)	(0.008)
Observations	2,175,144	1,881,184	293,960
Adjusted R^2	0.845	0.809	0.849
Panel B: Counties within 50 Miles of FCs			
PostFC	-0.020^{***}	-0.022^{***}	-0.001
	(0.004)	(0.004)	(0.008)
Observations	5,643,934	4,744,111	899,823
Adjusted R^2	0.865	0.828	0.856
Panel C: Counties within 100 Miles of FCs			
PostFC	-0.023^{***}	-0.024^{***}	-0.007^{*}
	(0.002)	(0.002)	(0.004)
Observations	11,141,092	9,549,880	1,591,212
Adjusted R^2	0.858	0.821	0.863
Worker FE	✓	✓	✓
YearQtr FE	✓	✓	✓

trends in online shopping and downward trends in sales at traditional brick-and-mortar retailers. In this case, the negative income effect may be driving the FC establishment and not vice versa. As such, our difference-in-differences assumption is only valid if treatment and control groups follow parallel trends before the shock. To test this, we directly examine the dynamic temporal effects by including leading and lagging indicators of FC establishment by estimating the following:

$$Log(Total\ Income_{i,c,t}) = \alpha + \sum_{j=2}^{4} \beta_j PreFC_{c,t}(-j) + \sum_{j=0}^{4} \gamma_j PostFC_{c,t}(j) + \eta_i + \theta_t + \epsilon_{i,c,t}. \tag{2.2}$$

To increase the power of our estimates, *PreFC* and *PostFC* dummies are defined at half-year intervals. The variable $PreFC_{c,t}(-j)$ ($PostFC_{c,t}(j)$) is a dummy that takes a value 1 if it is *j* half-years before (after) the establishment of FCs. Also, *PreFC(-4)* equals 1 if it is two or more years before the establishment of FCs, and *PostFC (+4)* equals 1 if it is two or more years after the establishment of FCs. The variable *PreFC(-1)* is dropped from the estimation so that all coefficient estimates can be treated as percentage changes relative to the income workers received six months before the establishment of FCs.

In Figure 2.3, Panel A, we show the dynamic effect of FCs on income for counties with FCs by plotting the coefficients from the specification in Equation 2.2. The shaded area around the coefficients represent 95% confidence intervals. Coefficients on *PreFC(-4)*, *PreFC(-3)*, *and PreFC(-2)* are all statistically insignificant from the income of workers in *PreFC(-1)* (the omitted category). That suggests that there is no pre-trend in the data, thus supporting the validity of our parallel trends assumption. Within six months of the establishment of an FC, the income of hourly workers decreases by 2.1% relative to the half-year shortly before the FC's establishment. The negative effect further increases to -4.1% two years after the FC's establishment. We find a similar pattern in Panels B and C of Figure 2.3, where we focus our analysis on counties within 50 miles of the county in which an FC opened and within 100 miles, respectively.

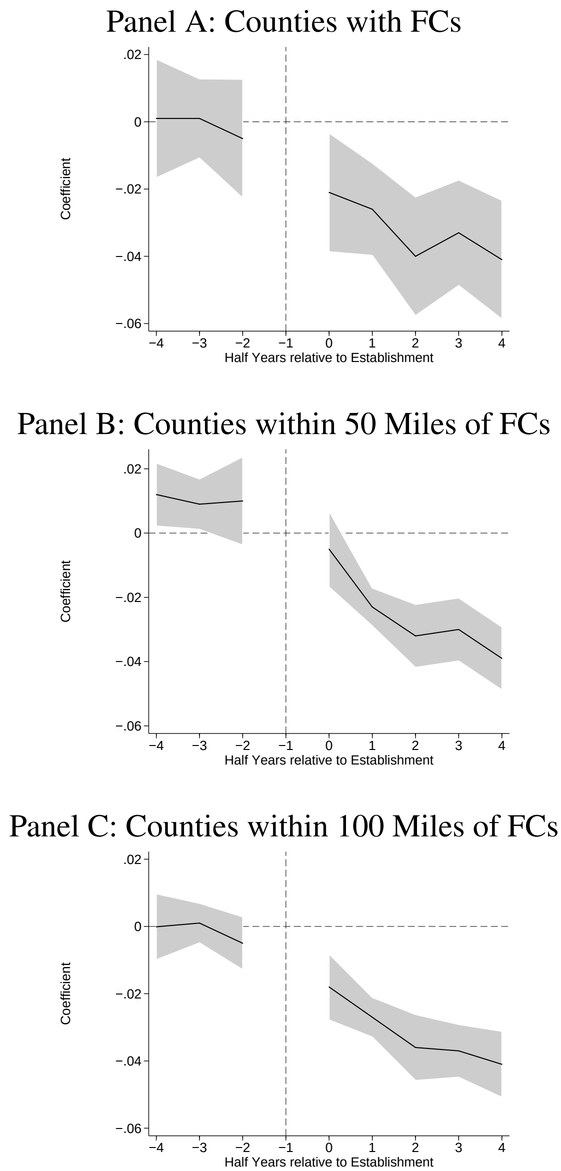


Figure 2.3: Dynamic Effect of FCs on Income

These figures present the dynamic effect of FCs on the income of hourly retail workers. We estimate Equation 2.2 and plot the estimated coefficients from $PreFC$ (j=-4 to j=-2) and $PostFC$ (j=0 to j=4) dummies which are defined at the semiannual frequency. $PreFC(-1)$ is dropped from the estimation so that all coefficient estimates can be treated as percentage changes relative to the income within two quarters before the establishment of FCs. The shaded area around the coefficients represents 5% confidence intervals. Panel A includes workers in counties with FCs. Panel B (C) includes workers in counties within 50 (100) miles of FCs but not in counties with FCs.

Can Unobservable Firm-Specific Variables or Local Economic Conditions Be Driving the Results?

Our results so far suggest a robust and negative relationship between the arrival of an FC and a worker's income. However, absent truly exogenous variation in both the geographic location and temporal timing of FC establishment, we may still be concerned that the arrival of an FC is correlated with unobservables present in the error term of Equation 2.1. These unobservables may include firm-specific characteristics or local economic conditions that jointly affect both the likelihood of an FC arriving in the county and the income of workers in local brick-and-mortar establishments.

For example, if a major lender or supplier to a brick-and-mortar retailer files for bankruptcy and this attracts the e-commerce retailer to establish an FC in the county as a result, then this negative correlation between our error term and the FC establishment amplifies our negative effect sizes. To address this concern, we include firm-year-quarter fixed effects to absorb all time-specific characteristics of our sample firms and identify our parameter of interest by exploiting varition within-firm-time across counties. As such, we can only estimate our *PostFC* variable from firms that operate in more than one county.

We see in Column (1) that when we include these firm-year-quarter fixed effects, the establishment of FCs in the county results in lower total income for retail hourly workers in the brick-and-mortar stores. We find that the magnitude diminishes from the baseline magnitude (2.5%) to 2.1% for counties with FCs. When we extend our analysis to focus on counties within 50 and 100 miles of the county in which the FC was established, we continue to see precisely estimated negative effects.

As discussed in Section 2.2.2, local economic conditions may also play an important role in the establishment of FCs by the e-commerce retailer. States with and without FCs may have different economic and regulatory environments, which could correlate with the establishment of an FC. For example, regions with suppressed economic activity (which would negatively impact retail sales) may be more inclined to offer sizable incentives for e-commerce retailers to establish an FC in their region. To control for time-varying unobservables at the state level, we include state-year-quarter fixed effects and report the results in Table 2.4, Column (2). The estimated effect for counties with FCs drops from -2.5% to -0.8%, but the result remains statistically significant from 0. Further, the results are strong both economically and statistically for workers within 50 or 100 miles of FCs. Our results remain robust when we combine firm-year-quarter and state-year-quarter fixed effects in Column (3).

While the state-year-quarter fixed effects may control for state-level heterogeneity, they may be insufficient to fully absorb any time-varying heterogeneity that arises at the county level. For example, it may be that the e-commerce retailer decides to build an FC in a county at the same time that an unexpected negative economic shock occurs in that county (or possibly even because of such a shock). Thus, it is possible that our baseline estimates may be driven by unexpected local economic shocks rather than by competition from e-commerce. To control for county-specific time-varying shocks, we expand our sample threefold to include data on hourly workers at non-retail firms. In doing so, we can employ a triple difference (difference-in-difference-in-differences) methodology whereby we exploit within county-year-quarter variation across industry type (retail versus non-retail). In doing so, we can carefully control for county-time specific shocks and identify our parameter of interest by comparing retail workers to non-retail workers in FC-treated counties.

If FCs are being established in regions that experience economic hardship, then we should observe no difference in incomes between retail and non-retail workers. In Column (4), we interact the *PostFC* dummy with a *Retail* dummy set to 1 if the focal worker works in a retail industry and 0 otherwise. We find that the income of retail hourly workers in counties with FCs is reduced by 4.4% compared to all other hourly workers within the same county, after controlling for county-level time-varying unobservables. As a result, it seems unlikely that a local negative shock that solely affects a county's retail firms, but not its non-retail firms, is driving our results.

Table 2.4: Firm-Specific Unobservables and Local Economic Conditions
This table presents results of worker-level panel regressions assessing the effect of FCs on the income of *hourly workers* after controlling for firm-specific unobservables and local economic conditions. All columns include worker fixed effects. We replace year-quarter fixed effects in Equation 2.1 with firm-year-quarter, state-year-quarter, and firm-year-quarter and state-year-quarter fixed effects in Column (1), (2), (3), respectively. In Column (4), we include all hourly workers in other industries in addition to retail workers. We interact *PostFC* with *Retail*, where *Retail* identifies retail workers, and we control for county-year-quarter fixed effects. Panel A includes workers in counties with FCs. Panel B (C) includes workers in counties within 50 (100) miles of FCs but not in counties with FCs. Standard errors clustered by county are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% level.

	Log(Total Income)			
	Hourly Workers			
	(1)	(2)	(3)	(4)
Panel A: Counties with FCs				
PostFC	-0.021^{***}	-0.008^{***}	-0.011^{***}	
	(0.004)	(0.002)	(0.002)	
PostFC*Retail				-0.044***
				(0.006)
Observations	1,881,184	1,881,184	1,881,184	5,596,632
Adjusted R^2	0.841	0.811	0.842	0.850
Panel B: Counties within 50 Miles of FCs				
PostFC	-0.017^{***}	-0.020^{***}	-0.018^{***}	
	(0.003)	(0.003)	(0.002)	
PostFC*Retail				-0.029***
				(0.007)
Observations	4,744,111	4,744,111	4,744,111	14,247,756
Adjusted R^2	0.862	0.830	0.863	0.852
Panel C: Counties within 100 Miles of FCs				
PostFC	-0.021^{***}	-0.015^{***}	-0.013^{***}	
	(0.002)	(0.002)	(0.001)	
PostFC*Retail				-0.026***
				(0.004)
Observations	9,549,880	9,549,880	9,549,880	26,914,902
Adjusted R^2	0.853	0.824	0.854	0.856
Worker FE	✓	✓	✓	✓
Firm-YearQtr FE	✓		✓	
State-YearQtr FE		✓	✓	
County-YearQtr FE				✓

Overall, the results presented in this subsection reduce our concerns that our results are being driven by firm-specific unobservable variables or some other omitted local economic conditions that coincide with the staggered establishment of FCs.

Decomposing the Impact of FC Establishment on Wages

Our detailed payroll data on workers allows us to decompose their total income into wage income and bonus income. We can further decompose wage income into hours worked and wage rate. We run Equation 2.1 using different components of total income as our dependent variables. Table 2.5, Column (1) reports results for wage income as the dependent variable. All regressions include worker, firm-year-quarter, and state-year-quarter fixed effects. We continue to use this tighter specification for all our worker-level regression estimates.[13] We find a significant negative impact on wage income across all three panels. The economic magnitude ranges from -0.7% to -1.4%. In Column (2), we find that bonuses decline by 0.3% to 2.6%. To further investigate the source of this wage reduction, we decompose wage income into hours worked and wage rate. In Column (3), we report results for hours worked and find that the estimated coefficients are almost the same as those in Column (1). We do not find economically significant changes in the wage rate (Column (4)).

The results documented in Table 2.5 suggest that the negative impact of FC establishment on local retail workers is mainly driven by the reduction in hours they work. As previously documented, the negative impact is concentrated in hourly workers. At the same time, the wages of many of the hourly retail workers are bound by the applicable minimum wage. So, our results suggest that in the presence of this wage floor, firms cut down on the number of hours demanded from their part-time, hourly workers.

[13]Note that our results are robust and in fact are stronger for our baseline model with only worker and year-quarter fixed effects.

Table 2.5: Decomposition of Income Effect for Hourly Workers

This table presents results of worker-level panel regressions assessing the effect of FCs on the wage income, bonus, hours worked, and wage rate of hourly retail workers using Equation 2.1. Panel A includes workers in counties with FCs. Panel B (C) includes workers in counties within 50 (100) miles of FCs but not in counties with FCs. All regressions include worker, firm-year-quarter, and state-year-quarter fixed effects. Standard errors clustered by county are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% level.

	Log(Wage Income)	Log(Bonus)	Log(Hours Worked)	Log(Wage Rate)
	(1)	(2)	(3)	(4)
Panel A: Counties with FCs				
PostFC	-0.007***	-0.012	-0.007***	0.001***
	(0.002)	(0.008)	(0.002)	(0.000)
Observations	1,875,176	1,764,173	1,866,683	1,866,683
Adjusted R^2	0.842	0.740	0.731	0.970
Panel B: Counties within 50 Miles of FCs				
PostFC	-0.014***	-0.026***	-0.014***	-0.000
	(0.002)	(0.009)	(0.002)	(0.001)
Observations	4,730,532	4,414,584	4,711,647	4,711,647
Adjusted R^2	0.856	0.744	0.761	0.968
Panel C: Counties within 100 Miles of FCs				
PostFC	-0.012***	-0.003	-0.011***	-0.001*
	(0.001)	(0.008)	(0.001)	(0.000)
Observations	9,521,947	8,921,138	9,476,710	9,476,710
Adjusted R^2	0.849	0.746	0.744	0.969
Worker FE	✓	✓	✓	✓
Firm-YearQtr FE	✓	✓	✓	✓
State-YearQtr FE	✓	✓	✓	✓

Are All Retail Workers Affected Equally by the Establishment of FCs?

Our rich worker-level data allow us to analyze *which* workers are more impacted by the negative wage shock due to the establishment of the FCs. This rich demographic information allows us to consider heterogeneity along worker dimensions such as age, gender, and worker status (part-time versus full-time).

We start by exploring worker heterogeneity by age. We split the retail workers in our sample into six age groups. Results from running Equation 2.1 over these six age groups are reported in Table 2.6. We find evidence for a stronger negative impact on the total income of young and old workers. We observe that for young workers under the age of 25 within 50 miles of FCs, total income decreases by 2.7%. The negative effect is lower for age group of 25-34 years (-1.8%). We find that this negative effect increases with age. For age groups 35-44 years, 45-54 years, and 54-64 years, the effect is -1.2%, -1.6%, and -2.0%, respectively. For the oldest group, i.e., workers older than 64 years of age, the negative income effect is as high as -3.5%. These results appear similar for focal counties and those within 100 miles of FCs.

These results suggest that a worker's age has a large moderating impact on the arrival of new technologies. One explanation for this result may be that a worker's age proxies for their productivity and accumulated firm-specific human capital. On the one hand, young workers may be more productive, but firms may not have invested much in enhancing their firm-specific human capital. On the other hand, old workers may have accumulated firm-specific human capital but may be less productive compared to younger workers. Explaining how age plays a prominent role is outside the scope of our paper; however, our results suggest that both younger and older workers shoulder a disproportionate share of the negative impact of FC establishment as opposed to middle-age workers.

Next, we test how the negative income effect varies across worker's working status, i.e., part-time workers versus full-time workers. Similar to age, a worker's working status may reflect his or her underlying level of firm-specific human capital accumulation. Firms may

Table 2.6: Heterogeneous Income Effect: Worker Age

This table presents results of worker-level panel regressions assessing the heterogeneous effect of FCs on the income of hourly retail workers based on the age of workers using Equation 2.1. Panel A includes workers in counties with FCs. Panel B (C) includes workers in counties within 50 (100) miles of FCs but not in counties with FCs. All regressions include worker, firm-year-quarter, and state-year-quarter fixed effects. Standard errors clustered by county are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% level.

	Log(Total Income)					
Age (in years)	<25	25-34	35-44	45-54	55-64	> 64
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Counties with FCs						
PostFC	-0.010***	-0.008***	-0.009***	-0.009***	-0.018***	-0.044***
	(0.003)	(0.002)	(0.003)	(0.002)	(0.004)	(0.006)
Observations	168,518	382,786	420,745	493,286	299,402	95,063
Adjusted R^2	0.782	0.815	0.858	0.851	0.846	0.860
Panel B: Counties within 50 Miles of FCs						
PostFC	-0.027***	-0.018***	-0.012***	-0.016***	-0.020***	-0.035***
	(0.008)	(0.004)	(0.004)	(0.003)	(0.004)	(0.005)
Observations	381,232	900,078	998,542	1,289,569	858,504	262,581
Adjusted R^2	0.786	0.834	0.879	0.885	0.865	0.864
Panel C: Counties within 100 Miles of FCs						
PostFC	-0.021***	-0.012***	-0.009***	-0.012***	-0.018***	-0.026***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)
Observations	834,541	1,883,249	2,035,922	2,532,303	1,644,437	508,696
Adjusted R^2	0.783	0.828	0.871	0.874	0.853	0.853
Worker FE	✓	✓	✓	✓	✓	✓
Firm-YearQtr FE	✓	✓	✓	✓	✓	✓
State-YearQtr FE	✓	✓	✓	✓	✓	✓

invest more in the human capital development of full-time workers than part-time workers. We define a worker as a part-time worker if the hours worked is less than 32 hours per week, otherwise the worker is considered a full-time worker. We define the worker's employment status in a time-invariant fashion by categorizing each worker by their work status at the beginning of our sample period, i.e., in 2010Q1. Table 2.7, Panel A reports the differential effect on part-time and full-time workers. In line with the hour reduction results previously reported, we find that the negative effect is stronger for part-time workers, i.e., the impact on part-time workers is about -1% more.

It is possible that the negative impact of FCs varies by gender, given the significant fraction of female retail employees. We test whether there is any differential effect of FCs on male versus female workers. Table 2.7, Panel B suggests that there is no difference in the effect based on worker gender.

In summary, the heterogeneity in the negative impact of establishment of FCs may be relevant for designing remedial responses to the negative impact of establishment of FCs on local retail employees. We find that young and old workers (as opposed to middle-aged workers) and part-time workers (as opposed to full-time workers) experience disproportionately more negative effects from the establishment of FCs in their focal and proximate counties.

Additional Robustness Tests

We further conduct additional tests to ensure the robustness of the results reported so far. In our main analysis (i.e., results from estimating Equation 2.1), we focus on post-2010 FCs, since our sample starts with 2010. We include all FCs in Table 2.8, Panel A as a robustness test. The income effect is about -1% and is still significant in Column (2) and Column (3). The lower magnitudes can be attributed to the non-availability of pre-treatment worker data for FCs established before 2010.

