ESSAYS IN ECONOMICS AND CORPORATE FINANCE

Tao Li

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2013

ABSTRACT

ESSAYS IN ECONOMICS AND CORPORATE FINANCE

Tao Li

This dissertation consists of two distinct chapters. In the first chapter, I study the outsourcing of corporate governance to proxy advisory firms, which are third-party advisors that help institutional investors decide which way to vote on corporate governance issues. Advising equity assets in trillions of dollars, these advisors play a powerful role in shaping corporate governance. First, I model how conflicts of interest arise when a proxy advisor provides advisory services to investors as well as consulting services to corporations on the same governance issues. The advisor can issue biased voting recommendations when expected reputation costs are low, compared to consulting fees. I then study how increased competition can alleviate these conflicts. Using a unique dataset on voting recommendations, I show that the entry of a new advisory firm reduces favorable recommendations for management proposals by the incumbent advisor. This is consistent with our theory as the incumbent is subject to conflicts of interest by serving both investors and corporations. These results inform the policy debate on whether and how to regulate the proxy advisory industry.

The second chapter of the thesis assesses the value of access to public transportation in Beijing, a megacity suffering from severe traffic congestion. Existing urban economic theory states that traffic congestion is welfare reducing. In practice, policymakers in congested cities invest heavily in public transit systems to reduce transportation costs. However, not all public transit modes are created equal – those that help alleviate traffic congestion are the most desirable. Using a unique panel dataset of Beijing's residential properties on sale between 2003 and 2005, I find strong evidence that traffic delays translate into lower housing prices, confirming that congestion is costly. Moreover, I show that announcements of metro

line construction inflate prices of properties near future stations, and the increase is even more staggering for more congested areas. This suggests that metro lines are expected to reduce adverse impacts of congestion. However, additional bus routes are not capitalized into prices because buses move slowly in the gridlocked city, often exacerbating rather than alleviating congestion. These findings suggest that the overall quantity of public transit services does not necessarily increase welfare.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am extremely grateful to my advisor Patrick Bolton for his guidance and support on my dissertation. Over the past four years, he has guided me through each step of my development with encouragement, enthusiasm and great insights. He serves as a constant inspiration as I embark on the next phase of my career.

I would also like to thank Wei Jiang, Suresh Naidu, Andrea Prat and Eric Verhoogen for their advice. Eric and Wei have been very helpful with the empirical part of this thesis, improving the quality of the work substantially. Andrea and Suresh brought me up to speed on urban economic issues.

This dissertation benefited from helpful comments and discussions from Charles Calomiris, Pierre-André Chiappori, John C. Coffee, Andrew Ellul, Fabrizio Ferri, Paolo Fulghieri, Jeffrey Gordon, Marina Halac, Jonas Hjort, Kate Ho, Pierre Liang, Marcos Nakaguma, Brendan O'Flaherty, Ailsa Röell, Bernard Salanié, Chester Spatt, Paul Tetlock, and David Yermack.

Special thanks to Susan Elmes for inspiring me to study economics while I was working as an architect. I would also like to thank my colleagues and friends for making my time at Columbia a memorable experience. Among others, I thank Adonis Antoniades, Li An, Ritam Chaurey, Guojun Chen, Jaehyun Cho, Alejo Czerwonko, Jonathan Dingel, Jeong Hwan Lee, Andres Liberman, David Monroe, Maya Rossin-Slater, Fangzhou (Ark) Shi, Assaf Shtauber, Xing Xia, and Zhanna Zhanabekova.

I have the greatest gratitude to my family and friends. In particular, I thank my parents for their unconditional love and support. I also thank Yu Zheng for intriguing conversations.

The first chapter of the dissertation was supported by a generous grant from Columbia Business School. The views expressed in that chapter are those of the author and do not necessarily represent those of Columbia Business School.

Dedication

To my parents.

Chapter 1

Outsourcing Corporate Governance: Conflicts of Interest and Competition in the Proxy Advisory Industry

1.1 Introduction

A proxy advisor is a third-party advisory firm that helps institutional shareholders decide which way to vote on corporate issues. For a fee,[1] this advisor provides independent proxy research[2] and voting recommendations. Some advisors also provide voting platforms for their institutional clients. The proxy advisory industry has grown exponentially over the past 25 years due to a convergence of market and regulatory developments, and a steady

[1] Most institutional investors subscribe to proxy advisory services on an annual basis.

[2] Sometimes referred to as proxy analysis, proxy research involves analyzing proxy statements and financial statements of public companies, as well as multiple external original research sources to evaluate board effectiveness and corporate governance risk profiles. It allows institutional investors to understand governance risk within portfolio companies and take appropriate voting action.

growth in institutional equity ownership[3] has substantially increased these investors' voting power as well as their obligations. However, they often lack proper incentives[4] or necessary expertise to do research in order to make informed votes. Many institutional investors rely on the advice of proxy advisory firms. Increased shareholder activism after the dot-com bust further increased demand for governance advice and proxy voting.

In 2003, the SEC adopted rules that require mutual funds to publicly disclose their voting records, as well as adopt policies to ensure that they vote proxies in the best interests of clients, which triggered a sharp increase in demand for proxy advisory services. Recent changes in financial regulation have further accentuated impacts of proxy votes and influences of the proxy advisory industry. These changes include (1) a shift towards majority voting for directors, (2) the Dodd-Frank Act's requirements on advisory vote on executive compensation ("say-on-pay" vote), (3) the elimination of broker discretionary voting[5] in uncontested director elections, and on executive compensation matters.

While the importance of a viable proxy advisory industry is clear, the provision of accurate proxy research and recommendations is complicated by peculiar industry features. First, this industry is dominated by only two firms – Institutional Shareholder Services, Inc. (ISS) and Glass, Lewis & Co. (Glass Lewis) – with Glass Lewis gaining prominence only in the past several years. Second, and perhaps most importantly, proxy advisors' business mod-

[3]In the U.S., institutional investors have become the dominant players in the stock market, owning 50.6% of total equity outstanding, and 73% of the largest 1,000 US companies at the end of 2009 (Tonello and Rabimov, 2010). These investors include mutual funds, public and private pension funds, hedge funds and other fiduciaries. Retail investors usually do not bother to vote for company policies.

[4]An institutional investor usually holds a diversified portfolio. Its stake in a particular company is typically small, and how it votes is unlikely to affect vote outcomes. As one investor put it, "[researching proxy voting issues] does not add a lot of value in terms of making [clients] money..." (see Bew and Fields, 2012).