In our baseline tests, we assigned FC treatment to workers based on the the ZIP code

Table 2.7: Heterogeneous Income Effect: Hours Worked and Gender

This table presents results of worker-level panel regressions assessing the heterogeneous effect of FCs on the income of hourly workers based on hours worked and gender of workers using Equation 2.1. Panel A compares part-time workers (less than 32 hours per week) and full-time workers (equal to or more than 32 hours per week). Panel B compares female workers and male workers. All regressions include worker, group-specific firm-year-quarter, and group-specific state-year-quarter fixed effects. Standard errors clustered by county are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% level.

	Log(Total Income)		
	Counties with FCs	Counties within 50 Miles of FCs	Counties within 100 Miles of FCs
	(1)	(2)	(3)
Panel A: Part-time vs. Full-time			
PostFC*Part-time (1)	-0.016***	-0.028***	-0.021***
	(0.003)	(0.004)	(0.002)
PostFC*Full-time (2)	-0.010***	-0.014***	-0.010***
	(0.002)	(0.002)	(0.001)
Difference ((1)-(2))	-0.005	-0.014***	-0.011***
	(0.004)	(0.003)	(0.002)
Observations	1,872,894	4,725,658	9,505,116
Adjusted R^2	0.843	0.864	0.854
Worker FE	✓	✓	✓
PartTime-Firm-YearQtr FE	✓	✓	✓
PartTime-State-YearQtr FE	✓	✓	✓
Panel B: Female vs. Male			
PostFC*Female (3)	-0.011***	-0.016***	-0.014***
	(0.004)	(0.003)	(0.002)
PostFC*Male (4)	-0.019***	-0.020***	-0.015***
	(0.003)	(0.004)	(0.003)
Difference ((3)-(4))	0.008*	0.004	0.002
	(0.004)	(0.005)	(0.003)
Observations	644,483	1,910,447	3,367,008
Adjusted R^2	0.850	0.853	0.849
Worker FE	✓	✓	✓
Female-Firm-YearQtr FE	✓	✓	✓
Female-State-YearQtr FE	✓	✓	✓

Table 2.8: Robustness Check

This table presents robustness checks for the effect of FCs on the income of hourly retail workers using Equation 2.1. Panel A uses all FCs including those established before 2010. Panel B excludes migrants whose first and last ZIP codes in the data are different. Panel C uses the modeled annualized income as an alternative measure of income. All regressions include worker, firm-year-quarter, and state-year-quarter fixed effects. Standard errors clustered by county are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% level.

	Log(Total Income)		
	Counties with FCs	Counties within 50 Miles of FCs	Counties within 100 Miles of FCs
	(1)	(2)	(3)
Panel A: All FCs			
PostFC	-0.006*	-0.011***	-0.010***
	(0.003)	(0.002)	(0.001)
Observations	2,789,307	6,703,938	13,858,442
Adjusted R^2	0.852	0.870	0.863
Panel B: Excluding Migrants			
PostFC	-0.013***	-0.019***	-0.014***
	(0.002)	(0.002)	(0.001)
Observations	1,632,007	4,041,405	8,208,699
Adjusted R^2	0.848	0.869	0.860
Panel C: Alternative Measure of Income			
PostFC	-0.009***	-0.013***	-0.010***
	(0.002)	(0.002)	(0.001)
Observations	1,885,938	4,753,073	9,570,559
Adjusted R^2	0.911	0.918	0.914
Worker FE	✓	✓	✓
Firm-YearQtr FE	✓	✓	✓
State-YearQtr FE	✓	✓	✓

of their residence in 2010Q1. It is possible that some workers move to avoid the negative income shocks caused by FCs. Therefore, as a further robustness test, we remove migrants whose last observed ZIP code in the data is different from the first observed ZIP code. Results documented in Panel B of Table 2.8 indicate that our main results remain unaffected.

In our analysis so far, we have used quarterly income computed from raw payroll data under the assumption that it is a timely reflection of the impact of FCs. But, it is possible that quarterly income is subject to seasonal variation. Since income is the key outcome measure in our analysis, we show robustness to our dependent variable by using an alternative income measure that is based on the projected annual income of a worker every month as computed by the credit bureau. We rerun our analysis with this projected annual income instead of the quarterly income we have used so far. Results documented in Panel C of Table 2.8 show a significant negative effect, suggesting that our analysis is robust to this measure of income.

Credit Outcomes

Our results so far indicate that the establishment of FCs by a major online retailer has a negative impact on the wages of workers at traditional brick-and-mortar retail stores in the focal county and geographically proximate counties. The effects are predominantly borne by young and old workers as opposed to middle-aged workers, part-time workers as opposed to full-time workers, and hourly workers as opposed to salaried employees.

However, to the extent that labor markets are frictionless (i.e., workers can easily switch jobs, and skills are completely transferable) the short-term displacement of some traditional retail store workers that we document may not matter for the workers or the local economy. However, in the presence of labor market frictions, the short-term impact on the workers and local economy can be negative. Moreover, to the extent that the scope of work differs between traditional retail stores and warehouses, at least some workers can be worse off.

Our data prevent us from identifying any other source of income for the affected workers, specifically part-time and hourly workers, except income from their primary employer in the credit bureau payroll database. So, we are unable to directly verify whether the affected workers offset the reduced hours with brick-and-mortar retail stores by picking up additional working hours with another employer (who may not be part of the payroll database that we use).

We test for this possibility indirectly by considering the credit outcomes of the workers. If workers can easily substitute their sources of income, then it should have no effect on their credit outcomes. Otherwise, the declines in income may lead to worse credit outcomes, especially for the workers who are already living at the margin (i.e., workers with high bank card utilization). We use credit score as a measure of the credit outcomes for the affected workers. We assign a worker to the high utilization group if her bank card utilization is higher than the median utilization ratio, and we assign workers to the low utilization group if their bank card utilization is lower than the median.

We report our results in Table 2.9. We find that the credit score for workers with high utilization of bank credit cards declines significantly. In counties with FCs, the decreases in credit scores of workers in the high utilization group are 3 points more than that of workers in the low utilization group. It seems that the decline in credit scores is driven by a higher bank credit card delinquency among the affected workers.

Overall, the evidence suggests that technological change, as manifested by an e-commerce retailer establishing an FC, leads to a decline in wages for workers in traditional brick-and-mortar stores. Among the affected workers, those who have a prior higher credit card utilization and those who are otherwise more financially vulnerable experience higher credit card delinquencies and a subsequent decline in their credit score. These results suggest that some of the affected retail workers experience some frictions in the labor market that preclude them from mitigating the extent to which the establishment of e-commerce FCs in their county depresses their wage income and subsequently their credit scores.

Table 2.9: Heterogeneous Effect on Credit Scores: Bank Card Utilization
This table presents results of worker-level panel regressions assessing the heterogeneous effect of FCs on credit scores and bank card 90+ day delinquency of hourly workers based on bank card utilization using Equation 2.1. All regressions include worker, group-specific firm-year-quarter, and group-specific state-year-quarter fixed effects. Standard errors clustered by county are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% level.

	Counties with FCs (1)	Counties within 50 Miles of FCs (2)	Counties within 100 Miles of FCs (3)
	Panel A: Credit Scores		
PostFC*Low (1)	0.593	0.004	0.110
	(0.786)	(0.433)	(0.325)
PostFC*High (2)	-2.500***	-0.835*	-0.495*
	(0.220)	(0.506)	(0.282)
Difference ((2)-(1))	-3.089***	-0.849	-0.605*
	(0.735)	(0.609)	(0.367)
Observations	1,210,611	3,191,674	6,441,603
Adjusted R^2	0.812	0.826	0.822
Worker FE	✓	✓	✓
Low-Firm-YearQtr FE	✓	✓	✓
Low-State-YearQtr FE	✓	✓	✓
	Panel B: Bank Card 90+ Day Delinquency		
PostFC*Low (3)	-0.001*	-0.000	0.001
	(0.001)	(0.001)	(0.000)
PostFC*High (4)	0.006***	0.002**	0.000
	(0.001)	(0.001)	(0.001)
Difference ((4)-(3))	0.007***	0.003***	-0.000
	(0.001)	(0.001)	(0.001)
Observations	1,081,133	2,879,747	5,778,895
Adjusted R^2	0.108	0.111	0.112
Worker FE	✓	✓	✓
Low-Firm-YearQtr FE	✓	✓	✓
Low-State-YearQtr FE	✓	✓	✓

2.4.2 How Do FCs Affect Local Brick-and-Mortar Retail Stores? Evidence from NETS Data

So far, we have used detailed worker level data and the staggered establishment of FCs of the e-commerce retailer to understand the impact on the wages of the workers at traditional brick-and-mortar retail stores in the focal county and geographically proximate counties. We next use NETS establishment-level data in order to understand the impact of the establishment of FCs on traditional brick-and-mortar retail stores themselves. This analysis examines the aggregate implications of the negative income effects that workers suffer when they work fewer hours.

Effect on Retail Store Sales

In Table 2.10, we use NETS data to understand the effect of FCs on the sales of local brick-and-mortar stores. Column (1) reports difference-in-differences estimates for all stores in the counties with FCs. In all the specifications, we include establishment fixed effects, 6-digit NAICS-year fixed effects, and state-year fixed effects. We find that the annual sales of local brick-and-mortar retail stores decrease by 2.8% (equivalent to $63,639 per store) after the establishment of an FC of an e-commerce retailer.

In the next three columns, we partition the incumbent establishment/store sample into terciles based on sales one year before the establishment of FCs in the county. We find that for the bottom tercile (*Small*), the annual sales decrease by 2.9%, which is equivalent to $7,565 per store. For the medium group (*Medium*), we find that sales decrease by 2%, which is equivalent to $12,348 per store. For the top group (*Large*), sales decrease by 3.4%, which is equivalent to $200,389 per store. These results suggest that the establishment of FCs negatively affects the sales of local brick-and-mortar stores, especially large retail stores. The effect is diminished in counties that are 50 or 100 miles from FCs.

These results also suggest that after the staggered establishment of the e-commerce firm's FCs, sales decline significantly in the focal county of the FC. This decline in sales

Table 2.10: Effect of FCs on Sales of Retail Stores

This table presents results of establishment-level panel regressions assessing the heterogeneous effect of FCs on the sales of retail establishments/stores based on the size of stores. Panel A includes establishments in counties with FCs. Panel B (C) includes establishments in counties within 50 (100) miles of FCs but not in counties with FCs. Column (1) reports results for all stores, while Column (2)-Column (4) report results for terciles based on sales one year before the establishment of FCs in the county or neighboring county. All regressions include establishment, industry-year and state-year fixed effects. Standard errors clustered by county are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% level.

	Log(1+Sales)			
	All	Small	Medium	Large
	(1)	(2)	(3)	(4)
Panel A: Counties with FCs				
PostFC	-0.028***	-0.029**	-0.020***	-0.034***
	(0.009)	(0.012)	(0.005)	(0.012)
Observations	184,829	63,636	58,799	62,394
Adjusted R^2	0.959	0.808	0.728	0.918
Panel B: Counties within 50 Miles of FCs				
PostFC	-0.009	0.010	-0.024***	-0.017
	(0.008)	(0.008)	(0.008)	(0.018)
Observations	509,182	172,954	164,848	171,362
Adjusted R^2	0.956	0.801	0.744	0.914
Panel C: Counties within 100 Miles of FCs				
PostFC	-0.011*	-0.018***	-0.009*	-0.005
	(0.005)	(0.003)	(0.005)	(0.011)
Observations	1,073,929	364,588	348,315	361,015
Adjusted R^2	0.958	0.800	0.765	0.921
Establishment FE	✓	✓	✓	✓
Ind-Year FE	✓	✓	✓	✓
State-Year FE	✓	✓	✓	✓

may result in financial stress on the store, or it may motivate the parent company to focus on improving the operational performance of the store. One consequence of the decline in sales may be a reduction in the number of hours of work assigned to part-time and hourly workers.

Effect on Retail Store Employment

So far, we find that the establishment of FCs negatively affects the income of retail workers and the sales of local brick-and-mortar stores. Thus, it is instructive to understand how stores respond to lower sales after the increase in competition due to the establishment of the e-commerce retailer's FCs. Do they also adjust overall employment levels in addition to reducing the number of hours of part-time and hourly workers?

Table 2.11 reports results for the effect of FCs on local establishment-level employment. The results appear similar to sales. Column (1) reports difference-in-differences estimates for all stores in counties with FCs. We find that for all stores, employment decreases by 2.4%, which is equivalent to a reduction in employees of 41 workers per 100 stores for a store with an average of 22 employees. For *small* stores, employment decreases by 2.7%, which is equivalent to reducing 11 workers per 100 stores for a store with an average of 4 employees. For large stores, employment decreases by 1 worker per store for a store with an average of 40 employees. Similar to sales results, the effect is diminished in counties that are 50 or 100 miles of an FC.

Based on the results presented in the previous two subsections, it appears that after the establishment of the e-commerce retailer's FCs, traditional brick-and-mortar retail stores in the focal county adjust to the decline in the sales both by reducing the number of hours of work assigned to part-time and hourly workers and also by reducing employment levels.

Table 2.11: Effect of FCs on Employment of Retail Stores

This table presents results of establishment-level panel regressions assessing the heterogeneous effect of FCs on the employment of retail establishments/stores based on the size of stores. Panel A includes establishments in counties with FCs. Panel B (C) includes establishments in counties within 50 (100) miles of FCs but not in counties with FCs. Column (1) reports results for all stores, while Column (2)-Column (4) report results for terciles based on sales one year before the establishment of FCs in the county or neighboring county. All regressions include establishment, industry-year and state-year fixed effects. Standard errors clustered by county are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% level.

	Log(1+Employment)			
	All	Small	Medium	Large
	(1)	(2)	(3)	(4)
Panel A: Counties with FCs				
PostFC	-0.024***	-0.027***	-0.019***	-0.025***
	(0.007)	(0.006)	(0.005)	(0.009)
Observations	184,829	63,636	58,799	62,394
Adjusted R^2	0.973	0.871	0.907	0.955
Panel B: Counties within 50 Miles of FCs				
PostFC	-0.007	-0.003	-0.015***	-0.005
	(0.004)	(0.003)	(0.005)	(0.007)
Observations	509,182	172,954	164,848	171,362
Adjusted R^2	0.973	0.875	0.912	0.956
Panel C: Counties within 100 Miles of FCs				
PostFC	-0.008**	-0.011***	-0.007**	-0.005
	(0.003)	(0.003)	(0.003)	(0.004)
Observations	1,073,929	364,588	348,315	361,015
Adjusted R^2	0.978	0.883	0.924	0.969
Establishment FE	✓	✓	✓	✓
Ind-Year FE	✓	✓	✓	✓
State-Year FE	✓	✓	✓	✓

Closures of Retail Stores

Next, we analyze whether the increase in competition and the consequent decline in store sales after the establishment of the e-commerce retailer's FCs can lead, in extreme cases, to an increase in retail store closures.

In Tables 2.12 and 2.13, we attempt to understand whether the establishment of FCs leads to store closures and how this effect varies with store size and age. Here we define *exit*, our dependent variable, as a dummy equal to 1 if the establishment ceases to exist one year before the end of the sample period, and 0 otherwise. Table 2.12, Panel A, Column (1) reports results for all stores. We find that the exit rate increases by almost 3%. The average exit rate in our sample is almost 13.6%. The effect is negatively correlated with the ex ante size of the store, i.e., small stores are more likely to exit than large stores. This effect is consistent for counties 50 miles or 100 miles from a FC.

We further test the role of a store's age on exits, and we report the results in Table 2.13. Here, we partition the *All* stores further into terciles, i.e., *young*, *medium*, and *old* based on ex ante age. The average age in the bottom tercile is about eight years. We find that young stores are more likely to close.

So, based on the analysis in Tables 2.12 and 2.13, it appears that there is an increase in the exit rate of local brick-and-mortar retail stores after the establishment of the FCs of the e-commerce retailer. This exit rate impact is more pronounced for young and small retail stores, as they are likely to be more financially stressed and may not be able to survive the decline in sales after the FCs are established in the focal county.

Entry of Retail Stores

In all of our previous analysis, we focus on the effect of the establishment of FCs on incumbent brick-and-mortar retail stores. But it is possible that entry into the local retail sector is discouraged by the establishment of the FC, the consequent increase in competition, and the decline in sales and closure of some of the incumbent brick-and-mortar retail stores. We

Table 2.12: Retail Store Closures: Size Effect

This table presents results of establishment-level panel regressions assessing the heterogeneous effect of FCs on the exit rates of retail establishments/stores based on the size of stores. Here, we define *exit* dummy, our dependent variable, as 1 if the establishment ceases to exist one year before the end of the sample period, and 0 otherwise. Panel A includes establishments in counties with FCs. Panel B (C) includes establishments in counties within 50 (100) miles of FCs but not in counties with FCs. Column (1) reports results for all stores, while Column (2)-Column (4) report results for terciles based on sales one year before the establishment of FCs in the county or neighboring county. All regressions include establishment, industry-year and state-year fixed effects. Standard errors clustered by county are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% level.

	Exit			
	All	Small	Medium	Large
	(1)	(2)	(3)	(4)
Panel A: Counties with FCs				
PostFC	0.029***	0.038**	0.027***	0.023***
	(0.010)	(0.014)	(0.009)	(0.007)
Observations	161,520	53,753	53,470	54,297
Adjusted R^2	0.214	0.222	0.213	0.207
Panel B: Counties within 50 Miles of FCs				
PostFC	0.023***	0.035***	0.022***	0.012**
	(0.007)	(0.009)	(0.007)	(0.005)
Observations	446,770	148,440	148,708	149,604
Adjusted R^2	0.202	0.207	0.201	0.200
Panel C: Counties within 100 Miles of FCs				
PostFC	0.014***	0.020***	0.012**	0.011***
	(0.004)	(0.005)	(0.005)	(0.003)
Observations	973,208	322,744	324,191	326,261
Adjusted R^2	0.204	0.207	0.199	0.204
Establishment FE	✓	✓	✓	✓
Ind-Year FE	✓	✓	✓	✓
State-Year FE	✓	✓	✓	✓

Table 2.13: Retail Store Closures: Age Effect

This table presents results of establishment-level panel regressions assessing the heterogeneous effect of FCs on the exit rates of retail establishments/stores based on the age of stores. Here, we define *exit* dummy, our dependent variable, as 1 if the establishment ceases to exist one year before the end of the sample period, and 0 otherwise. Panel A includes establishments in counties with FCs. Panel B (C) includes establishments in counties within 50 (100) miles of FCs but not in counties with FCs. Column (1) reports results for all stores for which we observe the store's age, while Column (2)-Column (4) report results for terciles based on store's age one year before the establishment of FCs in the county or neighboring county. All regressions include establishment, industry-year, and state-year fixed effects. Standard errors clustered by county are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% level.