[5]For most routine proposals, brokers once were allowed to vote on behalf of their retail investors in "street name" (broker discretionary votes or broker non-votes). Research finds that brokers historically voted uninstructed shares in accordance with management's recommendation (see Brickley, Lease and Smith, 1994; Bethel and Gillan, 2002). The SEC has recently prohibited broker discretionary voting as required by the Dodd-Frank Act.

els suffer from conflicts of interest. For example, ISS provides services to both institutional investors and corporate issuers on the same governance issues, potentially creating ample opportunities for making biased voting recommendations. Third, the extent of inaccuracies and lack of transparency in proxy analyses coupled with limited engagement with issuers, have raised concerns among industry experts and securities regulators.

These features of the industry have raised questions about the quality of proxy research and recommendations provided by these prominent players. In particular, a consensus has developed among policymakers and academics as to potential benefits of increasing competition among proxy advisors as a tool for improving quality. For example, SEC's July 2010 concept release on the U.S. proxy system explicitly asked "whether these issues are affected – and if so, how – by the fact that there is one dominant proxy advisory firm in the marketplace, Institutional Shareholder Services ('ISS'), whose long-standing position, according to the Government Accountability Office, 'has been cited by industry analysts as a barrier to competition."'

Although the argument for increased competition has gained traction, its theoretical and empirical merits are not at all well established. Recent research on the credit ratings industry,[6] a similar information intermediary market,[7] sheds light on why the role of competition needs to be better understood. A growing body of literature (see Bolton, Freixas and Shapiro, 2012; Becker and Milbourn, 2011) has shown that competition among credit rat-

[6]There are a number of parallels between the credit ratings industry and the proxy advisory industry. These common features include (1) various conflicts of interest (see footnote 9 for details), (2) a lack of competition due to barrier of entry, (3) a lack of transparency in decision making, and potential inaccuracies. However, their fundamental business models are different. Credit ratings agencies use an issuer-pay model – the bulk of their ratings-related income comes from issuers whose products they rate (see White, 2002). Proxy advisory firms, on the other hand, adopt a buyer-pay model that their principal revenue stream comes from investors who purchase their services.

[7]According to Rose (1999), an information intermediary is "an independent, profit maximizing economic information processing system performing its activities (information acquisition, processing, and dissemination) on behalf of other agents' information needs."

ings agencies actually contributes to ratings inflation and lower consumer surplus, contrary to the popular perception that competition improves ratings quality. In the issuer-pay business model adopted by ratings agencies, corporate issuers can shop for better credit ratings. In equilibrium, credit ratings agencies may loosen their standards for fear of revenue loss. A careful study of the role of competition among proxy advisors thus is warranted.

This is the first paper that examines effects of increased competition in the proxy advisory industry. Heightened competition is found to have reduced favorable recommendations for management proposals by ISS, the incumbent advisor widely believed to suffer from conflicts of interest arising from serving both shareholders and corporate issuers. This stands in sharp contrast to the role competition plays among credit ratings agencies. In the buyer-pay model adopted by proxy advisory firms, institutional investors now have an outside option (a competitor's reports) generated by competition (see Hörner, 2002). The existence of a competitor, especially when it is perceived as less conflicted, can discipline the incumbent advisor. Given this competitor's reports, investors may make a more informed guess about incumbent's truthfulness.

I begin this study by developing a model to analyze strategic behavior of a proxy advisor that sells investors a report on a management proposal, as well as consulting services to the management. I look at different market structures (monopoly and competition from a new entrant that sells only proxy reports) where proxy advisors obtain noisy information about the proposal's quality and issue reports to communicate the information. A central innovation of this model is how it captures demand for proxy advisory services. Investors may be either litigation-averse, in which case they follow the incumbent proxy advisor to avoid lawsuits from their clients in the case of a wrong vote, or rational, in which case they understand the structure of the game and can find out the advisors' incentives. The information reported is non-verifiable, but proxy advisors may suffer reputation costs (e.g., loss of future business) for misleading investors. This model shows that under monopoly, the incumbent advisor tends to inflate quality of the proposal when its expected reputation cost is low, compared to the consulting fee. Because of the disciplinary effect under competition (a higher probability of getting caught for issuing biased recommendations), the incumbent advisor is likely to be more truthful than under monopoly.

Before empirically analyzing potential conflicts of interest and the role of competition in reducing these conflicts, I provide evidence concerning proxy advisors' influence on vote outcomes. A major concern regarding conflicts of interest is that if biased recommendations translate into actual votes, shareholder value may be adversely affected. I show that endorsement by either of the dominant advisors, ISS or Glass Lewis, substantially increased the percent of "For" votes for management proposals, independent of ballot types and firm characteristics. This finding is consistent with prior research (see Choi, Fisch and Kahan, 2010; Ertimur, Ferri and Oesch, 2012). On the other hand, when these two advisors give conflicting recommendations, a proposal receives fewer favorable votes. Interestingly, as competition began to heat up, the ability of ISS's recommendations to predict vote outcomes diminished, while Glass Lewis became more influential. One plausible explanation is that investors were increasingly following Glass Lewis's recommendations, as it established itself as an alternative to ISS. The purpose of this exercise is to show that advisory firms can play an important role in shareholder voting, rather than claiming any causal relationship between voting recommendations and vote outcomes. In practice, investors may select a proxy advisor due to prior agreement with the advisor's governance philosophy.

A key contribution of this paper is to show empirically that increased competition brought by Glass Lewis's entry into the proxy market has reduced ISS's favoritism to corporate managers. Since entering the market in early 2003, Glass Lewis has grown into a credible competitor of ISS, capturing a market share of over 40% in 2011, measured by client assets (see Figure 1.1). Empirically, I use two methods to measure potential effects of competition. First, I examine whether or not there was a convergence of recommendations at the firm level when Glass Lewis was achieving a higher market share. Glass Lewis does not provide consulting services to corporate issuers, and is thus widely considered to be less conflicted. With more institutional shareholders subscribing to both companies, ISS's recommendations would be under intense scrutiny. It would potentially react to market pressure and engage in more truthtelling. This would increase the correlation between ISS's recommendations and Glass Lewis's. I find that the firm-level spread between ISS's and Glass Lewis's "For" recommendations shrank as Glass Lewis's market share increased, indicating that with Glass Lewis's entry, ISS was more likely to switch from making "For" recommendations to "Against/Withhold" than from making "Against/Withhold" recommendations to "For". To identify the direction of changes, I show that with a 10 percentage points increase in Glass Lewis's market share, the fraction of differing recommendations (ISS recommended "For," Glass Lewis recommended "Against/Withhold") indeed went down by 6 percentage points for the period 2004-2011. However, the fraction of differing recommendations (ISS recommended "Against/Withhold," Glass Lewis recommended "For") dropped by only 1 percentage point. These results suggest that competition had a disciplinary effect of making ISS less friendly to corporations.