	Exit			
	All	Young	Medium	Old
	(1)	(2)	(3)	(4)
Panel A: Counties with FCs				
PostFC	0.032**	0.041**	0.043***	0.007
	(0.012)	(0.016)	(0.013)	(0.010)
Observations	75,162	26,080	25,093	23,984
Adjusted R^2	0.204	0.217	0.199	0.185
Panel B: Counties within 50 Miles of FCs				
PostFC	0.032***	0.040***	0.030***	0.025***
	(0.008)	(0.011)	(0.009)	(0.007)
Observations	226,182	79,011	72,694	74,465
Adjusted R^2	0.198	0.212	0.190	0.182
Panel C: Counties within 100 Miles of FCs				
PostFC	0.016***	0.022***	0.015***	0.009***
	(0.005)	(0.008)	(0.004)	(0.003)
Observations	508,929	179,539	165,739	163,630
Adjusted R^2	0.202	0.211	0.196	0.191
Establishment FE	✓	✓	✓	✓
Ind-Year FE	✓	✓	✓	✓
State-Year FE	✓	✓	✓	✓

analyze the impact of the establishment of the FC on entry into the local retail market in Table 2.14. Column (1) of Panel A reports county-level results on the number of entrants. We find that after the establishment of an FC in the affected county, the entry rate for small stores is significantly reduced by 11.8%. This low entry rate effect is not just limited to counties with FCs; it also persists in counties 50 or 100 miles from an FC.

Effect on Sales of Full-Service Restaurants

We have attempted to rule out the key competing alternate explanation for the pattern we observe in our data: the possibility that our results are driven by some omitted local economic conditions and not by the establishment of the e-commerce retailer's FC in the county. As a further test to reduce the concerns about this interpretation, we test whether FCs have any effect on the sales of local full-service restaurants. If some omitted local economic shock is positively correlated with the establishment of an FC in the affected counties, we would expect that the sales of full-service restaurants (another important non-tradable sector) also respond to this negative economic shock and experience a decrease in the focal counties.

Table 2.15 reports the results for the effect of FCs on the sales of full-service restaurants. Panel A, column (1) reports difference-in-differences estimates for all full-service restaurants. We do not find any negative effect on sales of full-service restaurants with the establishment of FCs in the counties. Similar to the previous subsection, we partition the data into terciles based on ex ante sales. We find that sales of *small* restaurants increase in the focal counties, but we do not find any effect on sales for medium- or large-sized restaurants nor any effect within 50 or 100 miles of FCs. These results indicate that it is unlikely that an omitted local economic shock is responsible for the negative impact on the sales of traditional brick-and-mortar retailers.

In summary, using detailed establishment level data from NETS, we find that after the staggered establishment of the FCs of the e-commerce retailer, the geographically proxi-

Table 2.14: Opening of Retail Stores

This table presents results of county-level panel regressions assessing the heterogeneous effect of FCs on the entry rates of retail establishments/stores based on the size of stores. Here, we define *entry rate* as the logged value of the number of entrants in a given county. Panel A includes entrants in counties with FCs. Panel B (C) includes entrants in counties within 50 (100) miles of FCs but not in counties with FCs. Column (1) reports results for all stores, while Column (2)-Column (4) report results for terciles based on the first-year sales after the entry. All regressions include establishment, industry-year, and state-year fixed effects. Standard errors clustered by county are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% level.

	Log(1+Entrants)			
	All	Small	Medium	Large
	(1)	(2)	(3)	(4)
Panel A: Counties with FCs				
PostFC	-0.015	-0.118**	0.095	0.019
	(0.040)	(0.052)	(0.083)	(0.078)
Observations	180	180	180	180
Adjusted R^2	0.989	0.974	0.981	0.968
Panel B: Counties within 50 Miles of FCs				
PostFC	-0.088**	-0.054	-0.043	-0.133**
	(0.039)	(0.053)	(0.073)	(0.055)
Observations	2,205	2,205	2,205	2,205
Adjusted R^2	0.966	0.937	0.933	0.934
Panel C: Counties within 100 Miles of FCs				
PostFC	-0.138***	-0.132***	-0.070*	-0.126***
	(0.030)	(0.039)	(0.038)	(0.039)
Observations	5,690	5,690	5,690	5,690
Adjusted R^2	0.954	0.919	0.916	0.913
Establishment FE	✓	✓	✓	✓
Ind-Year FE	✓	✓	✓	✓
State-Year FE	✓	✓	✓	✓

Table 2.15: Placebo Tests: Sales of Full-Service Restaurants

This table presents results of establishment-level panel regressions assessing the heterogeneous effect of FCs on the sales of full-service restaurants (NAICS 722511) based on the size of restaurants. Panel A includes establishments in counties with FCs. Panel B (C) includes establishments in counties within 50 (100) miles of FCs but not in counties with FCs. Column (1) reports results for all restaurants, while Column (2)-Column (4) report results for terciles based on sales one year before the establishment of FCs in the county or neighboring county. All regressions include establishment and state-year fixed effects. Standard errors clustered by county are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% level.

	Log(1+Sales)			
	All	Small	Medium	Large
	(1)	(2)	(3)	(4)
Panel A: Counties with FCs				
PostFC	0.002	0.012***	-0.002	-0.002
	(0.002)	(0.004)	(0.002)	(0.004)
Observations	110,043	36,097	40,546	33,400
Adjusted R^2	0.979	0.895	0.856	0.944
Panel B: Counties within 50 Miles of FCs				
PostFC	0.003	0.010	-0.003	0.000
	(0.003)	(0.008)	(0.004)	(0.008)
Observations	311,132	100,077	115,726	95,329
Adjusted R^2	0.980	0.892	0.806	0.956
Panel C: Counties within 100 Miles of FCs				
PostFC	-0.001	-0.003	-0.001	-0.001
	(0.001)	(0.003)	(0.001)	(0.005)
Observations	616,376	198,337	227,175	190,859
Adjusted R^2	0.978	0.879	0.822	0.949
Establishment FE	✓	✓	✓	✓
State-Year FE	✓	✓	✓	✓

mate traditional brick-and-mortar retail stores experience a decline in sales, a decline in employment, a decline in entry in the local retail sector, and an increase in closures among the incumbent firms. The impact on store closures is more pronounced for young stores and small stores, whereas the decline in sales is more pronounced for larger stores. In addition, as a placebo, we find that the staggered establishment of the FCs of the e-commerce retailer does not correlate with the sales of full-service restaurants. Overall, our results suggest that the establishment of the e-commerce retailer's FCs has a negative impact on the financial health of the local traditional brick-and-mortar retail stores.

2.4.3 How Do FCs Affect Employment and Wage Growth? Evidence from BLS County-Industry Data

Finally, to understand the aggregate effect at the county level, we use county-level QCEW data on employment and total wages for each NAICS 2-digit sector. We estimate county-industry specific quarterly employment growth and wage growth and report the results of the analysis in Table 2.16. In Panel A, Column (1), we compare the effect of FCs on the employment growth in retail (NAICS 44-45), transportation and warehouse (NAICS 48-49) and restaurants (NAICS 72). Here, we keep all the other industries and all US counties to absorb the time-varying county-specific and industry-specific unobservables. The interaction term *Post FC* $\times$ *Retail* identifies the effect of establishment of FCs on the employment growth in the retail sector compared to all other sectors within the same county.

We find that the establishment of an FC has a negative effect on the employment growth in the retail sector. This is consistent with evidence using payroll data and NETS establishment-level data. A 2.9% decline in employment growth implies a loss of 938 jobs per county per quarter. While, on the other hand, the establishment of FCs do create 256 jobs in transportation and warehousing sector. Further, we find a mildly significant positive

Table 2.16: Effect of FCs on Wages and Employment: County-Industry Evidence
This table presents results of county-industry level panel regressions assessing the aggregate effect of FCs on employment growth and total wage growth. Retail (NAICS 44-45), Warehouse(NAICS 48-49) and Restaurant (NAICS 72) dummies identifies respective industries. All regressions include county-year-quarter and industry-year-quarter fixed effects. Standard errors clustered by county are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% level.

	Counties with FCs			Counties within 50 Miles			Counties within 100 Miles		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				Panal A: Employment Growth					
Post FC × Retail	-0.029***			-0.005*			-0.009***		
	(0.003)			(0.003)			(0.002)		
Post FC × (Retail		-0.025***			-0.005*			-0.010***	
+Warehouse)		(0.003)			(0.003)			(0.002)	
Post FC × (Retail			-0.033***			-0.008**			-0.012***
+Warehouse+Restaurant)			(0.003)			(0.003)			(0.002)
Post FC × Warehouse	0.021***			0.003			-0.006		
	(0.005)			(0.005)			(0.003)		
Post FC × Restaurant	0.005*	0.005*		-0.001	-0.001		-0.002	-0.002	
	(0.003)	(0.003)		(0.004)	(0.004)		(0.003)	(0.003)	
Adjusted R^2	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Obs.	1,178,848	1,122,988	1,056,669	1,178,848	1,122,988	1,056,669	1,178,848	1,122,988	1,056,669
				Panal B: Total Wage Growth					
Post FC × Retail	-0.028***			-0.004			-0.007***		
	(0.003)			(0.003)			(0.002)		
Post FC × (Retail		-0.025***			-0.004			-0.008***	
+Warehouse)		(0.003)			(0.003)			(0.002)	
Post FC × (Retail			-0.030***			-0.006*			-0.010***
+Warehouse+Restaurant)			(0.003)			(0.003)			(0.002)
Post FC × Warehouse	0.017***			0.006			-0.003		
	(0.005)			(0.005)			(0.003)		
Post FC × Restaurant	0.003	0.003		-0.001	-0.001		-0.002	-0.002	
	(0.003)	(0.003)		(0.004)	(0.004)		(0.003)	(0.003)	
Adjusted R^2	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Obs.	1,179,588	1,123,728	1,057,409	1,179,588	1,123,728	1,057,409	1,179,588	1,123,728	1,057,409
County-YearQtr FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Ind-YearQtr FE	✓	✓	✓	✓	✓	✓	✓	✓	✓

spillover effect in restaurants, which leads to 143 more jobs per county per quarter.[14]

In Column (2) and Column (3), we combine retail and warehouse, and retail, warehouse and restaurants, respectively and estimate the growth rates again for combined sectors. The aggregate employment growth effect is negative compared to all the other sectors within the same county with FC. In Column(4)-Column(9), we include counties within 50 and 100 miles of FCs but not in counties with FCs. Note that the positive effect on transportation and warehousing sector and restaurants disappears while the negative effect on the retail sector in the far away counties diminishes but remains negative and statistically significant. Panel B reports the results for total wage growth and we find consistent results.

Overall, the results using county-level QCEW data are largely supportive of our findings using administrative employment data and NETS data.

2.5 Conclusion

The recent disruption in the retail sector is attributed to the rise of e-commerce. As of 2018Q3, e-commerce sales accounted for 9.1% of total retail sales in the U.S., compared to 3.8% in 2010. We use the staggered rollout of a major e-commerce retailer's FCs as a proxy for local e-commerce presence. Using a payroll dataset for 2.6 million retail workers, we find that the labor income of retail workers in counties with FCs, on average, decreases by 2.4% after the establishment of FCs. Wages of hourly workers decrease significantly by 2.5%, equivalent to $825. Most of the effect comes from a reduction in the number of hours worked.

Further, using sales and employment data for 3.2 million stores, we find that retail stores in counties with FCs experience a reduction in sales and employees. We find that for stores in the top tercile based on sales one year before the FC, after the establishment of FCs in their county, their sales decrease by almost 3.4%, which is equivalent to $200,389 per store.

[14]In our sample, i.e., during 2010-2016, an average county with fulfillment center employs 32,373 workers in the retail sector, about 12,155 workers in transportation and warehousing, and nearly 28,589 workers in restaurants and accommodation.

For these stores, after the establishment of FCs in their county, their employment decreases by almost 2.5%, which is equivalent to one worker per store for a store with an average of 40 employees. Also, there is a decrease in entry and an increase in exits for stores in the retail sector, with small and young retail stores exiting at a higher rate. We find that the opening of FCs has no impact on the sales of a full-service restaurant, suggesting that negative local economic shocks may not be driving our results.

Overall, our results highlight how the dramatic increase in e-commerce retail sales can have adverse consequences for workers at traditional brick-and-mortar stores. At the same time, our results should be interpreted carefully in light of the many benefits of e-commerce. In this paper, we do not consider the impact of e-commerce on consumers, the increase in employment by the e-commerce firm, or the e-commerce firm's ecosystem and the ancilliary benefits to the county. Further, we do not consider the long-term dynamics of the labor market in the counties affected by the FCs nor do we consider the long-term effects on the traditional brick-and-mortar retail workers who are affected by the establishment of e-commerce FCs in the focal county. Given the limited scope of this paper, we do not aim to quantify the aggregate effect of e-commerce on the retail sector. Our results can only show that the growth of e-commerce has some adverse consequences for some traditional brick-and-mortar retail workers, and they can provide one piece of evidence to help fully quantify the impact of e-commerce.

CHAPTER 3
THE SPEED OF INFORMATION AND THE SELL-SIDE RESEARCH INDUSTRY

3.1 Introduction

Sell-side analysts play an important role as information providers in financial markets. For the analyst and brokerage house, time is of utmost importance – the business model of sell-side research is predicated on having the time to pitch clients on the information in their recommendations.[1]

In recent years, the sell-side business model has been challenged by third-party financial technology firms that quickly leak research reports. Notably, *Theflyonthewall.com (FLY)* quickly leaks a significant percentage of analysts' recommendations before the market open. A 2007 *Wall Street Journal* article citing Candace Browning, head of Global Securities Research at Merrill Lynch, illustrates the issue. Browning describes a recommendation upgrade by her firm that resulted in large stock price gains, but the firms' clients were not able to act quickly enough to profit from the recommendation. "What happened was that within 60 seconds of releasing our opinion change, the same information was being copied by a New Jersey-based digital financial news source."

In this paper, we examine the importance of speed for the sell-side research industry by examining the redistribution of recommendations by *FLY*. We focus on *FLY* for three primary reasons. First, *FLY* was one of the first movers in redistributing analyst research and they have a broad distribution network. Second, *FLY* was sued by three major brokers and several of our empirical tests rely on key dates in the court case. Finally, *FLY* received

[1]For instance, brokers conduct daily morning meetings going over its best ideas and recommendation changes from the night before. The sales staff will subsequently contact clients through messaging, calls, or email to attract their attention to a recommendation hoping that the client will place a trade. See Barclay's Capital Inc. v. *Theflyonthewall.com*.

substantial attention in the popular press, in legal commentary, and is mentioned anecdotally in several academic studies (i.e., [82, 80, 3, 84]). Despite wide-spread attention, the economic consequences of such third-party news aggregators are largely unknown.

Ex ante, it is not clear what impact speedy redistribution of analyst recommendations has on the brokerage industry. Brokers invest heavily in their research departments to provide their own paying clients access to research. If their research becomes a free good, then this creates disincentives to produce it. However, if broker house clients have the opportunity to act on the research first, and redistribution draws new attention and price movement, then these third-party news aggregators may actually be beneficial to brokers. In the context of [45], *FLY* could essentially allow a broker's clients to trade against 'dumb money.'[2]

We have three main findings. First, *FLY* systematically leaks a meaningful number of analyst recommendations. Over the 2009 to 2013 period, we find that 58.4% of the 83,950 recommendation revisions in Thomson Reuters I/B/E/S database are leaked on *FLY*. The median difference between *FLY* and I/B/E/S announcement times is only 27 minutes. [3] demonstrate that there is delay between when I/B/E/S receives analyst recommendations from brokers and when they are 'activated' and widely disseminated to third-party clients. We show that *FLY* disseminates recommendation changes to its clients almost a full day quicker, on average, compared to the Thomson Reuters' activation time.

Second, we show that *FLY* speeds up the price adjustment process. After eliminating confounding events, day 0 market reactions for leaked upgrades (downgrades) are 1.86% (-2.06%) with subsequent price drift over days (+1, +4) of 0.11% (-0.50%). This compares to corresponding non-leaked upgrades (downgrades) on day 0 of 0.33% (-0.48%) and drift over days (+1, +4) of 0.77% (-1.05%). To further bolster our evidence, we focus on a subset of recommendations where the I/B/E/S time stamp is released before the market open and

[2] In the context of the [45] model, as financial technology improves, the incentives to produce new information fall. In the case of *FLY* and other third-party news aggregators, traders may develop strategies to profit from the recommendations that are distributed on these platforms with delay.

FLY is released after 9:45 AM. This timing gap allows us to disentangle the market impact of the recommendation from the subsequent *FLY* disclosure. Using 5-minute intraday returns, we find an economically large and statistically significant 5-minute return in the *FLY* release window. Specifically, we find that the 5-minute *FLY* return is approximately 1/3 (1/2) of the magnitude of the initial response to upgrades (downgrades). Thus, the redistribution of analyst recommendations to *FLY* subscribers, despite delayed by at least 15 minutes, has a large *incremental* impact on prices.

We next assess the influence of *FLY* on brokers' ability to provide informational value to their clients by examining the trade execution quality of recommending brokers. While clients are not obligated to trade through the issuing broker, many do.[3] Similar to [52], we find evidence of improved trade execution quality for recommending brokers, but *only* for non-leaked recommendations. Brokers do not improve execution quality for leaked recommendations. This provides compelling evidence that the increased speed of disclosure eliminates benefits of trading through the issuing broker.

We use the court case of Barclays Capital, Merrill Lynch, and Morgan Stanley against *Theflyonthewall.com* to show that the increased speed of information disclosure has negatively impacted the scope of the sell-side research industry. We analyze two key court dates of the lawsuit. On March 18, 2010, the District Court ruled in favor of the plaintiffs and prohibited *FLY* from redistributing recommendations for the greater of 30 minutes after the market open or 10:00am. We find that publicly-traded brokerage houses experienced an abnormal return of 0.77% over the [0,+1]-window surrounding the court's decision. On June 20, 2011, however, the judgement of the lower court was reversed in the United States Court of Appeals Second Circuit suggesting that *FLY* did not violate the misappropriation doctrine. Upon this widely publicized announcement, we find that publicly-traded brokerage houses experienced a statistically significant abnormal return of -1.33% over the [0,+1]

[3][68] finds that analysts' forecasts in general produce significant trading commissions for their broker. Likewise, [67] finds analysts' buy recommendations increase brokers' trading volume in the recommended firm whereas [69] finds that more optimistic and reputable analysts generate more trading commissions for their firm.

window.[4] This translates into an aggregate wealth loss of nearly $11 billion for the sample of publicly-traded brokerage firms. The results are stronger for brokers that have higher leak rates and for brokers that derive a greater proportion of their revenues from commission business. Given that *FLY*'s leakage of recommendations affects the recommending broker's ability to add informational value to their clients, we examine the impact on the scope of the industry. Around the June 20, 2011 reversal decision, we find a precipitous drop in both the number of analysts employed by brokerage houses and the total number of firms covered.

Overall, our evidence suggests that by increasing the speed by which recommendations are distributed, *FLY* has negatively impacted the scope of the sell-side research industry. While leaked recommendations facilitate price discovery, rapid price discovery makes it harder for brokerage clients to take advantage of the gains created by their recommendations. These results are consistent with the theoretical models of [57, 41, 104] and contribute to an emerging literature on the proliferation of big data in capital markets and how financial technology impacts firms' information environment ([56, 43, 44, 107]). Our empirical results tie to [45]'s model, which suggests that financial technology boosts information processing efficiency, but simultaneously reduces incentives to produce information pertaining to fundamentals.

The rest of this paper proceeds as follows. Section 3.2 describes the literature on speed in financial markets and describes the court case of *FLY* vs. several major brokers. 3.3 presents the data and provides empirical evidence on leakage rates, market responses to recommendations, and broker execution quality. Section 3.4 analyzes the impact of the court case. Section 3.5 presents concluding comments and discusses implications for the sell-side research industry.

[4] For instance, see https://dealbook.nytimes.com/2011/06/20/wall-street-banks-lose-ruling-on-research/.

3.2 *Theflyonthewall* and the Speed of Disclosure

3.2.1 Speed and Analysts in Financial Markets

The digital age has revolutionized financial markets. From high frequency trading to the use of big data to crowdsourced financial technology platforms, the way in which investors gather and process information has witnessed a remarkable change in just the past two decades. Ultimately, the digital economy has reduced information acquisition costs and increased speed, but the question of whether this has increased price efficiency is open for debate ([45]). An emerging literature is focused on how financial technology and speed has impacted price discovery from order flow ([24, 15, 47]). For instance, investors are willing to pay considerable sums for real-time access to consumer sales data and/or satellite imagery of retail firms parking lots ([48, 107]) or crowdsourced investment platforms ([70, 56]).

Speed is of critical importance for the business model of sell-side research. Brokers rely on speed to get information to their clients so they have an opportunity to act before non-clients free-ride on that information. [3] demonstrate the importance of speed for sell-side research. They exploit 'activation delay'–the delay in processing time between when the recommendation is received in the I/B/E/S system (announcement time) and when it is distributed and disseminated to clients on the platform (activation time). Delays are associated with price inefficiencies: price reactions are muted and post-recommendation drift is observable. The introduction of financial technology firms that redistribute analysts' recommendations has the potential to increase speed to market participants. On one hand, akin to [3] increased speed should facilitate price discovery. However, rapid disclosure to non-clients potentially threatens the business model of sell-side research by creating disincentives to produce information. In fact, several large brokers sued *FLY* for this very reason. The lawsuit is discussed below.

3.2.2 Lawsuit: Barclays Capital, Merrill Lynch, and Morgan Stanley vs. *Theflyonthewall.com*

In June 2006, Barclays Capital, Merrill Lynch, and Morgan Stanley sued *Theflyonthewall* to prevent them from redistributing their recommendations. In the lawsuit, the brokers had two claims. First, they claimed *FLY* violated copyright laws by taking excerpts directly from their reports (this was easily settled). Second, and the most important and heavily debated, they claimed *FLY* violated the hot news misappropriation doctrine. This tort prevents the defendant from free-riding on time sensitive information.

The key issue in the case is the ability of the broker to communicate recommendations to their clients with sufficient time for them to place a trade. If their recommendations are redistributed before the market opens, the brokerage firms do not have sufficient time to make their sales pitch to generate commissions.[5] Note that while *FLY* was targeted in the case, it is not the only news aggregator that redistributes research. As the court documents illustrate, the outcome of the case would have major implications on other players with similar business models.

"It bears noting that it does not matter to the Firms whether the unauthorized distribution is through a small internet company like Fly or through media giants like Bloomberg, Thomson Reuters, or Dow Jones. The damage is caused not by the identity of the publisher, but by the timely and systematic unauthorized redistribution of the Firms' Recommendations, whatever the medium. To that end, through conference calls and face-to-face meetings with mainstream media, the Firms have objected to the systematic publication of their Recommendations. At least one mainstream publisher of financial news has represented that it is watching this litigation against Fly closely and will adjust its practices based on its evaluation of the outcome of this litigation. The Firms have also sent cease-and-desist

[5] For example, Sidoti & Company, a small boutique sell-side research company pulled their research from the Bloomberg platform because they found the *FLY* would distribute their rating changes before the market open, which is often sooner than they could convey this information to their own clients. "Sidoti's fear is that buy-side clients will value Sidoti's research less, or they will be willing to pay less for it, because their recommendations are available quite inexpensively over a service like the Fly." http://www.integrity-research.com/waiter-theres-a-fly-in-my-distribution-platfom/.

letters to several of Fly's competitors in the online newsfeed niche market."

As we describe later, in March 2010, the District Court initially found in favor of the plaintiffs. However, *FLY* quickly appealed and in June 2011, the decision was overturned. We exploit this court- imposed variation to determine the wealth consequences of increased speed on brokerage houses.[6]

3.3 Leaked Recommendations and Market Impact

3.3.1 Data Sources

We collect recommendation data from I/B/E/S from 2009 to 2013. We include only recommendation upgrades and downgrades issued on U.S. firms and exclude recommendations with an analyst code equal to zero. We assign each analyst into one of 24 GICS industry groups based on the industry in which the analyst issued the largest fraction of his or her recommendations. Several papers suggest that recommendations issued by all-star analysts are significantly more informative than those issued by non-stars. Similar to [35], we obtain all-star analyst data from *Institutional Investor* magazine.

We collect *FLY* data from *LexisNexis* from 2009 to 2013. We include only recommendation upgrades and downgrades. For each observation, we capture the firm name and ticker, time stamp, and the brokerage house issuing the recommendation. We merge this data with our I/B/E/S sample.

Institutional trade data are obtained from Ancerno Solutions. This data source has been used in a number of previous studies such as [52, 65, 66]. The data cover equity transactions and include information on the stock traded, the direction of the trade (buy or sell), the date of the transaction, the number of shares traded, the transaction price, and the commission paid by the institution. We filter the Ancerno data to exclude non-U.S. stocks

[6] In overturning the verdict, the Court essentially ruled that recommendations were news and *FLY* was free to report the news. As Judge Robert Sack wrote in the court's opinions, "A firm's ability to make new–by issuing a recommendation is likely to affect the market price of a security–does not give rise to a right for it to control who breaks that news and how." The court documents can be found here: https://www.eff.org/files/fotwopinion.pdf

and transactions without information on either the broker or the stock symbol. We compile data on the top 200 brokerage houses in terms of total dollar volume over the sample period. These brokerage houses account for 98.5% of the total market share.

For our sample period, we find that the data from Ancerno captures 3.95% of trading volume for the universe of CRSP listed stocks. This is similar to the findings in [66], who find that the Ancerno data captures 4.68% of total market trading volume in 2010. Appendix C.2 describes the key variables in our analysis.

3.3.2 Descriptive Statistics

Table 3.1 provides descriptive statistics of our recommendation sample from I/B/E/S along with recommendations that were leaked on *FLY*. Of the 83,950 recommendation revisions roughly half are upgrades and have an average announcement return of 2.3%. The corresponding announcement return for downgrades is -2.5%. We find that 58.4% of recommendations are leaked on *FLY*. [23] find that the majority of recommendations are released during the pre-market. Conditional on *FLY* leaking the recommendation, 83.9% are reported before the market opens. Consistent with prior studies, we find that approximately 20% of recommendations are confounded by other events.

[3] suggest that the activation time in I/B/E/S captures when the recommendation is distributed widely to I/B/E/S clients. We analyze the difference between *FLY* and I/B/E/S activation times relative to I/B/E/S announcement times. Similar to [3], Panel A of Table 3.2 shows that there is considerable delay between I/B/E/S announcement times and activation times. Not surprisingly, these differences are statistically significant each year. Interestingly, the reporting delay on *FLY* has monotonically increased and more than doubled (44 minutes to 98 minutes) throughout our 5-year sample. This is at least consistent with the view that brokers' efforts to plug leaks may have mitigated the speed in which their recommendations end up on third-party platforms such as *FLY*. On the other hand, I/B/E/S processing delay has declined over time.

Table 3.1: Descriptive Statistics

This table provides descriptive statistics for our sample of analyst recommendations over the 2009 to 2013 period. # Recs is the number of recommendation revisions. % Upgrades is the percentage of recommendation revisions that are upgrades. Upgrade (Downgrade) return is the 2-day [0, 1] CAR around the revision announcement. % Disclosed on FLY is the percentage of revisions that were leaked on *Theflyonthewall.com*. % *FLY* Recs released pre-market is the percentage of *Theflyonthewall.com* recommendations that are released before the market opens (i.e., 9:30am). % Recs Non-Confounded are the percentage of recommendations that are not confounded by other news events. Analyst data are from I/B/E/S and *Theflyonthewall.com*. Stock price data are from CRSP.

Year	# Recs	% Upgrades	Upgrade return	Downgrade return	% Disclosed on FLY	% FLY Recs released Pre-market	% Recs Non-Confounded
2009	18,809	49.03%	3.32%	-3.39%	56.43%	82.98%	77.40%
2010	16,085	50.39%	2.10%	-2.13%	61.04%	82.43%	80.55%
2011	17,997	48.78%	1.78%	-1.97%	55.65%	84.08%	78.88%
2012	16,964	41.48%	1.77%	-2.40%	55.99%	84.08%	79.32%
2013	14,095	42.66%	2.35%	-2.54%	64.19%	85.98%	80.72%
Total	83,950	46.64%	2.30%	-2.50%	58.36%	83.86%	79.27%

Panel B of Table 3.2 shows the distribution of time delay for *FLY* and I/B/E/S activation times. Note that in 32% of cases *FLY* has a time stamp that *precedes* the I/B/E/S announcement time stamp. This is consistent with the findings in [23] that the I/B/E/S *announcement* time stamp is often delayed. *FLY* leaks approximately 26% recommendations within one hour of the I/B/E/S announcement time. Thus, close to 60% of sample recommendations either precede or occur within one hour of I/B/E/S announcement times. In comparison, 21% of sample recommendations are processed within one hour on the I/B/E/S platform. The largest percentage of recommendations (25%) take more than 24 hours to be processed and disseminated in I/B/E/S.

3.3.3 Which Recommendations Are Likely to Be leaked?

In this subsection, we examine the determinants of whether a recommendation is leaked by *FLY*. We model the disclosure decision on *FLY* by including firm, bank, analyst, and recommendation characteristics. We estimate both linear probability and logistic regression models in an effort to understand the determinants that *FLY* leaks a recommendation. The linear probability model allows for the inclusion of fixed effects, which is likely important in this context. We include analyst and day fixed effects. For the reader's interest, we also report results using the logistic model.

The dependent variable is equal to one if a recommendation is leaked on *FLY* and zero otherwise. We include log market capitalization at time t-1 relative to the recommendation release date (SIZE) following [20], Tobin's Q computed at the fiscal year end prior to the recommendation (TOBIN_Q) following [79], and the absolute value of past returns ([38]). Specifically, we use the abnormal return over the (-20,-1) window before the recommendation release date (RUNUP). We include a dummy variable to capture the recommendation revision type (UPGRADE) and the magnitude of the revision change (REVISION_MAGNITUDE) based on I/B/E/Ss numerical scale. Following [100, 58] and others, we capture if the analyst issuing the recommendation change is an all-star analyst in year t based on *Institutional Investor*'s annual poll (ALL_STAR), whether the brokerage house is one of the banks named in the lawsuit (PLAINTIFF), log of the size of the brokerage house based on the number of analysts employed (BROKER_SIZE) following [36], a neglected stock indicator (NEGLECTED) that takes the value one if the firm has less two or fewer analysts covering the stock, and the natural logarithm of one plus the percentage of shares held by institutional investors (INSTITUTIONAL_OWNERSHIP). Formally, our model takes the form:

$$\begin{aligned}
&LEAK = \beta_1 SIZE + \beta_2 TOBIN_Q + \beta_3 RUNUP + \beta_4 UPGRADE \\
&+ \beta_5 REVISION_MAGNITUDE + \beta_6 ALL_STAR + \beta_7 PLAINTIFF \\
&+ \beta_8 BROKER_SIZE + \beta_9 NEGLECTED + \beta_{10} INSTITUTIONAL_OWNERSHIP \\
&+ Analyst\ FE + Day\ FE + \epsilon
\end{aligned} \tag{3.1}$$

Table 3.3 provides these results. Model 1 includes all firm, recommendation, and brokerage house characteristics, but no fixed effects. SIZE and TOBIN_Q are statistically significant indicating that the recommendations of smaller, growth firms are more likely to be leaked. The coefficient on the recent stock price performance, RUNUP, is 0.09, which implies that *FLY* is more likely to leak recommendations when there is more extreme stock price performance before the recommendation date. The greater the revision magnitude, the higher the probability that a recommendation will be leaked on *FLY*. Importantly, while larger brokers appear to be leaked more frequently, plaintiff's recommendations are significantly more likely to be leaked. In economic terms, plaintiffs' recommendations are approximately 10% more likely to be leaked. Recommendations for neglected stocks are less likely to be leaked.

Models 2-4 introduce analyst and day fixed effects. The results are similar across these alternative specifications. In specification (4) with analyst and day fixed effects, all-star analysts are 3% more likely to have their recommendations leaked by *FLY*. The coefficients on PLAINTIFF and BROKER_SIZE remain robust in these specifications. Specification (5) reports estimates from a logit model. The results are similar to the linear probability models reported in specifications (1) to (4).

3.3.4 Market Impact of Leaked Recommendations

In this section, we consider the market impact of leaked *FLY* recommendations. We use standard market model methods to compute abnormal returns around recommendation revisions. The maximum estimation window is 255 days and ends 46 days before the recommendation release date. [81] and [23] acknowledge the need to control for confounding events when analyzing recommendations because they often coincide with other news events such as earnings announcements. We follow the criteria in [81] and eliminate recommendations occurring on earnings announcement days, days when multiple analysts simultaneously issue recommendations on the same stock on the same day, and outlier returns above/below the 5% tail.

Panel A of Table 3.4 provides univariate results. For upgrades, pre-announcement returns over (-20, -1) are negative, but are not significantly between recommendations leaked on *FLY* and those that are not leaked. On day 0, however, we find large differences between leaked and non-leaked recommendations. For instance, leaked upgrades generate an average market response of 1.86% compared to 0.33% for non-leaked upgrades. This difference of almost 1.5% is highly significant. Interestingly, the Day +1 return for recommendations not leaked on *FLY* generates a return of 0.64%, while leaked recommendations display a corresponding return of 0.12%. The difference is also highly statistically significant. Most of the post-recommendation reaction is captured in day +1 as summing days (+1, +4) reveals a similar pattern.

Table 3.2: *FLY* Versus I/B/E/S Dissemination Delay

This table provides the delay in minutes relative to the I/B/E/S announcement times. I/B/E/S Delay (*FLY* Delay) is the difference in minutes between the I/B/E/S activation time stamp (*FLY* time stamp) relative to the I/B/E/S announcement time. Analyst data are from I/B/E/S and *Theflyonthewall.com*. The sample includes only those observations for which there is a *FLY* time stamp and I/B/E/S activation time stamp.

Panel A: Difference in Delay Between *FLY* and I/B/E/S

Year	Fly Delay	I/B/E/S Delay	P-value diff
2009	44.83	1323.64	0.00
	(11)	(477)	(0.00)
2010	64.92	1712.18	0.00
	(19)	(657)	(0.00)
2011	72.25	1596.23	0.00
	(21)	(542)	(0.00)
2012	84.04	1252.06	0.00
	(34)	(274)	(0.00)
2013	97.52	1029.37	0.00
	(64)	(132)	(0.00)
Overall	71.78	1389.55	0.00
	(27)	(419)	(0.00)

Panel B: Delay Relative to I/B/E/S Announcement Time

	Fly time	I/B/E/S Activation time
Before I/B/E/S Announcement	32.27%	0.00%
0 to 1 hour	26.24%	21.09%
1 to 4 hours	17.34%	16.35%
4 to 8 hours	22.74%	15.24%
8 to 24 hours	1.41%	21.45%
Greater than 24 hours	0%	25.87%
Total	100.00%	100.00%

Table 3.3: Probability That a Recommendation Gets Leaked

This table presents results from a linear probability model (Column 1-4) and a logistic model (Column 5) where the dependent variable is binary and equal to one if a recommendation gets leaked on *Theflyonthewall.com* and zero otherwise. See Appendix C.2 for variable definitions. Analyst and day fixed effects are included in various models. Analyst data are from I/B/E/S and *Theflyonthewall.com*. Stock price data are from CRSP. P-values are reported in parentheses. *,**,*** indicates statistical significance at the 10%, 5%, and 1% level.

	(1)	(2)	(3)	(4)	(5)
Firm Characteristics:					
SIZE	-0.03***	0.00*	-0.03***	-0.00	-0.11***
	(0.00)	(0.06)	(0.00)	(0.96)	(0.00)
TOBIN_Q	0.01***	0.01***	0.01***	0.00**	0.03***
	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)
RUNUP	0.11***	0.10***	0.14***	0.13***	0.50***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Rec Characteristics:					
UPGRADE	-0.01***	-0.01*	-0.01**	-0.00	-0.05***
	(0.00)	(0.07)	(0.02)	(0.39)	(0.00)
REVISION_MAGNITUDE	0.09***	0.06***	0.07***	0.03***	0.38***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Bank & Analyst Characteristics					
ALL_STAR	-0.02**	-0.01	0.01	0.03***	-0.09**
	(0.04)	(0.22)	(0.17)	(0.01)	(0.04)
PLAINTIFF	0.10***	0.13***	0.07***	0.14***	0.47***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
BROKER_SIZE	0.02***	0.05***	0.03***	0.05***	0.09***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
NEGLECTED	-0.08***	-0.05***	-0.09***	-0.05***	-0.35***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
INSTITUTIONAL_OWNERSHIP	0.05***	0.01***	0.04***	0.01***	0.20***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
Analyst FE		✓		✓	
Day FE			✓	✓	
Observations	76,187	75,483	76,124	75,418	76,187
Adj (Pseudo) R2	0.022	0.306	0.111	0.373	0.017

Table 3.4: CARs Based on Leaked Versus Non-Leaked Recommendations

This table provides abnormal returns to recommendations based on whether recommendation revisions are released on *Theflyonthewall.com* after removing confounding events, which include recommendations occurring on earnings announcement days, days when multiple analysts simultaneously issue recommendations on the same stocks, and outlier returns above/below the 5% tail. Panel A provides recommendation CARs for all non-counfounded observations and Panel B focuses on all-star analysts. Panel C provides recommendations CARs after matching leaked and non-leaked recommendations by firm characteristics used in Table 2 and analyst using coarsened exact matching (CEM). Panel D provides recommendations CARs after matching leaked and non-leaked recommendations on exact firm and analyst. See Appendix C.2 for variable definitions. Analyst data are from I/B/E/S and *Theflyonthewall.com*. Stock price data are from CRSP. P-values are reported in parentheses. *,**,*** indicates statistical significance at the 10%, 5%, and 1% level.