Second, I look at how ISS's recommendations for a company changed when Glass Lewis began to cover it for the first time. When Glass Lewis (or any other proxy advisor) obtains a new institutional client, it has to cover all portfolio firms of the client. Prior to establishing the relationship, however, Glass Lewis does not know which companies are in the client's portfolio. Thus the event that Glass Lewis began to cover a new company served as an *exogenous* shock to ISS's recommendations for that company. I find that when Glass Lewis began to cover a company for the first time, ISS's average "For" recommendations for its management proposals dropped by 1.9-2.3%.

This order of magnitude seems small because it is the effect of competition on ISS's recommendations for an average firm, which might not be ISS's corporate client. The effect of competition is expected to be larger for ISS's corporate clients. Absent competition, ISS might be more likely to issue favorable recommendations for these companies due to the conflict of interest discussed in Section 1.2. Were the corporate client data not proprietary, I might be able to determine directly the magnitude of any conflict of interest – that is, whether or not companies that subscribed to ISS's consulting services were more likely to receive favorable recommendations from ISS, compared with similar firms that did not.

This paper contributes to the policy debate on whether and how to regulate the proxy advisory industry. My results suggest that conflicts of interest may be a real concern, and increasing competition can help to alleviate them to a certain extent. With respect to conflicts of interest, the SEC intends to issue an interpretive guidance that directs a proxy advisor to disclose "any significant relationship" with issuers or shareholder proponents. Another popular approach[8] is to issue directives similar to those addressing conflicts by credit ratings agencies, prohibiting certain conflicts of interest[9] and requiring specific disclosures, procedures, etc.

While this paper supports the view that greater competition is needed among proxy advisors the readiness of investors to support multiple advisory firms remains doubtful. A 2007 Government Accountability Office (GAO) study recognizes this dilemma, noting that some investors "questioned whether the existing number of firms is sufficient, while others questioned whether the market could sustain the current number of firms." Alternatively, some industry experts have argued that not-for-profit proxy advisors are conflict-free and can

---

[8]The SEC stated in its July 2010 concept release on the U.S. proxy system that "in light of the similarity between the proxy advisory relationship and the 'subscriber-paid' model for credit ratings, we could consider whether additional regulations similar to those addressing conflicts of interest on the part of Nationally Recognized Statistical Rating Organizations ('NRSROs') would be useful responses to stated concerns about conflicts of interest on the part of proxy advisory firms."

[9]The following conflicts credit ratings agencies face are similar to what proxy advisors face: conflicts of interest involved in both rating and helping to design the same securities; conflicts of interest in the provision of ancillary services to issuers whose securities they rate. (see Center On Executive Compensation, 2011). Other conflicts may arise when issuers shop for ratings (see Bolton, Freixas and Shapiro, 2012).

better serve the public interest. The entry of organizations like the Sustainable Investments Institute (Si2) which assists retail investors in voting seems to support such a “public utility” model. But eventual effects remain to be seen.

1.1.1 Related Literature

This paper is related to a large literature on information intermediaries in both microeconomics and finance (see e.g., Biglaiser 1993; Lizzeri, 1999). A parallel topic within this literature is credit ratings agencies. Bolton, Freixas and Shapiro (2012) analyze credit ratings agencies’ conflicts of interest, and find that increased competition leads to more ratings inflation, as issuers are able to more easily shop for ratings. Similar papers that study ratings inflation and shopping include Mathis, McAndrews and Rochet (2009), Skreta and Veldkamp (2008), Sangiorgi, Sokobin and Spatt (2009) and Camanho, Deb, and Liu (2012). Becker and Milbourn (2011) provide empirical support, finding that competition in corporate bond markets led to higher and less informative ratings. Griffin and Tang (2011) and Strobl and Xia (2012) provide further evidence of ratings inflation.

This work also relates to the literature on conflicts of interest in financial markets. Davis and Kim (2007) study mutual funds’ business ties with their portfolio firms, and find a positive relation between business ties and the propensity to vote with management. Hong and Kacperczyk (2010) find that competition among stock analysts reduces their optimism bias in their research. Similarly, a number of papers find that analysts from brokerage houses with underwriting relationship to a company tends to provide more positive forecasts than those from unaffiliated houses (e.g., Dugar and Nathan, 1995; Lin and McNichols, 1998; Dechow, Hutton and Sloan, 1999; Michaely and Womack, 1999).

Regarding the proxy advisory industry itself, a handful of papers have documented a strong association between proxy advisors’ recommendations and shareholder votes. The effect of ISS’s recommendations has been estimated at 14-21% for management proposals

(Bethel and Gillan 2002), between 13% and 30% for director elections, depending on the context and time period (Cai, Garner and Walkling, 2009; Choi, Fisch and Kahan, 2010; Ertimur, Ferri and Maber, 2012), and 25% for compensation-related shareholder proposals (Ertimur, Ferri and Muslu, 2011). Also, Alexander, Chen, Seppi, and Spatt (2010) find that an ISS recommendation in favor of a dissident in proxy contests increases the likelihood of the dissident's victory by 14%.

To my knowledge, only a few papers have studied impacts of Glass Lewis's recommendations, in addition to ISS's. Choi, Fisch and Kahan (2010) find that for director elections in 2005 and 2006, a Glass Lewis "Withhold" recommendation has a greater impact on a vote if ISS has issued a "For" recommendation than if ISS has issued a "Withhold" recommendation. This suggests the possibility that some institutional investors automatically will vote in favor of the board's nominees if both ISS and Glass Lewis issue "For" recommendations, but not if one of them issues a "Withhold" recommendation. Ertimur, Ferri and Oesch (2012) focus on say-on-pay votes in 2011, and find a negative recommendation from ISS (Glass Lewis) is associated with 24.7% (12.9%) more votes against the say-on-pay proposal. These papers, however, explore only a small portion of the data. My paper is the first to use Glass Lewis's comprehensive voting recommendations for the period 2004-2011. Together with ISS's voting recommendations, this panel dataset enables me to study effects of competition on incumbent advisor ISS's recommendations during that period.

Another strand of literature questions the value of proxy advisors' recommendations. Daines, Gow and Larcker (2010) find governance ratings do not predict governance-related outcomes, and there is little relation between ISS's governance ratings and its voting recommendations. Larker, McCall, and Ormazabal (2012) find that companies following proxy advisors' guidelines on stock option repricing had worse subsequent performance.