	(-20, -1)	0	+1	+2	+3	+4	(+1,+4)	(0,+4)
Panel A: Recommendation CARS								
Upgrades:								
Not Leaked	-0.93%***	0.33%***	0.64%***	0.05%***	0.03%	0.04%*	0.77%***	1.10%***
Leaked	-0.82%***	1.86%***	0.12%***	0.03%*	0.01%	-0.05%***	0.11%***	1.97%***
p-value difference	(0.45)	(0.00)	(0.00)	(0.31)	(0.49)	(0.00)	(0.00)	(0.00)
Downgrades:								
Not Leaked	0.05%	-0.48%***	-0.75%***	-0.19%***	-0.05%***	-0.05%**	-1.05%***	-1.52%***
Leaked	0.48%***	-2.06%***	-0.23%***	-0.15%***	-0.05%***	-0.07%***	-0.50%***	-2.55%***
p-value difference	(0.00)	(0.00)	(0.00)	(0.15)	(0.91)	(0.38)	(0.00)	(0.00)
Panel B: Recommendation CARs for All-Star Analysts								
Upgrades:								
Not Leaked	-0.93%	0.42%***	1.47%***	0.02	%-0.07%	-0.11%	1.28%***	1.70%***
Leaked	-0.42%	2.30%***	0.06%	0.00%	-0.06%	-0.05%	-0.04%	2.27%***
p-value difference	(0.49)	(0.00)	(0.00)	(0.89)	(0.94)	(0.64)	(0.00)	(0.10)
Downgrades:								
Not Leaked	0.29%	-0.96%***	-1.86%***	-0.01%	-0.05%	0.11%	-1.86%***	-2.81%***
Leaked	-1.11%**	-2.58%***	-0.18%*	-0.12%	0.08%	-0.23%***	-0.46%***	-3.02%***
p-value difference	(0.11)	(0.00)	(0.00)	(0.50)	(0.44)	(0.01)	(0.00)	(0.53)
Panel C: Recommendation CARs After Exact Matching on Firm and Analyst								
Upgrades:								
Not Leaked	-1.01%***	0.39%***	0.50%***	0.17%***	0.00%	0.03%	0.71%***	1.11%***
Leaked	-0.94%***	1.68%***	0.13%**	0.09%**	-0.03%	-0.10%	0.11%	1.79%***
p-value difference	(0.85)	(0.00)	(0.00)	(0.29)	(0.75)	(0.06)	(0.00)	(0.00)
Downgrades:								
Not Leaked	-0.63%**	-0.58%***	-0.76%***	-0.25%***	-0.07%	-0.08%	-1.14%***	-1.71%***
Leaked	0.31%	-1.94%***	-0.32%***	-0.04%	0.00%	-0.06%	-0.41%***	-2.34%***
p-value difference	(0.01)	(0.00)	(0.00)	(0.00)	(0.27)	(0.80)	(0.00)	(0.00)

We observe the same pattern when focusing on downgrades. A significant price drift occurs for non-leaked downgrades. Downgrades leaked on *FLY* generate an initial Day 0 price reaction of -2.06% with a drift over the next four days of -0.50%. Downgrades that are not leaked by *FLY* result in an announcement period return of -0.48% and a drift of -1.05% over the next four days. Overall, these results suggest that leaked recommendations appear to aid in price discovery. From the perspective of the brokerage industry, this evidence is consistent with the claim that an increase in the speed of price discovery is likely detrimental to their clients' interests; faster price adjustments may prevent them from benefiting from the stock price changes associated with the recommendations. These findings are consistent with the results in [63] and [61] that information slowly diffuses into stock prices, particularly when information is released to different parties with delay.

An alternative explanation for these market reaction results is that the *FLY* picks up more important recommendations resulting in larger price reactions and rapid price discovery. This is a legitimate concern as Table 3.3 shows that leaked recommendations are more likely to come from more reputable brokers and all-star analysts. In Panel B, we report results for all-star analysts. The results mirror the full sample evidence in both panels. For instance, all-star upgrades and downgrades released on *FLY* have significantly larger market reactions at time 0 compared to all-star recommendations not leaked on *FLY*. Leaked recommendations have little corresponding post-recommendation drift whereas non-leaked recommendations are associated with drift. In addition, the aggregate reaction over the (0, +4)-day period are similar for recommendations leaked on*FLY* and the ones not leaked on *FLY*, mitigating the concern that the immediate response to recommendations leaked on *FLY* is due to their salient information.[7] Thus, while systematic differences in what recommendations *FLY* chooses to leak is unlikely to explain the differences between leaked versus non-leaked market reactions, some caution is warranted in interpreting the evidence as such.

[7] We also find that the same pattern emerges for recommendations from reputable banks. We define top brokers as those that are in the top 10% in terms of number of analysts employed in a given year.

In Panel C, we compare the stock price reaction to leaked and non-leaked recommendations after requiring an *exact* match on firm and analyst. For example, we compare the market reaction to upgrade recommendations for analyst John S. at Goldman Sachs on Walmart that was leaked on *FLY* to his upgrade on Walmart that wasnt leaked on *FLY* on different dates. For both upgrades and downgrades, we see significantly larger Day 0 price reaction for leaked recommendations. The price drift for recommendations not released on *FLY* is significantly larger than those that are leaked. The economic magnitudes are roughly similar across all specifications.

Despite matching on firm and analyst, we cannot be certain that *FLY* releases recommendations from John S. that are more interesting and will deliver greater market returns. Nonetheless, the result survives a number of matching controls that suggests leaked recommendations on *FLY* facilitate price discovery. These results are consistent with the empirical findings in [107] and [56] that FinTech companies are improving price informativeness by decreasing the costs of information acquisition.

3.3.5 Intraday Price Discovery Around *FLY* Announcements

In this section, we provide additional evidence that *FLY* announcements facilitate price discovery. We focus on the subset of recommendations where the I/B/E/S time stamp precedes the market open (i.e., before 9:30am) and *FLY* releases the recommendation during market hours (after the 9:30am open). We identify 823 upgrades and 988 downgrades that meet this criteria.

Table 3.5, Panel A shows the distribution of this subset of recommendations. Conditional on the *FLY* time stamp occurring after the market open, most occur very early in the trading day. For instance, approximately 45% (46%) of upgrades (downgrades) are leaked between 9:30 and 9:35, while 18.71% (15.89%) of upgrades (downgrades) are released between 9:35 and 9:40.

Table 3.5 Panel B examines intraday returns surrounding recommendation releases on

FLY. In order to disentangle the price effects of *FLY* from the initial price reaction to the recommendation at the market open, we require that the *FLY* time stamp be after 9:45am. This minimum 15-minute time gap should allow for a clean identification of any incremental impact from the *FLY* release. This particular test has the added benefit that it should help ease selection concerns.

The overnight return (4pm to 9:30am) to I/B/E/S pre-market recommendation upgrades (downgrades) is a statically significant 1.22% (-0.81%), both of which are similar in magnitude to the overnight returns documented in [23]. The 9:30am to 9:35am return to recommendation upgrades is an insignificant 0.03%, while the corresponding return to downgrades is a statistically significant -0.39%.

Five-minute returns around the *FLY* time stamp suggests that *FLY* indeed aids in the price discovery process. Recommendation upgrades (downgrades) generate a statistically significant return of 0.36% (-0.37%) in the five minute window containing the *FLY* time stamp. These returns are large compared to the initial price reactions for upgrades and downgrades (1/3 and 1/2 of the initial market reaction, respectively). There is minimal drift over the five minute windows from +1 to +5 for both upgrades and downgrades. This evidence suggests that *FLY* disclosure provides an incremental market reaction as the information becomes more broadly diffused to market participants.

As a final test of the price reaction to leaked recommendations, we focus on the stock price reaction around the I/B/E/S activation date. As discussed previously, [3] suggest that the activation time corresponds to when forecasts are widely distributed to I/B/E/S clients. As in Table 3.4, we filter out confounded recommendations and require that the activation date occur at least one trading day after the announcement of the recommendation. Similar to [3], we observe a statistically significant stock price reaction on the activation date for both recommendation upgrades and downgrades. Upgrades (downgrades) generate a 0.46% (-0.57%) abnormal return.

If *FLY* speeds up the price adjustment process, then we should see a smaller return on the activation date for leaked recommendations than for non-leaked recommendations. We find that this is the case. For recommendation downgrades, leaked recommendations generate a -0.21% return on the activation date versus -0.81% for non-leaked recommendations. The difference is both economically and statistically significant. We find a similar result for upgrades. Leaked upgrades generate a 0.12% return on the activation date, while non-leaked upgrades generate a 0.70% return. Note that even for leaked recommendations, the activation return is statistically significant suggesting that there is still incremental information in these recommendations.

3.3.6 Execution Quality

In this section, we explore whether the early release of recommendations of *FLY* impacts the price improvement provided by brokerage houses. [52] examine institutional trading around recommendation revisions and find that clients trading through the recommending broker receive a price that is significantly better than trades through non-recommending brokers. They argue that this is a particularly powerful test of the informational advantage of being a client of the broker since trades routed through the broker, by definition, identifies clients.

We replicate the analysis in [52] using our sample of recommendations. This analysis is reported in Table 3.7. We require that more than 30 trades are executed on the recommended stock day. Panel A reports the number of trades, price improvement relative to the value-weighted average price (VWAP), price improvement relative to the close price, average share and dollar volume, and the corresponding market share.

We find non-issuing and issuing brokers do not provide price improvement relative to VWAP, but issuing brokers provide about 1.4 cents per share improvement from the close price relative to non-issuing brokers (4.33 cents versus 2.91 cents).

While the above results are consistent with [52], they suggest that commissions and price improvement surrounding recommendations has dropped significantly. For example,

[52] also find significant price improvement relative to VWAP and Close of 4.4 cents and 10.2 cents, respectively. We find no price improvement relative to VWAP and our reported price improvement relative to close is significantly smaller. This suggests that the value of being a long-term client of a full-service broker has declined.

More importantly, in Panel B of Table 3.7, we compare the price improvement for the sample of trades placed with the recommending broker for leaked versus non-leaked recommendations. In this case, we find significant price improvement for non-leaked recommendations relative to leaked ones for both the VWAP and close price. The spread is 4.57 cents and 2.46 cents, respectively. Consistent with the market reaction results, brokers' clients also benefit when recommendations are not leaked in terms of execution quality. This is not the case for their leaked recommendations. This suggests that the increased disclosure speed by recommendations leaked on *FLY* hampers the ability of brokers to offer price improvement.

In Panel C we perform coarsened exact matching on various characteristics related to trade complexity. Specifically, we match on: share (dollar) volume of the trade; and share (dollar) volume of the trade relative to the total day's trading volume. The results are robustbrokers can offer significantly more price improvement on non-leaked recommendations.

3.4 Impact of Court Case on Brokers and Analysts

In this section, we use the court case of Barclays Capital, Merrill Lynch, and Morgan Stanley against *Theflyonthewall.com* to examine the impact of the speed of information dissemination on the sell-side research industry. We analyze two key court dates of the lawsuit. On March 18, 2010, the District Court ruled in favor of the plaintiffs and concluded that *FLY* committed copyright infringement and misappropriation of hot news. In addition to monetary penalties, a permanent injunction was issued prohibiting *FLY* from redistributing recommendations for the greater of 30 minutes after the market open or 10:00am. On

June 20, 2011, however, the judgement of the lower court was reversed in the United States Court of Appeals Second Circuit. Section 3.4.1 presents event study results for brokers around key court dates. Section 3.4.2 investigates the impact on the scope of the research industry.

3.4.1 Event Study

We begin our analysis of the impact of the court case to brokers by examining the market reactions of banks around the initial lower court ruling and the subsequent Court of Appeals ruling. We identify a sample of 28 publicly-traded brokerage houses that produce sell-side research. We then employ standard event study methods and estimate equally-weighted abnormal returns for various windows. Like with recommendations, the maximum estimation window is 255 days and ends 46 days before the recommendation release date. Following [85], we compare announcement period returns to a control sample of Broker/Dealers (SIC code 6211) that do not produce sell-side research. The results are presented in Table 3.8, Panel A.

Table 3.5: Intraday Returns Around *FLY* Announcements

This table focuses on the subset of recommendations where the I/B/E/S timestamp occurs before the market open (9:30am) and the *FLY* time stamp is after the market open. Intraday trade data is obtained from TAQ. We require that sample firms have sufficient liquidity to calculate 5-minute returns. Panel A reports the timing of *FLY* announcements that occur after the market opens. Panel B computes 5-minute returns surrounding the recommendation announcement on *FLY*. We only include *FLY* announcements that occur on or after 9:45, so that we can separately identify the I/B/E/S announcement effect and *FLY* announcement effect. We report results separately for upgrades and downgrades. *,**,*** indicates statistical significance at the 10%, 5%, and 1% level.

Panel A: Distribution of *FLY* Announcements After the Market Opens

Time Window	Upgrades	Downgrades
9:30 to 9:35	44.71%	46.26%
9:35 to 9:40	18.71%	15.89%
9:40 to 9:45	6.93%	8.91%
9:45 to 9:50	4.25%	4.25%
9:50 to 9:55	1.82%	2.33%
9:55 to 10:00	1.58%	2.63%
10:00 to 4:00	22.00%	19.73%
# of observations	823	988

Panel B: Five Minute Returns Surrounding *FLY* Release Times

Time Window	Upgrades	Downgrades
4:30 to 9:30	1.22%***	-0.81%***
9:30 to 9:35	0.03%	-0.39%***
Five-minute returns around the FLY time stamp		
-1	0.17%***	-0.07%**
0	0.36%***	-0.37%***
1	0.04%	0.02%
2	-0.02%	0.05%*
3	0.07%***	0.01%
4	-0.03%	0.00%
5	0.03%	0.03%

Table 3.6: Announcement Period Returns Surrounding the I/B/E/S Activation Times

This table focuses on announcement period returns surrounding the I/B/E/S activation time, when [3] argue the recommendation is distributed widely to I/B/E/S clients. We examine abnormal returns on the activation date based on whether the recommendation revisions are released on *Theflyonthewall.com*. We remove confounding events, which include recommendations occurring on earnings announcement days, days when multiple analysts simultaneously issue recommendations on the same stocks, and outlier returns above/below the 5% tail. We also require that the recommendation announcement date and the activation date be different. We report results separately for upgrades and downgrades. *,**,*** indicates statistical significance at the 10%, 5%, and 1% level.

	Activation date abnormal return
Upgrades	0.46%***
Downgrades	-0.57%***
Downgrades:	
Not leaked	-0.81%***
Leaked	-0.21%***
p-value difference	0.00
Upgrades:	
Not leaked	0.70%***
Leaked	0.12%***
p-value difference	0.00

Table 3.7: Price Improvement for Brokerage Houses Around Recommendation Releases
This table examines price improvement and brokerage commissions for recommendations issued between 2009 and 2013. Institutional trade data are obtained from Ancerno for the day the recommendation was released. We restrict the sample to recommendation days with more than 30 trades. VWAP is the volume-weighted average price. Price Improvement relative to Close (cents) is the execution price relative to the close on the day of trading. We calculate price improvement using buy (sell) orders for upgrades (downgrades). Share volume is the number of shares transacted per trade. Dollar volume is the transaction price multiplied by share volume. Share volume market share is share volume divided by the daily total volume. Dollar volume market share is dollar volume divided by the daily total dollar volume. *,**,*** indicates statistical significance at the 10%, 5%, and 1% level.

Panel A. Issuing Versus Non-Issuing Brokerage House

	Non-issuing	Issuing	Diff
# of trades	998,062	26,111	
Price improvement relative to VWAP (cents)	-1.35	-1.03	-0.32
Price improvement relative to Close (cents)	2.91	4.33	-1.42***
Share volume	4,577	5,460	-884***
Dollar volume ($)	152,683	167,925	-15,243**
Share volume (%)	0.08	0.12	-0.04***
Dollar volume (%)	0.08	0.12	-0.04***

Panel B: Issuing Brokerage House Price Improvement: Leaked vs. Not Leaked

	Not leaked	Leaked	Diff
# of trades	4,821	21,290	
Price improvement relative to VWAP (cents)	2.69	-1.88	4.57***
Price improvement relative to Close (cents)	6.34	3.88	2.46***
Share volume	5,487	5,454	33
Dollar volume ($)	165,296	168,520	-3,224
Share volume (%)	0.11	0.13	-0.02
Dollar volume (%)	0.11	0.12	0.01

Panel C: Issuing Brokerage House Price Improvement: Leaked vs. Not Leaked (CEM on Volume, Dollar Volume, Share Volume Market Share, and Dollar Volume Market Share)

	Not leaked	Leaked	Diff
# of trades	4,809	4,809	
Price improvement relative to VWAP (cents)	2.73	-1.29	4.02***
Price improvement relative to Close (cents)	6.41	4.3	2.11**
Share volume	4,902	4,185	718**
Dollar volume ($)	142,351	127,806	14,545
Share volume (%)	0.10	0.10	0
Dollar volume (%)	0.10	0.10	0

Table 3.8: Wealth Effects for Brokers Around Key Lawsuit Dates

This table shows market model abnormal returns around key dates in the *Barclays, Inc. v. Theflyonthewall.com* lawsuit. Equal-weighted returns are presented. Dollar abnormal returns are calculated by multiplying the brokerage houses market capitalization from day -1 by the CAR over the [0,+1] window. Panel A presents average abnormal returns for the full sample of publicly-traded brokerages. High leak rate corresponds to those brokerage houses with above median leak rates in 2009. High commissions correspond to those brokerage houses with above median ratios of commission revenue to total revenue in 2009. Panel B presents results for individual brokerages. Stock price data are from CRSP. Inference is based on a bootstrapped p-value. *,**,*** indicates statistical significance at the 10%, 5%, and 1% level.

Panel A: Full Sample CARs

	March 18, 2010		June 20, 2011	
	(0, +1) CAR	%Positive	(0, +1) CAR	% Positive
All brokerage houses	0.77%*	71.43%	-1.33%***	28.57%
Control Sample	0.23%		-0.26%	
Difference	0.54%		-1.07%**	
High leak rate	1.14%**	85.71%	-1.44%***	21.43%
Low leak rate	0.40%	57.14%	-1.21%***	35.71%
Difference	0.74%		-0.21%	
High commissions	1.40%*	71.43%	-1.51***	21.43%
Low commissions	0.14%	71.43%	-1.17***	35.71%
Difference	1.26%		-0.34	

Panel B: Individual Brokerages

		March 18, 2010		June 20, 2011	
Plaintiffs:	Leak Rate 2009	CAR (0,+1)	Value Created	CAR(0,+1)	Value Created
MORGAN STANLEY DEAN WIT	69.85%	0.30%	124.12	-2.38%	-839.99
BARCLAYS PLC	82.80%	1.85%	1200.04	-1.44%	-732.57
BANK OF AMERICA CORP	52.05%	1.14%	1972.5	-1.00%	-1085.96
Other Brokerage Houses:					
THOMAS WEISEL PARTNERS	39.72%	3.63%	4.58	NA	NA
LADENBURG THALMANN FIN	71.74%	4.89%	8.37	-0.95%	-2.19
RODMAN & RENSHAW CAP GRP	54.12%	-2.41%	-3.71	-2.98%	-1.38
MERRIMAN CURHAN FORD	56.98%	7.10%	0.80	-2.21%	-0.15
J M P GROUP INC	53.85%	-2.20%	-4.20	-0.45%	-0.59
F B R CAPITAL MARKETS C	87.60%	2.16%	6.87	1.19%	2.61
OPPENHEIMER HOLDINGS IN	62.28%	-0.81%	-2.78	-2.67%	-9.55
BGC PARTNERS	92.86%	-4.37%	-15.22	-2.05%	-13.60
COWEN GROUP	46.43%	0.80%	3.31	0.67%	1.84
EVERCORE PARTNERS INC A	36.73%	-2.56%	-14.62	-3.44%	-24.13
PIPER JAFFRAY COMPANIES	64.27%	-1.72%	-15.57	-3.26%	-18.44
K B W INC	88.54%	1.46%	13.54	-4.86%	-35.52
STIFEL FINANCIAL CORP	79.53%	1.78%	30.58	1.47%	27.55
ALLIANCEBERNSTEIN HOLDINGS	31.82%	-1.13%	-34.32	0.33%	6.58
LAZARD LTD A	88.35%	1.08%	37.01	0.87%	38.34
RAYMOND JAMES FINANCIAL	83.30%	2.06%	71.38	-0.19%	-7.83
JEFFERIES GROUP INC NEW	36.18%	-1.98%	-91.76	-2.30%	-97.31
SUNTRUST BANKS INC	85.62%	0.38%	53.14	-4.09%	-574.38
DEUTSCHE BANK A G	41.16%	0.13%	61.31	0.30%	159.95
CREDIT SUISSE GROUP	63.02%	0.99%	658.46	0.05%	16.56
ROYAL BANK OF CANDA	76.06%	0.97%	816.70	-0.48%	-363.38
GOLDMAN SACHS GROUP INC	47.67%	2.52%	2289.40	-2.37%	-1681.41
CITIGROUP INC	30.57%	0.37%	421.14	0.27%	295.38
WELLS FARGO & CO	73.31%	3.10%	4905.57	-2.03%	-2941.11
JPMORGAN CHASE & CO	82.99%	2.02%	3523.03	-1.91%	-3100.06

Focusing on the March 18, 2010 announcement date, the initial court verdict produces a positive equally-weighted return of 0.77%. This announcement period return is significant at the 10% level, but not statistically different from the control sample's return of 0.23%. When we examine brokers with a high leak rate (above the median), we find that they experience significantly higher CARs compared to brokers with a low leak rate (i.e., 1.14% versus 0.40%, respectively). However, the difference is not statistically significant. We collect the proportion of revenue brokers derive from trading commissions from their 2009 10-K annual filings. Brokers with high commissions as a percent of revenue (based on medians) experience larger, announcement period returns than low commission firms. However, the difference is again not statistically significant.

On the announcement date of the 2^{nd} Court of Appeals reversal on June 20, 2011, banks experience a statistically significant abnormal return of -1.33%. This return is statistically different from zero and from a control sample of broker/dealers without research divisions and economically meaningful.

Table 3.8 of Panel B reports abnormal returns by individual banks. The three plaintiffs in the lawsuit are noticeably impacted by the court rulings. Bank of America, Morgan Stanley, and Barclays experienced a cumulative dollar gain on the initial verdict of more than $3 billion. Similarly, their market values fell by over $2.5 billion after the Appeal's court ruling. Panel B also highlights significant spillover effects. Brokerage houses not directly involved in the lawsuit are also impacted by the court ruling. In aggregate, banks lost close to $11 billion around the June 20^{th}, 2011 decision.

We interpret the results from this section to indicate that the court rulings had an economically meaningful impact on the sell-side analyst research industry. An alternative explanation is that the market is predicting how regulators will treat firms operating in the financial services industry in future dealings. However, we do not believe this is a valid explanation. If the market's reaction signaled a future unfavorable legal environment for the financial services industry we would expect that broker-dealers without research divisions would be equally impacted by the court case. As we demonstrate, this is not the case.[8]

[8] We also estimate a difference-in-difference regression where we examine dollar trading volume and commissions in the 3- and 6-month period following the court case. These regressions are estimated at the broker-month and include broker and month fixed effects. We find trading volume and commissions decline by an economically and statistically significant amount, which is consistent with the view that brokers lost trading revenue following the court's verdict. We could also examine client-broker switching around the court case. That is, we could identify clients of large brokers with research divisions before the court verdict and examine if they significantly changed their routing of trades to discount brokers after the final verdict. Unfortunately, Ancerno stopped providing client identifiers in 2010 so we are unable to perform this test.

3.4.2 Impact on the Scope of the Sell-Side Industry

In this section, we examine if the lawsuit also impacted the scope of the sell-side industry. We follow an approach similar to [87], who examine how the analyst industry has evolved through time. They find that having more analysts in an industry facilities information flow and price discovery resulting in more efficient markets.[9] They also find that following Regulation Fair Disclosure (Reg FD), the number of analysts in the industry has dropped, consistent with the view that this regulation diminished the value of their services. While the *FLY* case is fundamentally different than regulations specifically targeted at curbing analyst behavior, we would expect that a natural reaction to the ruling would be a decline in the number of analysts and the number of stocks that receive analyst coverage, particularly by larger brokers. Consequently, a reduction in competition would ensue along with the quality of the information environment.

To examine the impact of the court ruling on the scope of the industry, we generally follow the framework of [87]. As a first step, we classify each analyst into one of 24 GICS industries. For each industry and each month, we define an entering analyst as an analyst who issues her first recommendation or an analyst who issues a recommendation after not doing so for a period of twelve months. An exiting analyst is one who issued her last recommendation in the month or does not issue a recommendation over the next twelve months. Changes in the number of analysts are calculated as the difference between entering and exiting analysts. The total number of analysts employed in a given industry during a particular month is simply the running total based on these changes and the prior level. In sum, the output of the above process is the number of active analysts within an industry-month.

To understand the determinants of the scope of the sell-side research industry, we estimate regressions similar to [87]. The sample period runs from 2009 to 2013. The de-

[9] [33] show that hedge funds increase information production around exogenous losses of analyst coverage. In this sense, sophisticated investors act as a substitute in generating information and step in when analysts drop out.

pendent variables are the change in the number of stocks covered and the change in the number of analysts in the industry each month. Included in the model are the monthly portfolio median industry return (MEDIAN_INDUSTRY_RETURN); the monthly number of IPOs in each industry (#_OF_IPOS); the number of monthly delistings in each industry (#_OF_DELISTINGS); the monthly change in industry trading volume (Δ_TRADING_VOLUME); and an indicator variable representing the 18-month period after the appellate court's ruling (POST_2ND_COURT_OF_APPEALS_RULING).[10] Month fixed effects and industry fixed effects are included in the regressions. Inference is based on robust standard errors. The model takes the following form:

$$\Delta\_ANALYSTS_{t+1} = \beta_1 MEDIAN\_INDUSTRY\_RETURN + \beta_2 \#\_OF\_IPOS$$
$$+ \beta_3 \#\_OF\_DELISTINGS + \beta_4 \Delta\_TRADING\_VOLUME$$
$$+ \beta_5 POST\_2^{ND}\_COURT\_OF\_APPEALS\_RULING + Month\ FE + Industry\ FE + \epsilon \tag{3.2}$$

Table 3.9 presents the results. We estimate the full sample of brokers and separately estimate the effect for large brokers. Focusing on the 'All Brokers' Δ_ANALYSTS column, the coefficient on the post-ruling period is negative and significant. This is consistent with the view that the brokerage industry reacted to the ruling by trimming the number of analysts they employ. None of the other variables in the model are significant. [87] find that the number of IPOs and change in trading volume are positively related to the size of the analyst industry, but we do not. This is likely because our sample period is much smaller than theirs (2009-2013), as our focus is to estimate the impact of the court ruling as opposed to describe the evolution of the analyst industry.

In the second model, we estimate equation 2, but substitute the number of stocks covered as the dependent variable. We find that the number of stocks covered is negatively

[10] We obtain similar results if the indicator variable is defined using a 12- or 24-month period after the appellate courts ruling.

Table 3.9: Changes in Analyst Scope

This table presents regression results examining the scope of the analyst industry. The time period is from 2009 to 2013 and the unit of observation is at the month-industry level. The dependent variables are the monthly change in analysts and change in the number of stocks covered within a GICS industry group. The median industry return is the median monthly return in the industry group in month t. The number of IPOs and delistings are the number of IPOs filed and firms delisted in month t, respectively. Change in trading volume is the change in industry trading volume from month t-1 to t. The post 2^{nd} Court of Appeals ruling is an indicator variable for the 18-month time period after June 20, 2011. Industry fixed-effects are included to account for time invariant differences in analyst following across industries. We also include 11 calendar month fixed effects (i.e., February, March, etc.), to account for seasonal differences in the scope of the sell-side research industry. Large brokers are the 20 largest brokers by number of recommendations issued over the 2009 to 2013 time period. P-values are reported in parentheses. *,**,*** indicates statistical significance at the 10%, 5%, and 1% level.

	All Brokers		**Large Brokers**	
	Δ_ANALYSTS$_{t+1}$	Δ_STOCKS_COVERED$_{t+1}$	Δ_ANALYSTS$_{t+1}$	Δ_STOCKS_COVERED$_{t+1}$
MEDIAN_INDUSTRY _RETURN	-0.01	-0.15***	-0.04**	-0.14***
	(0.77)	(0.00)	(0.05)	(0.00)
#_OF_IPOS	-0.00	2.74**	0.63	1.66*
	(0.99)	(0.03)	(0.24)	(0.09)
#_OF_DELISTINGS	0.25	-0.65*	0.02	-0.58*
	(0.41)	(0.09)	(0.91)	(0.07)
Δ_TRADING_VOLUME	-0.03	0.10	-0.04	0.08
	(0.93)	(0.84)	(0.84)	(0.83)
POST_2ND_COURT_OF_APPEALS_RULING	-0.64***	-0.70***	-0.40***	-1.06***
	(0.00)	(0.01)	(0.00)	(0.00)
Month FE	✓	✓	✓	✓
Industry FE	✓	✓	✓	✓
Observations	1,530	1,530	1,508	1,508
Adj R2	0.54	0.15	0.35	0.09

related to the industry return and positively related to the number of IPOs in the industry. The latter result is consistent with results in [87]. Most importantly, we find that the number of stocks covered declines after the court ruling. We separate large brokers in the last two models. In general, the results are similar. In terms of the number of analysts employed and the number of stocks covered, both significantly decline after the appeals decision.

We also consider the impact of the court case on analyst effort. The previous results suggest that brokers, particularly large brokers, reduce the number of analysts and stocks they cover following the lawsuit. It is not clear however if the remaining analysts have to increase their workload and pick up coverage on at least some of the firms that were dropped as a result of analyst departures.

We focus on analysts that were active in the year before/year after the June 20^{th}, 2011 court verdict. We calculate five proxies for analyst effort: number of stocks covered; number of recommendations issued; number of annual earnings forecasts issued; number of quarterly earnings forecasts issued; and number of long-term growth forecasts issued. These results are presented in table 3.10.

Table 3.10: Changes in Analyst Effort

This table examines changes in analyst effort in the year after the June 20th, 2011 court verdict in favor of *FLY*. For analysts covering three or more stocks in the year prior to the court verdict, we calculate five measures of effort: (1) number of stocks covered; (2) number of recommendations issued; (3) number of annual earnings forecasts issued; (4) number of quarterly earnings forecasts issued; and (5) number of long-term growth forecasts issued. These measures are calculated at the analyst level. P-values from a paired t-test are reported in parentheses. *,**,*** indicates statistical significance at the 10%, 5%, and 1% level.

Effort Proxy	Obs	Pre	Post	Diff
Number of stocks covered	2,257	8.15	7.97	-0.18
				(0.10)
Number of recommendations issued	2,257	26.77	26.84	0.07
				(0.89)
Number of annual earnings forecasts issued	2,257	65.26	69.01	3.75***
				(0.00)
Number of quarterly earnings forecasts issued	2,257	56.45	59.32	2.87***
				(0.01)
Number of long-term growth forecasts issued	2,257	5.67	6.41	0.74***
				(0.01)

We find no evidence that analysts increased the number of stocks covered in the year after the June 20^{th}, 2011 court ruling. There is also no evidence that analysts issued more recommendations following the court ruling. We do find a statistically significant increase in the number of annual, quarterly, and growth forecasts issued in the year following the June 20^{th}, 2011 court ruling. Thus, there is some evidence that analysts increased effort on generating earnings forecasts.

Collectively, the results from these tests indicate that the number of stocks receiving coverage declines, but that the average analyst did not increase coverage following the Appeals Court Ruling. We find some evidence that analysts increase their production of earnings forecasts (annual, quarterly, and long-term growth forecasts), but do not produce more recommendations.

3.5 Conclusion

In this paper, we assess the impact of an increase in the speed of recommendation dissemination on the sell-side research industry. Our primary results are as follows. Leaking recommendations on *FLY* facilitates price discovery. Recommendation upgrades and downgrades not leaked on *FLY* experience a statistically significant and economically meaningful drift in the (+1,+4) window following the release of the recommendation. While *FLY*'s leaking of recommendations improved price discovery, the practice had a negative impact on the price improvement offered by brokerage houses to clients. Thus, we find direct evidence that the ability of *FLY* to quickly disseminate recommendations has disrupted the business model of sell-side research. Consistent with such disruption, we observe a decrease in the number of analysts employed and the number of stocks covered. An analysis of announcement period returns surrounding the court case suggests that publicly traded brokerage houses experienced an aggregate wealth loss in excess of $10 billion.

Our paper has important implications for the future of analyst research. The real economic consequences from the speedy dissemination of analyst research are large. As noted, in aggregate we estimate over $10 billion in market value is destroyed. This is likely a lower bound on losses. Many brokerages are private entities, and consequently, we cannot estimate the wealth impact of the court's decision for these firms.

Given the speed with which recommendations are distributed by third party sources, we expect either that brokers will continue to scale down their research divisions or devise a different platform to protect the speed advantage. Brokerage houses are likely to substitute towards broker-hosted investor conferences as in [55] or engage in other concierge and high-touch services for their clients as opposed to generating research reports. These types of research services mitigate the speed issue as private meetings are not easily substituted.

The increasing speed of disclosure also impacts how much brokers can charge buy-side clients. This is a key concern given the new Markets in Financial Instruments Directive (MiFID II) regulations that require the separation of analyst research from brokerage commissions for European firms. Coupling the devaluation of research with these new regulations forcing the unbundling of commissions and research will undoubtedly have a major impact on the industry.

Appendices

APPENDIX A

APPENDIX FOR "IMPACT OF VENTURE CAPITAL FLOWS ON INCUMBENT FIRMS: EVIDENCE FROM 70 MILLION WORKERS"

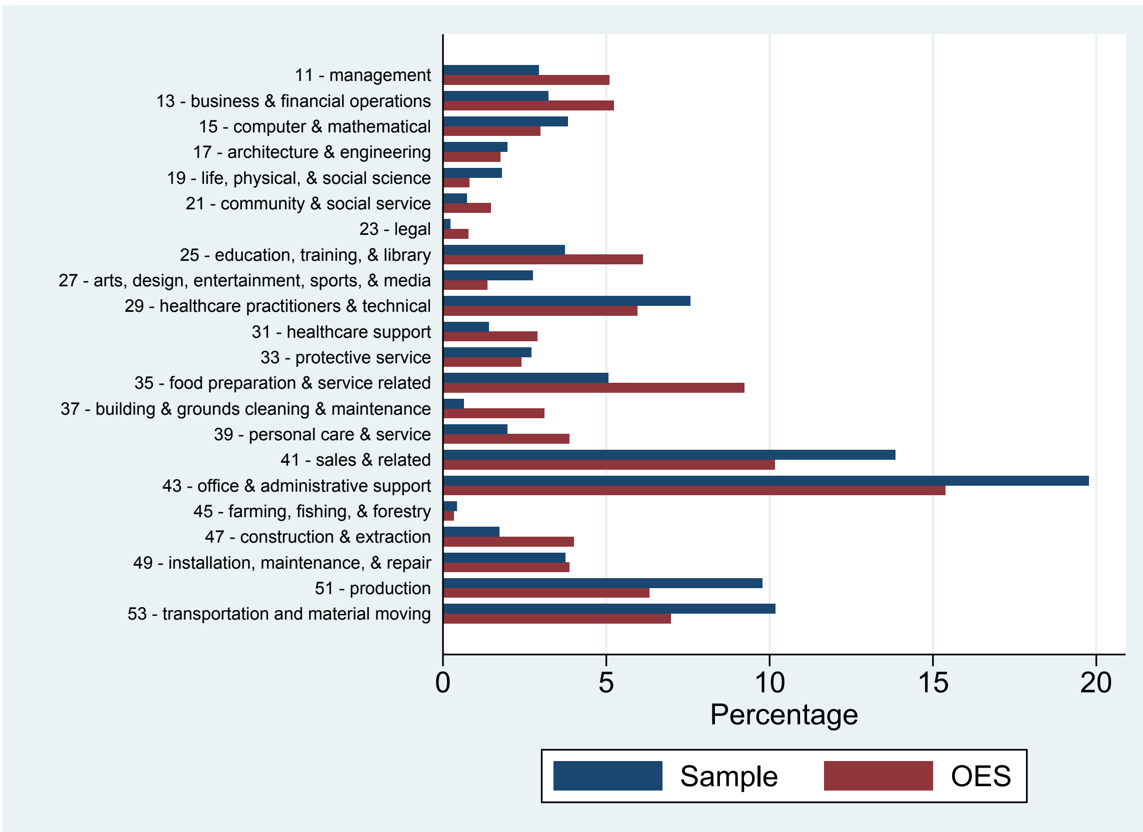


Figure A.1: Occupation Distribution
The figure presents the occupation distribution of the work-level data in the sample and the occupation distribution of the 2017 BLS OES data.