This work also relates to the broad shareholder voting literature that identifies various economic determinants of proxy voting outcomes (see e.g., Brickley, Lease and Smith, 1988;

Pound, 1988; Gordon and Pound, 1993; Morgan and Poulsen, 2001; Bethel and Gillan, 2002; Cremers and Romano, 2007; Cai, Garner and Walkling, 2009; Gillan and Starks, 2000; Maug and Rydqvist, 2009; and Matvos and Ostrovsky, 2010).

This chapter is organized as follows. In Section 1.2, I develop a model and derive testable implications. Section 1.3 provides more background information on the proxy advisory industry. Section 1.4 describes various data sources and presents summary statistics. In Section 1.5, I discuss the influence of voting recommendations. Section 1.6 discusses effects of competition, and Sections 1.7 and 1.8 evaluate effects of increased competition on the incumbent advisor's voting recommendations. Sections 1.9 and 1.10 offer further discussions and conclusion. References are listed in Section 1.11.

1.2 The Model

Consider a company owned by $N > 1$ institutional investors and run by management. Each investor holds one share, and casts exactly one vote during shareholder meetings (a "one-share-one-vote" rule). The management comes up with an exogenous proposal for a project that requires shareholder approval. The game lasts for one period.

Value of the proposal to investors is uncertain. Let $\omega \in \{a, o\}$ denote state of the world, in which a stands for "approve", and o stands for "oppose". If an "a" proposal is approved, nothing goes wrong. However, if an "o" proposal is approved, with probability p it leads to a loss and that is discovered by investors.

Investors share the same ex ante belief that the proposal is of type "a" with probability $\frac{1}{2}$. This is without loss of generality since an ex ante belief different from $\frac{1}{2}$ will not change our results. This creates a business opportunity for a proxy advisory firm (hereafter "PA"), which can use its technology to discover the state. The PA costlessly extracts a private signal $s \in \{a, o\}$, whose precision is $e \in (\frac{1}{2}, 1)$ and e is common knowledge:

$$\Pr(s = a|\omega = a) = \Pr(s = o|\omega = o) = e$$

Given the level of precision e, the PA proposes to sell its voting report to investors[10] for a fee f. This report will contain a recommendation of $m = A$ ("approve") or $m = O$ ("oppose"). If at least one investor subscribes to its service, the PA will retrieve a signal and make a report. Only investors who buy the report observe PA's recommendation.[11] At the same time, the PA provides consulting services to the company that investors own. The company will buy PA's services for ϕ if it makes a favorable recommendation $m = A$, otherwise the company refuses to purchase the product.[12] This creates potential conflicts of interest for the PA. We say the PA is conflicted when it receives a signal $s = o$ and reports $m = A$.

Institutional investors will be held liable if their clients find out that they made a wrong vote – that is, they voted for an "o" proposal. These investors are required by law (rules the SEC adopted in 2003) to vote proxies in the best interests of clients. The clients may sue[13] these institutional investors for breach of fiduciary duty[14] when they discover that their

[10]This buyer-pay business model is in stark contrast to the issuer-pay model adopted by credit ratings agencies. Competition among credit ratings agencies can lead to ratings inflation due to issuer shopping. Competition in the proxy advisory industry, however, may play a disciplinary role in the absence of shopping.

[11]For contested meetings in which a group of dissident shareholders seek shareholder support for their own slate of director nominees, investors may observe voting recommendations given by advisory firms even without purchasing their reports. The media tend to report high-profile proxy contests. On average there were only 52 contested meetings annually for the period 1994-2008, as documented by Fos (2011). This paper studies uncontested management proposals. For these proposals, the associated voting recommendations are usually not widely reported by the media.

[12]In practice, a company may buy the PA's consulting services in anticipation of its favorable recommendation. Modeling this will yield similar results. ϕ is normalized by investor size N as per capita profit.

[13]Alternatively, clients may withdraw funds from these institutions. Since management fees are based on assets under management, client withdrawal of funds is costly for such institutional investors.

[14]Institutional investors owe clients fiduciary duties under Section 36(b) of the Investment Company Act of 1940. Recent high-profile shareholder lawsuits for breach of fiduciary duty involve Janus Capital Group Inc. and the AXA Group Mutual Funds.

shares were voted for an "o" proposal. Investors will incur a litigation cost (both monetary and reputation costs) in case of such a lawsuit. On the other hand, voting against an "a" proposal does not involve any litigation cost because the proposal never leads to a loss, and the state therefore is not revealed.

There are two types of investors – litigation-averse and rational. A fraction α of investors are litigation-averse who wish to reduce litigation costs by relying on PA's advice. Since the PA is regarded as an expert in the corporate governance market, following its recommendation may convince clients that they acted in good faith. These investors often lack appropriate incentives or necessary expertise to do research in order to make informed votes. Without hiring the PA, expected litigation cost is C when a litigation-averse investor is held liable for voting for an "o" proposal. However, if the investor obtains the PA and follows its voting recommendation, then litigation cost is a smaller c (c is normalized to 0) when the PA is found out to be conflicted. In reality, many institutional investors with limited resources, such as small mutual funds, tend to be litigation-averse. The rest of investors are rational in that they understand PA's incentives and potential conflicts, and form their belief of PA's truthfulness.[15] However, they do not observe the state or signal of the PA. Rational investors will incur a cost $\tilde{C}$ if their clients discover that they voted for an "o" proposal. Rational investors can be thought of as well-incentivized asset management firms. Both types of investors will vote according to their ex ante belief – that is, each votes for the proposal with probability $\frac{1}{2}$, if they choose not to buy PA's report.[16]

If investors find out that the PA is conflicted, they will refuse to buy its reports in future periods. Investors, however, are able only to discover ex post whether the PA is conflicted

[15] Allowing the PA to disclose potential conflicts of interest will not qualitatively change the equilibria as the results are mainly driven by ligitation-averse investors, who will not use the disclosed information.

[16] We can allow rational investors to retrieve their own signal of the state, but this will not qualitatively change any result.

in the event of a loss. For example, after several months, investors involving in an M&A deal that the PA supported may see it fall apart. As a result, they investigate whether the PA received an "o" signal. If this is indeed the case, investors conclude that the PA did not relay its signal truthfully, and will not purchase any further reports. In practice it is difficult to determine if the PA is conflicted even ex post, but it is in general easier than ex ante. Formally, the PA will incur a reputation cost ρ in terms of the present value of future profits when it receives a signal $s = o$ and reports $m = A$, and a loss occurs. Reputation cost ρ is exogenous, as in Bolton, Freixas and Shapiro (2012) and Morgan and Stocken (2003).