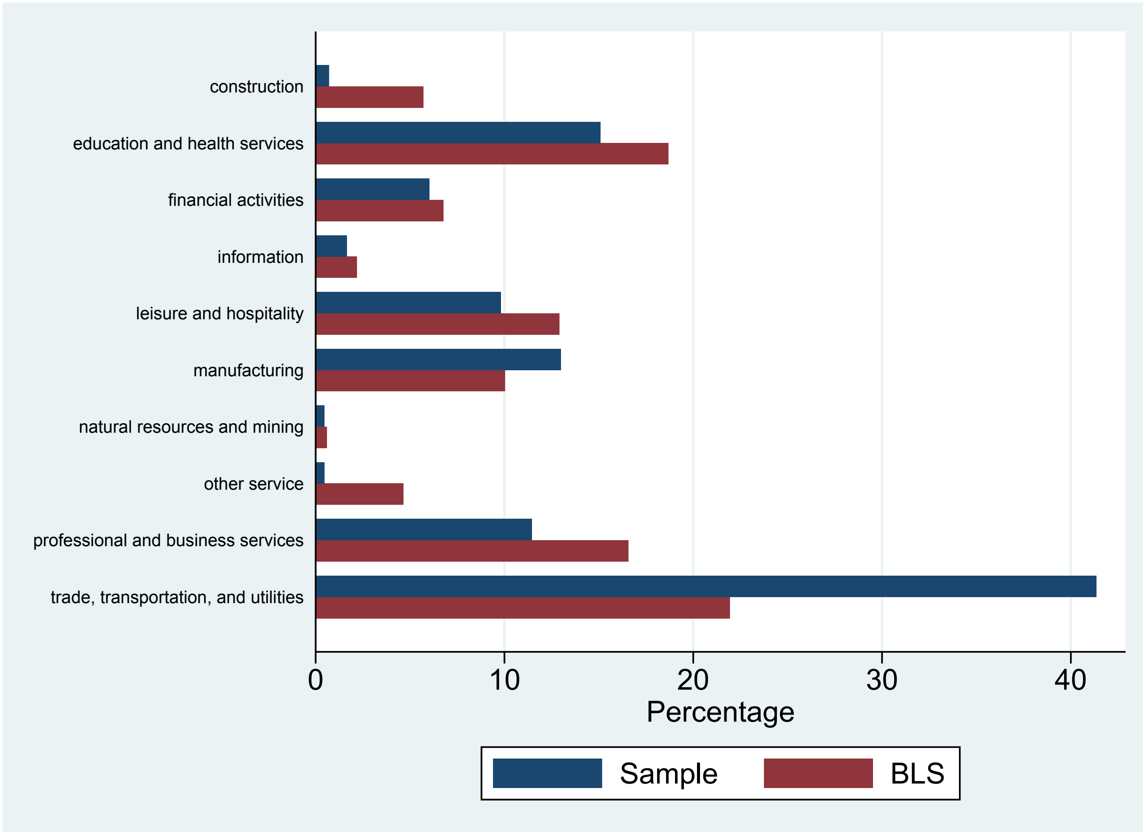
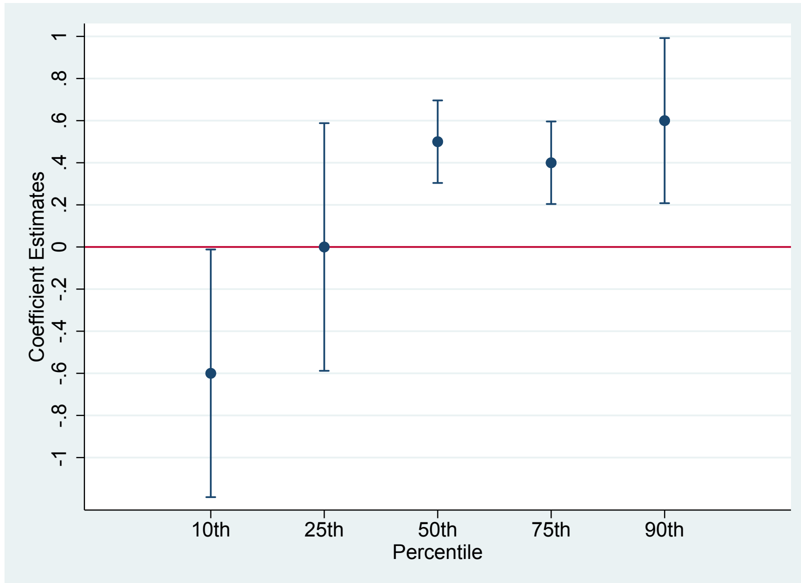


Figure A.2: Industry Distribution
The figure presents the industry distribution of the work-level data in the sample and the industry distribution of the June 2018 BLS data.

Panel A: High-Skilled Workers



Panel B: Low-Skilled Wokrers

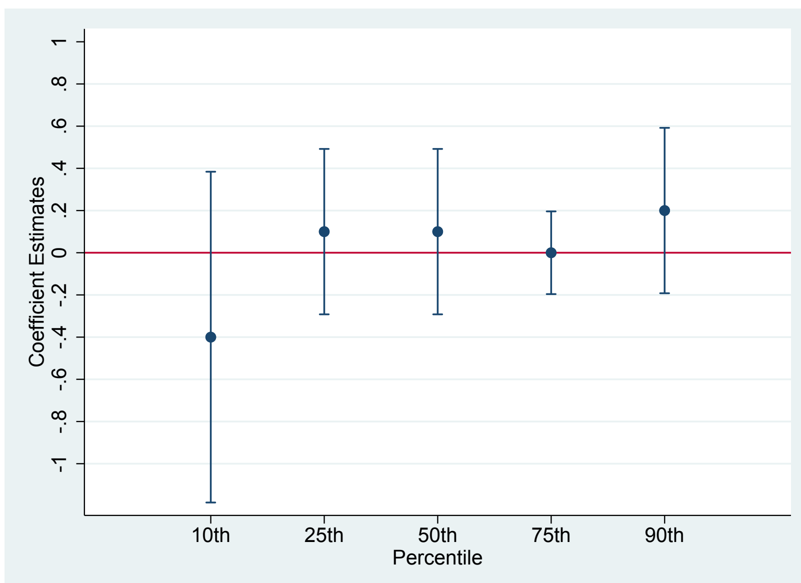


Figure A.3: Distributional Effects of VC Investments on Wages of Incumbent Establishments: Alternative Specification

The figure presents distribution effects of VC investments on wages of incumbent establishments using Equation 1.2 by using log changes in wage percentiles (in percentage points) as dependent variables.

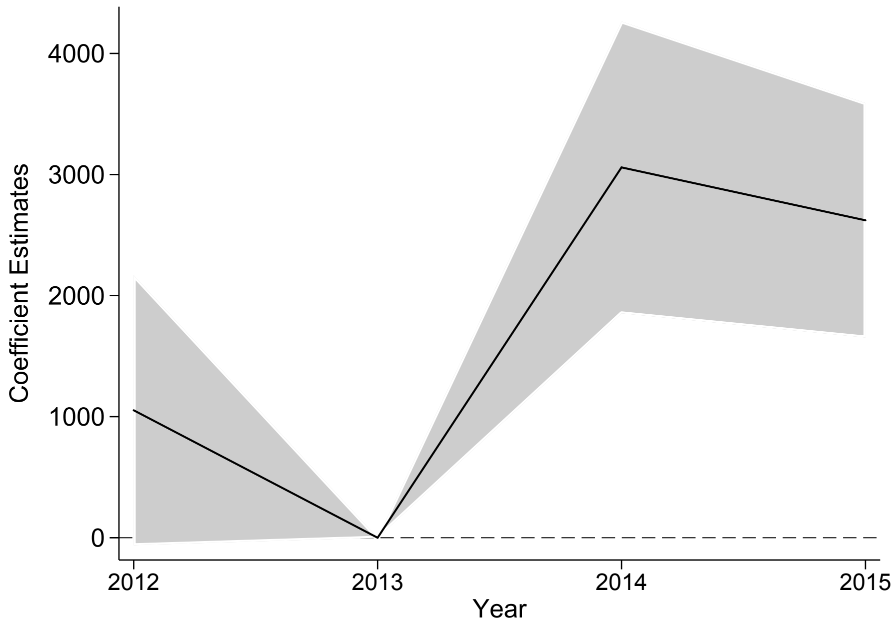


Figure A.4: Dynamic Effect of 2014 VC Boom
The figure presents the dynamic effect on wages of incumbent establishments around the 2014 VC boom.

Table A.1: Effect of VC on the Wage Dispersion of Incumbent Establishments

This table presents results of establishment-level panel regressions assessing the effect of VC investments on the wage dispersion using Equations 1.1, 1.2, and 1.3. The dependent variable is the dollar change in the gap between the 90^{th} percentile of high-skilled workers' wages and the 10^{th} percentile of low-skilled workers' wages. VCPerCapita (in thousands of dollars) is the dollar value of VC investments in a CZ scaled by the population in 2000. The regressions include establishments in CZs that received any VC investments from 2009 to 2017. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Δ(P90_HighSkilled - P10_LowSkilled)				
	(1)	(2)	(3)	(4)	(5)
VCPerCapita	859***	662***	815***	634***	492***
	(143)	(162)	(187)	(164)	(167)
Establishment FE	✓	✓	✓	✓	✓
Year FE	✓				
Industry × Year FE		✓			
State × Year FE			✓		
Industry × State × Year FE				✓	
Firm × Year FE					✓
Observations	337,032	337,032	337,032	337,032	337,032
R-squared	0.112	0.150	0.113	0.217	0.448

Table A.2: IV Estimation

This table presents results of establishment-level panel regressions assessing the effect of VC investments on the wages of high-skilled workers at incumbent establishments using the instrumental variable approach. The dependent variable is the dollar change in the 90^{th} percentile of high-skilled workers' wages. VCPerCapita (in thousands of dollars) is the dollar value of VC investments in a CZ scaled by the population in 2000. LP Return represents the distance-weighted portfolio returns of limited partners. Buyout Fund represents the distance-weighted inflows to buyout funds. The regressions include establishments in CZs that received any VC investments from 2009 to 2017. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	ΔP90_HighSkilled	1st Stage	ΔP90_HighSkilled	1st Stage
	(1)	(2)	(3)	(4)
$VC\widehat{PerC}apita$	3,735**		1,462***	
	(1,824)		(499)	
LP Return		0.019**		
		(0.008)		
Buyout Fund				0.021*
				(0.011)
Observations	337,032	337,032	337,032	337,032
R-squared	0.107		0.107	
F-stat (excl instr.)	5.2		3.6	
Establishment FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓

Table A.3: High Tech vs. Low Tech

This table presents results of establishment-level panel regressions assessing the differential effect of VC investments on wages of high-skilled workers at high-tech and low-tech incumbent establishments using Equations 1.1 and 1.2. The dependent variable is the dollar change in the 90^{th} percentile of high-skilled workers' wages. VCPerCapita (in thousands of dollars) is the dollar value of VC investments in a CZ scaled by the population in 2000. HighTech (LowTech) is a dummy variable indicating the establishment is in high-tech (low-tech) industries. The regressions include establishments in CZs that received any VC investments from 2009 to 2017. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	ΔP90_HighSkilled			
	(1)	(2)	(3)	(4)
VCPerCapita × HighTech	2,202***	1,460***	2,074***	1,067***
	(268)	(251)	(290)	(325)
VCPerCapita × LowTech	693***	650***	560*	743**
	(229)	(227)	(292)	(318)
Observations	337,032	337,032	337,032	337,032
R-squared	0.107	0.151	0.108	0.217
Establishment FE	✓	✓	✓	✓
Year FE	✓			
Industry × Year FE		✓		
State × Year FE			✓	
Industry × State × Year FE				✓

Table A.4: Heterogeneity in VC

This table presents results of establishment-level panel regressions assessing the effect of different VC investments on wages of high-skilled workers at incumbent establishments using Equation 1.2. The dependent variable is the dollar change in the 90^{th} percentile of high-skilled workers' wages. Hightech_VCPerCapita (Lowtech_VCPerCapita) is the dollar value of VC investments in high-tech (low-tech) industries of a CZ scaled by the population in 2000. Early_VCPerCapita (Late_VCPerCapita) is the dollar value of VC investments in early-stage (late-stage) startups of a CZ scaled by the population in 2000. The regressions include establishments in CZs that received any VC investments from 2009 to 2017. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	ΔP90_HighSkilled	
	(1)	(2)
Hightech_VCPerCapita	1,351**	
	(535)	
Lowtech_VCPerCapita	-1,377	
	(1,636)	
Early_VCPerCapita		-3,017*
		(1,597)
Late_VCPerCapita		1,994***
		(596)
Observations	337,032	337,032
R-squared	0.217	0.217
Establishment FE	✓	✓
Industry × State × Year FE	✓	✓

Table A.5: Effect of VC Funding on Startup Wages

This table presents results of event studies assessing the effect of VC funding on startup wages using Equation 1.7. A VC deal is an event. The dependent variable is the logarithm of the average wage of foreign workers being requested by LCA. Post is equal to 1 after the VC deal, and 0 otherwise. Treated is equal to 1 for startups receiving funding in the VC deal, and 0 otherwise. Column (1) includes windows 3 months before and 3 months after the events. Column (2) includes windows 6 months before and 6 months after the events. Standard errors clustered by CZ are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	ln(Wage)	
	[-3,3]	[-6,6]
	(1)	(2)
Post $\times$ Treated	-0.002	0.009**
	(0.014)	(0.004)
Event $\times$ Firm FE	✓	✓
Event $\times$ Month FE	✓	✓
Observations	227,644	451,532
R-squared	0.815	0.755

APPENDIX B

APPENDIX FOR "THE DARK SIDE OF TECHNOLOGICAL PROGRESS? IMPACT OF E-COMMERCE ON EMPLOYEES AT BRICK-AND-MORTAR RETAILERS"

Table B.1: County Demographics

This table presents demographics statistics of 3,135 counties based on Census 2010.

	Full Sample	Counties with FCs	Counties within 50 Miles of FCs	Counties within 100 Miles of FCs
N	3,135	50	445	1,141
Total Population	308,674,608	30,774,770	86,724,715	163,939,679
Population	98,460.80	615,495.40	194,887.00	143,680.70
Population Density	259.49	672.84	698.24	531.76
Retail Sales (in millions)	431.71	2,827.55	821.65	615.01
Retail Sales per Capita	3,552.72	4,623.48	3,692.95	3,679.81
Median Household Income	43,419.43	56,220.34	51,179.56	47,123.08
Unemployment Rate	9.36	9.64	10.17	10.27
Percent Age under 18	23.49	24.80	23.77	23.22
Percent Age over 65	15.93	12.16	14.08	15.31
Percent High School Graduate or Higher	82.51	85.84	83.04	82.31
Percent Bachelor's Degree or Higher	18.73	27.52	21.74	19.73

Table B.2: List of Retail Sectors
This table presents 6-digit NAICS industries that we include in our analysis.

NAICS	Industry Name
441310	Automotive parts and accessories stores
441320	Tire dealers
442110	Furniture stores
442210	Floor covering stores
442291	Window treatment stores
442299	All other home furnishings stores
443141	Household appliance stores
443142	Electronics stores
444110	Home centers
444120	Paint and wallpaper stores
444130	Hardware stores
444190	Other building material dealers
444210	Outdoor power equipment stores
444220	Nursery, garden, and farm supply stores
446120	Cosmetic and beauty supply stores
446191	Food, health, supplement stores
446199	All other health and personal care stores
448110	Men’s clothing stores
448120	Women’s clothing stores
448130	Children’s and infants’ clothing stores
448140	Family clothing stores
448150	Clothing accessories stores
448190	Other clothing stores
448210	Shoe stores
448310	Jewelry stores
448320	Luggage and leather goods stores
451110	Sporting goods stores
451120	Hobby, toy, and game stores
451130	Sewing, needlework, and piece goods stores
451140	Musical instrument and supplies stores
451211	Book stores
451212	News dealers and newsstands
451220	Precorded tape, cd, and record stores
452111	Department stores, except discount
452112	Discount department stores
452910	Warehouse clubs and supercenters
452990	All other general merchandise stores
453110	Florists
453210	Office supplies and stationery stores
453220	Gift, novelty, and souvenir stores
453310	Used merchandise stores
453910	Pet and pet supplies stores
453920	Art dealers
453930	Manufactured, mobile, home dealers
453991	Tobacco stores
453998	Store retailers not specified elsewhere

APPENDIX C

APPENDIX FOR “THE SPEED OF INFORMATION AND THE SELL-SIDE RESEARCH INDUSTRY”

Table C.1: Timeline of Barclays Capital, Merrill Lynch, and Morgan Stanley v. *Theflyonthewall.com*
This table shows key events in the lawsuit of Barclays Capital, Merrill Lynch, and Morgan Stanley against *Theflyonthewall.com* (Docket no. 10-1372-cv). The key dates where the courts decisions were rendered are highlighted in bold. On March 18, 2010, the district court ruled in favor of the plaintiffs claim that *FLY* misappropriated hot news. On June 20, 2011, the appeals court reversed this decision.

Date	Event
June 26,2006	Barclays Capital, Merrill Lynch, and Morgan Stanley file suit against Theflyonthewall.com alleging copyright infringement and “hot news” misappropriation.
March 18th, 2010	Based on the plaintiffs “hot news” misappropriation claim, the district court prohibits Theflyonthewall.com from disclosing recommendations until 30 minutes after the opening of the NYSE or 10:00AM, whichever is later. Plaintiffs were awarded statutory damages and attorney fees.
April 9th, 2010	Theflyonewall.com files an appeal.
May 7th, 2010	District court rejects Theflyonthewall.com’s motion to stay or modify the injunction pending appeal.
May 19th, 2010	The United States Court of Appeals for the Second Circuit granted the Theflyonthewall.com’s motion to stay the injunction.
June 20th, 2011	The judgment of the lower court is reversed by the United States Court of Appeals Second Circuit.

Table C.2: Variable Definitions
This table contains definitions for the key variables used in the empirical analysis.

Variable	Definition
SIZE	Log firm market capitalization at time t-1 relative to the recommendation date.
TOBIN_Q	Firms market capitalization at scaled by replacement value at the fiscal year end prior to the recommendation.
RUNUP	The absolute value of stock price performance over days (-20,-1) relative to the recommendation date.
UPGRADE (DOWNGRADE)	A dummy variable equal to one if the recommendation is revised upward (downward), zero otherwise.
REVISION_MAGNITUDE	The absolute value of the difference between the last numerical recommendation rating and the current recommendation.
ALL_STAR	A dummy variable equal to one if the analyst is an Institutional Investor all-star analyst in year t, zero otherwise.
PLAINTIFF	A dummy variable equal to one if the broker sued Theflyonthewall.com (Barclays, Morgan Stanley or Bank of America), zero otherwise.
BROKER_SIZE	The natural logarithm of the number of analysts employed by the broker.
NEGLECTED	An indicator that takes the value of one if two or fewer analysts are covering the stock, zero otherwise.
INSTITUTIONAL_OWNERSHIP	The natural log of one plus the percentage of shares held by institutional investors at the quarter end prior to the recommendation.
VWAP	Volume weighted price calculated from TAQ using trades between 9:30 and 4:00PM.
MEDIAN_INDUSTRY_RETURN	The median return in industry i at time t.
#_OF_IPOS	The number of IPOs filed in month t.
#_OF_DELISTINGS	The number of firms that delist in month t.
Δ_TRADING_VOLUME	The change in industry trading volume from month t-1 to t.

REFERENCES

[1] D. Acemoglu, "Technical change, inequality, and the labor market," *Journal of economic literature*, vol. 40, no. 1, pp. 7–72, 2002.

[2] A. Agrawal and P. Tambe, "Private equity and workers' career paths: The role of technological change," *Review of Financial Studies*, vol. 29, no. 9, pp. 2455–2489, 2016.

[3] F. Akbas, S. Markov, M. Subasi, and E. Weisbrod, "Determinants and consequences of information processing delay: Evidence from the thomson reuters institutional brokers estimate system," *Journal of Financial Economics*, vol. 127, no. 2, pp. 366–388, 2018.

[4] N. Amornsiripanitch, P. A. Gompers, and Y. Xuan, "More than money: Venture capitalists on boards," *Working Paper*, 2017.

[5] R. Ashraf and R. Ray, "Human capital, skilled immigrants, and innovation," *Working Paper*, 2017.