As in Bolton, Freixas and Shapiro (2012), I assume that there is a small amount of uncertainty on the reputation ρ:

Assumption 1 *The PA is uncertain about the value of ρ: $\rho \in [\tilde{\rho} - \epsilon, \tilde{\rho} + \epsilon]$ such that $\epsilon \to 0$. After the PA receives its signal, the uncertainty is resolved.*

This tiny uncertainty prevents the PA from using mixed strategies for its report. Since most institutional investors have to diversify in their investments, the company is likely to be owned by many investors. N thus is assumed to be a large number. And without loss of generality, I assume that N is an even number throughout the analysis. The analysis in which N is odd is similar and is briefly described in the Internet Appendix. Voting rule is simple majority.

1.2.1 The Monopoly PA

I first analyze the monopoly game. The timing of the moves is as follows:

1. The PA posts fee f for its proxy report of a company.

2. Institutional investors of the company decide whether to buy the report.

3. The company issues a management proposal for shareholder vote.

4. The PA retrieves a private signal $s \in \{a, o\}$ and makes a report of $m \in \{A, O\}$.

5. The company uses PA's consulting services for fee ϕ if recommendation is "A"; otherwise the company does not use PA's services.

6. Outcome of the proposal realizes if approved. With probability p a loss occurs and that is found out by investors only if the state is "o". The PA incurs a reputation cost from investors who buy its report.

Litigation-averse investors would like to buy the report if the value of information exceeds price f. The profit-maximizing PA will set fee f low enough to woo litigation-averse investors. Thus there are two reporting regimes for the PA – it always reports $m = A$ when consulting fee ϕ is greater than its expected reputation cost, and it truthfully relays the signal when consulting fee is less than its reputation cost. An equilibrium also depends on whether the fraction of litigation-averse investors α is greater than $\frac{1}{2}$ and rational investors' prior belief of the PA's truthfulness.

To simplify notations for the equilibria, I introduce these following definitions:[17] β is the probability that the proposal is approved when all litigation-averse investors vote for it and each rational investor votes for it with probability $\frac{1}{2}$. γ is the probability that the proposal is approved when all litigation-averse investors vote against it and each rational investor votes for it with probability $\frac{1}{2}$. Since N is large, if one investor deviates (e.g., joins the other group in voting), the approval probabilities are still approximately β and γ, respectively.[18]

I proceed to derive the following symmetric equilibrium under each informational regime:

[17] $\beta = 1 - F\left((\frac{1}{2} - \alpha)N; (1-\alpha)N, \frac{1}{2}\right)$ and $\gamma = 1 - F\left(\frac{1}{2}N; (1-\alpha)N, \frac{1}{2}\right)$. And $F(\cdot)$ follows a binomial distribution.

[18] If one litigation-averse investor votes the same way as the rational investors, the actual approval probabilities are $\hat{\beta} = 1 - F\left((\frac{1}{2} - \alpha)N + 1; (1-\alpha)N, \frac{1}{2}\right); \hat{\gamma} = 1 - F\left(\frac{1}{2}N + 1; (1-\alpha)N, \frac{1}{2}\right)$, respectively. Similarly, if one rational investor votes the same way as the litigation-averse investors, the actual approval probabilities become $\check{\beta} = 1 - F\left((\frac{1}{2} - \alpha)N - 1; (1-\alpha)N, \frac{1}{2}\right); \check{\gamma} = 1 - F\left(\frac{1}{2}N - 1; (1-\alpha)N, \frac{1}{2}\right)$, respectively.

Proposition 1 *When $\alpha < \frac{1}{2}$, the Nash equilibrium of this game is:*

1. *If $\phi > \beta ep\rho\alpha$, the PA always reports $m = A$, and sets fee $f = \frac{1}{2}\frac{1}{2}\beta pC$. Only litigation-averse investors buy the report. PA's profit is $\phi + f\alpha - \frac{1}{2}\beta ep\rho\alpha$.*

2. *If $\phi < ep\rho$, the PA reports truthfully, and sets fee $f = \min[\frac{1}{2}\frac{1}{2}(1-e)pC, \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}]$. Both litigation-averse and rational investors buy the report. PA's profit is $\frac{1}{2}\phi + f$.*

The proof is in the Appendix.

This proposition demonstrates that the PA can maximize the present value of its profits by choosing either of the two informational regimes depending on rational investors' belief and parameter values. If rational investors believe the PA is conflicted, it maximizes its profit by always reporting $m = A$ when consulting fee ϕ is greater than expected reputation cost $\beta ep\rho\alpha$. The PA sets price f at a level that equals the benefit of additional information for litigation-averse investors. This reputation loss is from litigation-averse investors only since they are the ones who buy PA's report. Since the PA recommends $m = A$, with probability β an "o" proposal passes and with probability ep it leads to a loss that is discovered by investors. Conditional on these events, the reputation cost is $\rho\alpha$. For a low reputation cost ρ and a small number of litigation-averse investors (a small α), the PA takes advantage of litigation-averse investors by always reporting $m = A$. Voting behavior of rational investors also plays a role here because collectively they determine the probability of approval β.

Note that for reasonable parameter values, there is no equilibrium in which (1) the PA reports truthfully (when $\phi < \beta ep\rho\alpha$), and (2) only litigation-averse investors buy the report. Because the PA will lower fee f to a level at which it is beneficial for both types of investors to purchase its report.

There is another equilibrium in which both types of investors buy PA's report. If rational investors believe the PA is truthful, it will report truthfully when its consulting profit is less than expected reputation cost. The PA sets a fee that is lower than the cost of voting based on ex ante beliefs for either type of investors. Note this equilibrium does not depend on the value of α.

When $\alpha > \frac{1}{2}$, on the other hand, the PA completely determines vote outcomes.

Proposition 2 *When $\alpha > \frac{1}{2}$, the Nash equilibrium of the game is:*

1. *If $\phi > ep\rho\alpha$, the PA always reports $m = A$, and sets fee $f = \frac{1}{2}\frac{1}{2}pC$. Only litigation-averse investors buy the report. PA's profit is $\phi + f\alpha - \frac{1}{2}ep\rho\alpha$.*

2. *If $\phi < ep\rho$, the PA reports truthfully, and sets fee $f = \min[\frac{1}{2}\frac{1}{2}(1-e)pC, \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}]$. Both litigation-averse and rational investors buy the report. PA's profit is $\frac{1}{2}\phi + f$.*

The proof is similar to that of Proposition 1. In the case that rational investors believe the PA is conflicted, it always reports $m = A$ when consulting fee ϕ is greater than expected reputation cost $ep\rho\alpha$. Note that this reputation cost is greater than the corresponding one when $\alpha < \frac{1}{2}$. It is because now litigation-averse investors are the majority, and with probability 1 an "o" proposal passes and it leads to a loss with some probability.

1.2.2 Competition among PAs

I now analyze a game where two PAs compete in selling reports to investors. Incumbent player PA 1 sells reports to investors as well as consulting services to corporations, while new entrant PA 2 sells reports only to investors. As in Bolton, Freixas and Shapiro (2012), we can think of these PAs as providing differentiated reports given that they receive imperfect signals of the proposal's type. For simplicity, I assume that both PAs retrieve independent signals of the same precision $e > \frac{1}{2}$.[19] An investor may want to purchase both reports to obtain more information. The timing of the moves is similar to the monopoly game:

[19] In a richer model, one can also differentiate the qualities of PAs' signals. For example, we may assume that PA 1 retrieves a signal of higher precision ($e_1 > e_2$). This reflects the fact that PA 1 is more experienced than new entrant PA 2, and it thus receives a more "precise" signal.

1. Two PAs post fees f_i for their proxy reports of a company, where $i \in \{1, 2\}$.

2. Institutional investors of the company decide whether to buy one, both or neither report.

3. The company issues a management proposal for shareholder vote.

4. The PAs retrieve their signals $s_i \in \{a, o\}$ and make reports of $m_i \in \{A, O\}$, where $i \in \{1, 2\}$.

5. The company uses PA 1's consulting services for fee ϕ if recommendation is "A"; otherwise the company does not use PA 1's services.

6. Outcome of the proposal realizes if approved. With probability p a loss occurs and that is found out by investors only if the state is "o". The PAs incur respective reputation costs ρ_1 and ρ_2 from investors who buy their reports.

Since PA 2 does not provide consulting services, in equilibrium it always truthfully relays its signal because recommending $m = A$ when retrieving an "o" signal does not yield additional profits. However, PA 1 may be conflicted when its consulting fee is greater than expected reputation cost. Understanding both PAs' incentives, rational investors will choose to purchase a report from PA 2 when it charges a fee that makes these investors indifferent toward buying its report or not.[20] PA 2 attempts to charge a price that makes it beneficial for both litigation-averse and rational investors to purchase its report. However, anticipating this, PA 1 will lower its price to a point that makes litigation-averse investors indifferent towards purchasing a report from either PA. PA 2, on the other hand, can lower its price only to a level that makes it no worse off by serving both types of investors than only serving

[20]In a "truthful" equilibrium in which both PAs relay their signals truthfully, rational investors will be indifferent about purchasing reports from either PA if the PAs charge the same price. For simplicity, I assume that rational investors will only purchase a report from PA 2 in this case.

rational investors. Litigation-averse investors thus will purchase a report from PA 1. Understanding PA 1's strategy, PA 2 raises its fee and sells reports to rational investors only. This price competition leads to the following lemma:

Lemma 1 *There exists no equilibrium in which both types of investors purchase reports from PA 2 only.*

The proof is in the Appendix.

Price competition from PA 1 leads to market segmentation: litigation-averse investors purchase a report from PA 1, while rational investors buy a report from PA 2. In such an equilibrium featuring segmentation, competition plays two roles. First, although litigation-averse investors stick to incumbent player PA 1, they now have an outside option (buying PA 2's report instead) created by competition. The existence of competitor PA 2, can serve as a disciplinary device for PA 1, at least ex post (see Hörner, 2002). In case a loss occurs, litigation-averse investors may be able to observe PA 2's report, and thus make a more informed guess about PA 1's truthfulness. When the PAs disagree on their recommendations, litigation-averse investors may examine PA 1's report more carefully. Choi, Fisch and Kahan (2010) suggest that some investors automatically vote for the board's nominees if both PAs issue "For" recommendations, but not if one of them issues a "Withhold" recommendation.

In practice, companies regularly learn proxy advisors' voting recommendations after votes are cast, and could pass on the information to investors if requested. It is reasonable to assume that with this extra information on hand, litigation-averse investors will be more likely to discover whether PA 1 is conflicted compared with the monopoly case. Define $\tilde{p}$ as the probability that a loss occurs and that is discovered by investors when the PAs disagree on their recommendations. So we have $\tilde{p} > p$, where p is the corresponding probability under monopoly.

For simplicity, I further assume that investors can not decide whether PA 1 is conflicted in case a loss occurs after the PAs issue the same recommendation. This is a somewhat extreme assumption and is not essential, but it helps to simplify the analysis. However, it is likely that when investors receive an identical recommendation from both PAs, they are less likely to perform due diligence themselves. The empirical section of this paper shows evidence that when two prominent PAs give the same recommendation for a proposal, the probability of approval is much higher than an average proposal. This suggests that more investors are likely to follow voting recommendations automatically when both PAs issue the same recommendation. It is worth noting that there are potentially other ways to model the effect of competition.[21]

Second, competition affects rational investors' voting behavior. In the monopoly game, these investors vote based on ex ante belief, and now they will follow PA 2 to make informed votes. This matters for vote outcomes when rational investors are the majority ($\alpha < \frac{1}{2}$). As in the monopoly case, I derive symmetric equilibria under each informational regime:

Proposition 3 *When $\alpha < \frac{1}{2}$, the equilibrium of the subgame is:*

1. *PA 1 always reports $m_1 = A$ and sets fee $f_1 = \min[\frac{1}{2}\frac{1}{2}(1-e)pC, \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}]$. PA 2 reports truthfully and sets fee $f_2 = \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}$. Litigation-averse investors purchase PA 1's report, and rational investors buy PA 2's report.*

The proof is in the Appendix.

In this equilibrium, rational investors are the majority and they determine vote outcomes. PA 1 understands that these investors will vote according to PA 2's recommendation. When PA 1 receives an "o" signal, it always reports $m_1 = A$ due to the following reason: if PA 2 retrieves an "a" signal, it reports $m_2 = A$. Since the PAs give an identical recommendation,

[21] In a richer dynamic model in which reputation ρ is endogenously determined, one can model competition as an option of switching to PA 2's services (see Hörner, 2002).

it is difficult ex post for litigation-averse investors to determine which PA is conflicted. Thus PA 1 will not incur a reputation cost. On the other hand, if PA 2 receives an "o" signal, the proposal is defeated. As a result, PA 1 will not be liable either.