[6] D. H. Autor, "Skills, education, and the rise of earnings inequality among the "other 99 percent"," *Science*, vol. 344, no. 6186, pp. 843–851, 2014.

[7] ——, "Why are there still so many jobs? the history and future of workplace automation," *Journal of economic perspectives*, vol. 29, no. 3, pp. 3–30, 2015.

[8] D. H. Autor, D. Dorn, and G. H. Hanson, "The china syndrome: Local labor market effects of import competition in the united states," *American Economic Review*, vol. 103, no. 6, pp. 2121–68, 2013.

[9] ——, "Untangling trade and technology: Evidence from local labour markets," *The Economic Journal*, vol. 125, no. 584, pp. 621–646, 2015.

[10] D. H. Autor, D. Dorn, G. H. Hanson, and J. Song, "Trade adjustment: Worker-level evidence," *The Quarterly Journal of Economics*, vol. 129, no. 4, pp. 1799–1860, 2014.

[11] D. H. Autor, L. F. Katz, and A. B. Krueger, "Computing inequality: Have computers changed the labor market?" *The Quarterly journal of economics*, vol. 113, no. 4, pp. 1169–1213, 1998.

[12] D. H. Autor, F. Levy, and R. J. Murnane, “The skill content of recent technological change: An empirical exploration,” *The Quarterly journal of economics*, vol. 118, no. 4, pp. 1279–1333, 2003.

[13] J. Azar, I. Marinescu, and M. Steinbaum, “Labor market concentration,” *NBER Working Paper*, 2017.

[14] T. Babina, P. Ouimet, and R. Zarutskie, “Going entrepreneurial? ipos and new firm creation,” *Working Paper*, 2017.

[15] J. Bai, T. Philippon, and A. Savov, “Have financial markets become more informative?” *Journal of Financial Economics*, vol. 122, no. 3, pp. 625–654, 2016.

[16] T. J. Bartik, “Who benefits from state and local economic development policies?” *W.E. Upjohn Institute of Employment Research*, 1991.

[17] E. Basker, “Job creation or destruction? labor market effects of wal-mart expansion,” *Review of Economics and Statistics*, vol. 87, no. 1, pp. 174–183, 2005.

[18] B. Baugh, I. Ben-David, and H. Park, “Can taxes shape an industry? evidence from the implementation of the amazon taxi,” *The Journal of Finance*, vol. 73, no. 4, pp. 1819–1855, 2018.

[19] E. Benmelech, N. K. Bergman, and H. Kim, “Strong employers and weak employees: How does employer concentration affect wages?” *NBER Working Paper*, 2018.

[20] R. Bhushan, “Firm characteristics and analyst following,” *Journal of accounting and economics*, vol. 11, no. 2-3, pp. 255–274, 1989.

[21] N. Bishara, “Fifty ways to leave your employer: Relative enforcement of covenants not to compete, trends, and implications for employee mobility policy,” *University of Pennsylvania Journal of Business Law*, vol. 13, no. 3, pp. 751–795, 2011.

[22] A. Borisov, A. Ellul, and M. Sevilir, “Access to public capital markets and employment growth,” *Working Paper*, 2017.

[23] D. Bradley, J. Clarke, S. Lee, and C. Ornthanalai, “Are analysts recommendations informative? intraday evidence on the impact of time stamp delays,” *The Journal of Finance*, vol. 69, no. 2, pp. 645–673, 2014.

[24] J. Brogaard, T. Hendershott, and R. Riordan, “High-frequency trading and price discovery,” *The Review of Financial Studies*, vol. 27, no. 8, pp. 2267–2306, 2014.

[25] E. Brynjolfsson, Y. Hu, and M. D. Smith, "Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers," *Management Science*, vol. 49, no. 11, pp. 1580–1596, 2003.

[26] E. Brynjolfsson and A. McAfee, *Race Against the Machine*. Digital Frontier Press, 2011.

[27] ——, *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company, 2014.

[28] E. Brynjolfsson and M. D. Smith, "Frictionless commerce? a comparison of internet and conventional retailers," *Management science*, vol. 46, no. 4, pp. 563–585, 2000.

[29] A. W. Butler, L. Fauver, and I. Spyridopoulos, "Local economic consequences of stock market listings," *Working Paper*, 2017.

[30] S. Chava, V. Nanda, and S. C. Xiao, "Lending to innovative firms," *Review of Corporate Finance Studies*, vol. 6, no. 2, pp. 234–289, 2017.

[31] S. Chava, A. Oettl, M. Singh, and L. Zeng, "Impact of e-commerce on employees at brick-and-mortar retailers," *Working Paper*, 2018.

[32] T. J. Chemmanur, K. Krishnan, and D. K. Nandy, "How does venture capital financing improve efficiency of private firms? a look beneath the surface," *Review of Financial Studies*, vol. 24, no. 12, pp. 4037–4090, 2011.

[33] Y. Chen, B. Kelly, and W. Wu, "Sophisticated investors and market efficiency: Evidence from a natural experiment," National Bureau of Economic Research, Tech. Rep., 2018.

[34] S. Chernenko, J. Lerner, and Y. Zeng, "Mutual funds as venture capitalists? evidence from unicorns," *Working Paper*, 2018.

[35] J. Clarke, A. Khorana, A. Patel, and P. R. Rau, "The impact of all-star analyst job changes on their coverage choices and investment banking deal flow," *Journal of Financial Economics*, vol. 84, no. 3, pp. 713–737, 2007.

[36] M. B. Clement, "Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter?" *Journal of Accounting and Economics*, vol. 27, no. 3, pp. 285–303, 1999.

[37] J. Cornaggia, M. Gustafson, J. D. Kotter, and K. Pisciotta, "Public ownership and the local economy," *Working Paper*, 2018.

[38] S. S. Crawford, D. T. Roulstone, and E. C. So, "Analyst initiations of coverage and stock return synchronicity," *The Accounting Review*, vol. 87, no. 5, pp. 1527–1553, 2012.

[39] M. Da Rin, T. Hellmann, and M. Puri, "A survey of venture capital research," *Handbook of the Economics of Finance*, vol. 2, pp. 573–648, 2013.

[40] A. Davila, G. Foster, and M. Gupta, "Venture capital financing and the growth of startup firms," *Journal of Business Venturing*, vol. 18, no. 6, pp. 689–708, 2003.

[41] D. W. Diamond and R. E. Verrecchia, "Information aggregation in a noisy rational expectations economy," *Journal of financial economics*, vol. 9, no. 3, pp. 221–235, 1981.

[42] C. Dougal, C. A. Parsons, and S. Titman, "Urban vibrancy and firm value creation," *Working Paper*, 2018.

[43] J. Dugast and T. Foucault, "Data abundance and asset price informativeness," *Journal of Financial Economics*, vol. 130, no. 2, pp. 367–391, 2018.

[44] M. Farboodi, A. Matray, and L. Veldkamp, "Where has all the big data gone?," 2018.

[45] M. Farboodi and L. Veldkamp, "Long run growth of financial technology," National Bureau of Economic Research, Tech. Rep., 2017.

[46] K. J. Forbes, "Skill classification does matter: Estimating the relationship between trade flows and wage inequality," *The Journal of International Trade & Economic Development*, vol. 10, no. 2, pp. 175–209, 2001.

[47] T. Foucault, J. Hombert, and I. Roşu, "News trading and speed," *The Journal of Finance*, vol. 71, no. 1, pp. 335–382, 2016.

[48] K. Froot, N. Kang, G. Ozik, and R. Sadka, "What do measures of real-time corporate sales say about earnings surprises and post-announcement returns?" *Journal of Financial Economics*, vol. 125, no. 1, pp. 143–162, 2017.

[49] M. J. Garmaise, "Ties that truly bind: Noncompetition agreements, executive compensation, and firm investment," *Journal of Law, Economics, and Organization*, vol. 27, no. 2, pp. 376–425, 2011.

[50] A. Ghose, M. D. Smith, and R. Telang, "Internet exchanges for used books: An empirical analysis of product cannibalization and welfare impact," *Information systems research*, vol. 17, no. 1, pp. 3–19, 2006.

[51] X. Giroud and H. M. Mueller, "Firms' internal networks and local economic shocks," *American Economic Review*, forthcoming.

[52] M. A. Goldstein, P. Irvine, E. Kandel, and Z. Wiener, "Brokerage commissions and institutional trading patterns," *The Review of Financial Studies*, vol. 22, no. 12, pp. 5175–5212, 2009.

[53] P. Gompers and J. Lerner, "Money chasing deals? the impact of fund inflows on private equity valuations," *Journal of Financial Economics*, vol. 55, no. 2, pp. 281–325, 2000.

[54] J. Gonzalez-Uribe, "Exchanges in venture capital portfolios," *Working Paper*, 2017.

[55] T. C. Green, R. Jame, S. Markov, and M. Subasi, "Broker-hosted investor conferences," *Journal of Accounting and Economics*, vol. 58, no. 1, pp. 142–166, 2014.

[56] J. Grennan and R. Michaely, "Fintechs and the market for financial analysis," *Michael J. Brennan Irish Finance Working Paper Series Research Paper*, no. 18-11, pp. 19–10, 2018.

[57] S. J. Grossman and J. E. Stiglitz, "On the impossibility of informationally efficient markets," *The American economic review*, vol. 70, no. 3, pp. 393–408, 1980.

[58] B. Groysberg, P. M. Healy, and D. A. Maber, "What drives sell-side analyst compensation at high-status investment banks?" *Journal of Accounting Research*, vol. 49, no. 4, pp. 969–1000, 2011.

[59] T. Hellmann and M. Puri, "The interaction between product market and financing strategy: The role of venture capital," *Review of Financial Studies*, vol. 13, no. 4, pp. 959–984, 2000.

[60] ——, "Venture capital and the professionalization of start-up firms: Empirical evidence," *Journal of Finance*, vol. 57, no. 1, pp. 169–197, 2002.

[61] D. Hirshleifer, S. S. Lim, and S. H. Teoh, "Driven to distraction: Extraneous events and underreaction to earnings news," *The Journal of Finance*, vol. 64, no. 5, pp. 2289–2325, 2009.

[62] T. J. Holmes, "The diffusion of wal-mart and economies of density," *Econometrica*, vol. 79, no. 1, pp. 253–302, 2011.

[63] H. Hong and J. C. Stein, "A unified theory of underreaction, momentum trading, and overreaction in asset markets," *The Journal of finance*, vol. 54, no. 6, pp. 2143–2184, 1999.

[64] J.-F. Houde, P. Newberry, and K. Seim, "Economies of density in e-commerce: A study of amazons fulfillment center network," National Bureau of Economic Research, Tech. Rep., 2017.

[65] G. Hu, "Measures of implicit trading costs and buy–sell asymmetry," *Journal of Financial Markets*, vol. 12, no. 3, pp. 418–437, 2009.

[66] G. Hu, R. D. McLean, J. Pontiff, and Q. Wang, "The year-end trading activities of institutional investors: Evidence from daily trades," *The Review of Financial Studies*, vol. 27, no. 5, pp. 1593–1614, 2014.

[67] P. J. Irvine, "Analysts' forecasts and brokerage-firm trading," *The Accounting Review*, vol. 79, no. 1, pp. 125–149, 2004.

[68] P. J. Irvine, "Do analysts generate trade for their firms? evidence from the toronto stock exchange," *Journal of Accounting and Economics*, vol. 30, no. 2, pp. 209–226, 2000.

[69] A. R. Jackson, "Trade generation, reputation, and sell-side analysts," *The Journal of Finance*, vol. 60, no. 2, pp. 673–717, 2005.

[70] R. Jame, R. Johnston, S. Markov, and M. C. Wolfe, "The value of crowdsourced earnings forecasts," *Journal of Accounting Research*, vol. 54, no. 4, pp. 1077–1110, 2016.

[71] J. S. Jeffers, "The impact of restricting labor mobility on corporate investment and entrepreneurship," *Working Paper*, 2018.

[72] P. Jia, "What happens when wal-mart comes to town: An empirical analysis of the discount retailing industry," *Econometrica*, vol. 76, no. 6, pp. 1263–1316, 2008.

[73] L. F. Katz and K. M. Murphy, "Changes in relative wages, 1963-1987: Supply and demand factors," *Quarterly Journal of Economics*, vol. 107, no. 1, pp. 35–78, 1992.

[74] S. Klasa, H. Ortiz-Molina, M. Serfling, and S. Srinivasan, "Protection of trade secrets and capital structure decisions," *Journal of Financial Economics*, vol. 128, no. 2, pp. 266–286, 2018.

[75] L. Kogan, D. Papanikolaou, L. Schmidt, and J. Song, "Technological innovation and the distribution of labor income growth," *Working Paper*, 2018.

[76] S. Kortum and J. Lerner, "Assessing the contribution of venture capital to innovation," *RAND Journal of Economics*, vol. 31, no. 4, pp. 674–692, 2000.

[77] A. B. Krueger, “How computers have changed the wage structure: Evidence from microdata, 1984–1989,” *The Quarterly Journal of Economics*, vol. 108, no. 1, pp. 33–60, 1993.

[78] S. Lagaras, “Corporate takeovers and labor restructuring,” *Working Paper*, 2017.

[79] M. H. Lang, K. V. Lins, and D. P. Miller, “Adrs, analysts, and accuracy: Does cross listing in the united states improve a firm’s information environment and increase market value?” *Journal of Accounting Research*, vol. 41, no. 2, pp. 317–345, 2003.

[80] E. X. Li, K Ramesh, M. Shen, and J. S. Wu, “Do analyst stock recommendations piggyback on recent corporate news? an analysis of regular-hour and after-hours revisions,” *Journal of Accounting Research*, vol. 53, no. 4, pp. 821–861, 2015.

[81] R. K. Loh and R. M. Stulz, “When are analyst recommendation changes influential?” *The review of financial studies*, vol. 24, no. 2, pp. 593–627, 2010.

[82] D. Lui, S. Markov, and A. Tamayo, “Equity analysts and the market’s assessment of risk,” *Journal of Accounting Research*, vol. 50, no. 5, pp. 1287–1317, 2012.

[83] W. Ma, P. Ouimet, and E. Simintzi, “Mergers and acquisitions, technological change and inequality,” *Working Paper*, 2018.

[84] S. Markov, V. Muslu, and M. Subasi, “Analyst tipping: Additional evidence,” *Journal of Business Finance & Accounting*, vol. 44, no. 1-2, pp. 94–115, 2017.

[85] M. W. Marr Jr, “Effects of the antitakeover provisions of pennsylvania act 36: A survey of empirical studies,” *Financial Analysts Journal*, vol. 48, no. 6, pp. 52–57, 1992.

[86] D. A. Matsa, “Capital structure and a firm’s workforce,” *Annual Review of Financial Economics*, vol. 10, pp. 387–412, 2018.

[87] K. Merkley, R. Michaely, and J. Pacelli, “Does the scope of the sell-side analyst industry matter? an examination of bias, accuracy, and information content of analyst reports,” *The Journal of Finance*, vol. 72, no. 3, pp. 1285–1334, 2017.

[88] A. Mian and A. Sufi, “What explains the 2007-2009 drop in employment?” *Econometrica*, vol. 82, no. 6, pp. 2197–2223, 2014.

[89] E. Moretti, “Workers’ education, spillovers, and productivity: Evidence from plant-level production functions,” *American Economic Review*, vol. 94, no. 3, pp. 656–690, 2004.

[90] ——, "Local labor markets," *Handbook of Labor Economics*, vol. 4b, pp. 1237–1313, 2010.

[91] A. Mukherjee, M. Singh, and A. Zaldokas, "Do corporate taxes hinder innovation," *Journal of Financial Economics*, vol. 124, no. 1, pp. 195–221, 2017.

[92] National Academies of Sciences, Engineering, and Medicine, *Information Technology and the US Workforce: Where Are We and Where Do We Go from Here?* National Academies Press, 2017.

[93] D. Neumark, J. Zhang, and S. Ciccarella, "The effects of wal-mart on local labor markets," *Journal of Urban Economics*, vol. 63, no. 2, pp. 405–430, 2008.

[94] A. Pozzi, "The effect of internet distribution on brick-and-mortar sales," *The RAND Journal of Economics*, vol. 44, no. 3, pp. 569–583, 2013.

[95] M. Puri and R. Zarutskie, "On the life cycle dynamics of venture-capital- and non-venture-capital-financed firms," *Journal of Finance*, vol. 67, no. 6, pp. 2247–2293, 2012.

[96] S. Samila and O. Sorenson, "Venture capital, entrepreneurship, and economic growth," *Review of Economics and Statistics*, vol. 93, no. 1, pp. 338–349, 2011.

[97] M. Schnitzer and M. Watzinger, "Measuring the spillovers of venture capital," *Working Paper*, 2017.

[98] M. Shen, "Skilled labor mobility and firm value: Evidence from a natural experiment," *Working Paper*, 2018.

[99] R. C. Silva, "Internal labor markets, wage convergence and investment," *Working Paper*, 2017.

[100] S. E. Stickel, "Reputation and performance among security analysts," *The Journal of Finance*, vol. 47, no. 5, pp. 1811–1836, 1992.

[101] B. Stone, *The everything store: Jeff Bezos and the age of Amazon*. Random House, 2013.

[102] M. Subramani and E. Walden, "The impact of e-commerce announcements on the market value of firms," *Information Systems Research*, vol. 12, no. 2, pp. 135–154, 2001.

[103] C. M. Tolbert and M. Sizer, "Us commuting zones and labor market areas: A 1990 update," *Rural Economy Division, Economic Research Service, U.S. Department of Agriculture Staff Paper*, no. 9614, 1996.

[104] R. E. Verrecchia, “Information acquisition in a noisy rational expectations economy,” *Econometrica: Journal of the Econometric Society*, pp. 1415–1430, 1982.

[105] S.-J. Xu, “Skilled labor supply and corporate investment: Evidence from the h-1b visa program,” *Working Paper*, 2018.

[106] D. Yagan, “Employment hysteresis from the great recession,” *Journal of Political Economy*, forthcoming.

[107] C. Zhu, “Big data as a governance mechanism,” *The Review of Financial Studies*, vol. 32, no. 5, pp. 2021–2061, 2019.

[108] R. H. Ziedonis, “Don’t fence me in: Fragmented markets for technology and the patent acquisition strategies of firms,” *Management Science*, vol. 50, no. 6, pp. 804–820, 2004.

[109] L. Zingales, “In search of new foundations,” *Journal of Finance*, vol. 55, no. 4, pp. 1623–1653, 2002.

VITA

Linghang Zeng completed his Ph.D. in Finance at the Scheller College of Business, Georgia Institute of Technology. Before that, Linghang received a Master's degree in Mathematical Finance from the University of Alberta and a Bachelor's degree in Industrial Engineering from Nanjing University.

Linghang's research interests include labor and finance, household finance, and empirical asset pricing. His research has won Best Paper Awards from the Northern Finance Association and Chicago Quantitative Alliance. His papers have been presented at numerous prestigious conferences such as NBER Summer Institute, SFS Cavalcade, and European Finance Association, and accepted for publication at the Journal of Financial and Quantitative Analysis. He teaches Finance and Investments.