It is worth noting that this case is a "limiting" result. To the extent that PA 1 believes that there is a certain probability that rational investors do not follow PA 2's recommendation (after purchasing its report), the expected reputation cost will be greater than zero. This is because now there is a positive probability that an "o" proposal will be approved. In turn, there is a positive probability that PA 1 will truthfully relay its signal. In other words, PA 1 can adopt a mixed strategy for its truthfulness.

I define q as the probability of being in the "o" state given PAs' signals $s_1 = s_2 = o$. As shown in the proof of Proposition 4, this is the probability that PA 1 incurs a reputation cost when it is conflicted (it recommends $m_1 = A$) in the case $\alpha > \frac{1}{2}$. I now derive the equilibrium of this game when litigation-averse investors are the majority.

Proposition 4 *When $\alpha > \frac{1}{2}$, the equilibrium of the subgame is:*

1. *If $\phi > \frac{1}{2}q\tilde{p}\rho_1\alpha$, PA 1 always reports $m_1 = A$ and sets fee $f_1 = \min[\frac{1}{2}\frac{1}{2}pC, \frac{1}{2}\frac{1}{2}p\tilde{C}]$. PA 2 reports truthfully and sets fee $f_2 = \frac{1}{2}\frac{1}{2}p\tilde{C}$. Litigation-averse investors purchase PA 1's report, and rational investors buy PA 2's report.*

2. *If $\phi < \frac{1}{2}q\tilde{p}\rho_1\alpha$, both PAs report truthfully. PA 1 sets fee $f_1 = \min[\frac{1}{2}\frac{1}{2}(1-e)pC, \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}]$ and PA 2 sets fee $f_2 = \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}$. Litigation-averse investors purchase PA 1's report, and rational investors buy PA 2's report.*

The proof is in the Appendix.

This proposition establishes that PA 1 can maximize the present value of its profits by choosing either of the two informational regimes. It maximizes its profit by always reporting

$m_1 = A$ when consulting fee ϕ is greater than its expected reputation cost $\frac{1}{2}q\tilde{p}\rho_1\alpha$.[22] This reputation cost is from litigation-averse investors only since they are the ones who buy the report. PA 1 sets fee f_1 at a level that makes litigation-averse investors weakly prefer buying a report from PA 1. Anticipating this, PA 2 raises the price and just sells the report to rational investors. As in the monopoly game, for a low reputation cost ρ_1 and a small number of litigation-averse investors (a small α), PA 1 takes advantage of the litigation-averse investors by always reporting $m_1 = A$. PA 1, however, will remain truthful when consulting fee ϕ is less than its expected reputation cost.

Each PA charges a different fee for its report between the two informational regimes. This is because for each type of investors, costs of voting based on ex ante beliefs are different between the two cases. In case 1, the proposal is always approved while in case 2, the proposal passes only when PA 1 receives an "a" signal. Thus costs of voting based on ex ante beliefs are higher for both types of investors in case 1.

Competition leads to more truthtelling if PA 1's reputation cost $\frac{1}{2}q\tilde{p}\rho_1\alpha$ is greater than $e\tilde{p}\rho\alpha$, which is its reputation cost under monopoly. This condition requires $\frac{\tilde{p}}{p} > \frac{2}{q} \cdot \frac{\rho}{\rho_1}$. It can be shown that $\frac{2}{q} < 4$.[23] Therefore, when $\frac{\tilde{p}}{p} > 4 \cdot \frac{\rho}{\rho_1}$, competition alleviates conflicts of interest arising from PA 1's consulting services. This, of course, is an empirical question. More competition, however, will likely push up $\tilde{p}$, where $\tilde{p}$ equals to the probability that investors discover that a loss has occurred.

In practice, it may be beneficial for litigation-averse investors to subscribe to PA 2's report, in addition to PA 1's. Some asset management firms indeed subscribe to multiple

---

[22] When PA 1 retrieves signal $s_1 = o$ and reports $m_1 = A$, and PA 2 retrieves signal $s_2 = o$ and reports $m_2 = O$, the proposal passes because litigation-averse investors are the majority. The associated reputation cost for PA 1 is $q\tilde{p}\rho_1\alpha$. However, when PA 1 retrieves signal $s_1 = o$ and reports $m_1 = A$, and PA 2 receives signal $s_2 = a$ and reports $m_2 = A$, PA 1 does not suffer a reputation cost because it is difficult for investors to determine which PA is conflicted given the same recommendation. Thus the expected reputation cost is $\frac{1}{2}q\tilde{p}\rho_1\alpha$.

[23] Since $q = \frac{e^2}{e^2+(1-e)^2}$, $\frac{2}{q} = 2\left[1 + \frac{(1-e)^2}{e^2}\right] < 4$. Note that $e > \frac{1}{2}$.

proxy advisory services. This enables these institutional investors to collect more information before votes are cast. When the PAs give conflicting recommendations on the same issue, investors may scrutinize the proposal. When the PAs issue an identical recommendation, investors may feel more confident that they cast the right vote. In the Appendix, I show that under certain assumptions there exist equilibria in which litigation-averse investors subscribe to both PAs' reports, while rational investors purchase a report from PA 2.

1.2.3 Predictions

The model allows me to address three issues regarding PAs' strategic behavior and shareholder votes. Propositions 1 and 2 show that when consulting profit ϕ is greater than expected reputation cost, the incumbent PA is less likely to report truthfully. This leads to the following predictions:

Prediction 1 *The incumbent PA is more likely to issue favorable recommendations for companies that subscribe to its consulting services, compared with the ones that do not.*

Prediction 2 *The incumbent PA is more likely to issue favorable recommendations for companies that pay more consulting fees.*

With competition, the likelihood that the incumbent PA is conflicted (receives an "o" signal and reports $m_1 = A$) depends on parameter α – the fraction of litigation-averse investors. When $\alpha > \frac{1}{2}$, the vote outcome is determined by litigation-averse investors. If the incumbent PA is conflicted, there is a greater chance that it will be discovered because investors now can access the new entrant's recommendation ex post. This is because when the PAs disagree on their recommendations, litigation-averse investors may examine the incumbent PA's report more carefully. The incumbent PA thus is more likely to report truthfully under competition than monopoly. When $\alpha < \frac{1}{2}$, however, the incumbent PA will not be caught if it is conflicted. The reason is that if the new entrant receives an "o" signal (and it reports $m_2 = O$), the proposal is defeated given that rational investors determine vote outcomes. We therefore have the following prediction:

Prediction 3 *When $\alpha > \frac{1}{2}$, the incumbent PA is more likely to report truthfully under competition than monopoly. The fraction of cases where the incumbent recommends "For", and the new entrant recommends "Against" are diminishing as competition intensifies. When $\alpha < \frac{1}{2}$, the incumbent PA is always conflicted.*

Prediction 3 reveals that effects of competition on the incumbent PA's truthfulness depend on the value of α, the proportion of litigation-averse investors. It is a priori not clear which group of investors are the majority. Whether competition plays a disciplinary role thus is an empirical question. I test the effect of competition in Section 1.7.

Prediction 4 *Under competition, when two PAs give conflicting recommendations, the proposal is less likely to be approved. However, when the PAs issue an identical recommendation, there is a stronger correlation between the recommendation and vote outcome.*

Note that this prediction is embedded in the derivation of Propositions 3 and 4. Due to market segmentation, litigation-averse (rational) investors tend to follow the incumbent (new entrant) PA. When the PAs give the same recommendation for a proposal, investors tend to vote the same way. When the PAs disagree on the proposal, the likelihood that it will pass is smaller.

The remaining sections will take these predictions to the data. Due to data constraints, Predictions 1 and 2 will not be tested directly.

1.3 Background and Business Model

This section explains in detail important market and regulatory developments in the proxy advisory industry. In 1988, the Department of Labor issued a letter mandating that pension funds have a fiduciary duty to vote their proxies in the best interest of their clients. This prompted managers of employee retirement plans to seek advice from ISS, which was established in 1985 and was the only proxy advisor at that time.

In the 1990s and early 2000s, ISS's dominance in the industry continued to rise, thanks to growing fiduciary obligations of institutional investors and increased shareholder activism in the aftermath of the dot-com bust. Institutional investors hired proxy advisory firms to help them assess corporate governance practices at public companies.

In 2003, the SEC reinforced fiduciary duties of investment advisors with respect to proxy voting through widened application of the Investment Advisers Act of 1940. These expanded rules require mutual funds to publicly disclose their voting records, as well as adopt policies and procedures to ensure that they vote proxies in the best interests of clients. These requirements led to a rapid increase in demand for proxy advisory and governance services.

Since 2004, many large corporations (80% of S&P 500 companies) have adopted some majority voting standard for director elections, thanks to a number of shareholder initiatives and a series of amendments[24] that facilitate the adoption of majority voting by company boards. This has greatly increased leverages that investors and proxy advisors have over directors. To curb bad compensation practices that potentially contributed to the 2008-2009 financial crisis, the Dodd-Frank Act (Section 951) requires companies to hold a non-binding shareholder say-on-pay vote at least once every three years to "approve" executive compensation. This new requirement applies to all shareholder meetings held after January

[24]Model Business Corporation Act (MBCA) and the Delaware General Corporation Law were amended to facilitate the adoption of majority voting by company boards or by shareholders.

21, 2011. "The overall effect of say-on-pay will be to increase the influence of proxy advisory firms as investors grapple with more than 16,000 additional proxy votes in 2011, many of which require an understanding of each company's pay philosophy and arrangements."[25]

As required by the Dodd-Frank Act, the SEC in 2010 approved rules to eliminate broker discretionary voting in uncontested elections as well as executive compensation matters, including say-on-pay votes. This rule change is thought to be significant because brokers tend to cast uninstructed broker votes in favor of management and can comprise up to 20 percent of total proxy votes. Combined with majority voting, it could result in more directors failing to achieve majority support from shareholders. The elimination of broker non-votes will likely enhance institutional investors' power. This in turn increases the influence of proxy advisory firms.

1.3.1 Major Proxy Advisors

Today, the proxy advisory industry is dominated by just two firms: ISS and Glass Lewis. This duopoly structure has allowed them to have a significant influence on pay and corporate governance policy. Since 1985, ISS has become a leading player in both proxy advisory services and corporate governance ratings. It is currently owned by MSCI Inc.,[26] a leading provider of investment decision support tools to investors worldwide. As of 2007, ISS had 1,700 institutional clients, and a market share of 61%, based on clients' equity assets. Its clients included 24 of the top 25 mutual funds, 25 of the top 25 asset managers, and 17 of the top 25 public pension funds. ISS's core business includes proxy research and voting recommendations. It also provides web-based voting services and consulting services to

[25] Center on Executive Compensation (2011).

[26] In early 2007, ISS was purchased by RiskMetrics Group Inc., a leading provider of risk assessment and wealth management products. In 2010, RiskMetrics was acquired by MSCI Inc. in a transaction valued at $1.57 billion.

corporate issuers through ISS Corporate Services, Inc. (ICS), a wholly-owned subsidiary of ISS. ICS provides products and services on executive compensation, corporate governance ratings, voting analytics and governance research. ISS's business model of selling data and consulting services to corporations while advising investors how to vote on proposals of the same issuers has led to charges that ISS is seriously conflicted. In 2011, approximately 21.2% of ISS's total revenue was generated from its ICS subsidiary.[27] Despite vehement criticism for potential conflicts of interest created by its consulting services, ISS has been reluctant to spin off this business because of its high profitability.[28] In fact, some industry experts believe that without this highly profitable business, ISS's operations would be, at best, only marginally profitable.

Glass Lewis[29] was founded in early 2003, and has quickly established itself as ISS's main competitor, controlling 37% of the market share in 2007. At the end of 2010, Glass Lewis acquired Proxy Governance, Inc.'s 100 clients after the latter exited the market, further increasing its market share. In 2011, it covered around 23,000 companies in more than 100 countries, inching closer to ISS's coverage of 26,000 companies. A 2004 New York Times article reported that "Glass Lewis has unseated [ISS] . . . from its position as the undisputed leader in the field." Like ISS, Glass Lewis provides proxy research and vote recommendations to institutional shareholders. Glass Lewis's ability to quickly cut into ISS's market share owes partly to the fact that it does not sell corporate governance services to corporations. Many investors view Glass Lewis as less conflicted. A Glass Lewis executive stated in an email:

[27] See MSCI Inc. Annual Report (Form 10-K) for fiscal year 2011.

[28] Much of the consulting revenue results from charging corporations for use of the ISS compensation model.

[29] According to co-founder Gregory P. Taxin, the firm was named for two Supreme Court justices who fought for individual rights and ethical corporate practices. "'Glass' is derived from the surname of William O. Douglas, a former Securities and Exchange Commission chairman and a justice from 1939 to 1975, while 'Lewis' is a bow to Louis D. Brandeis, a justice from 1916 to 1939 who wrote *Other People's Money and How the Bankers Use It*." (see Morgenson, 2004)

"We do not advise or consult with corporations regarding their proxies; we believe to do so would compromise our ability to objectively evaluate those proxies and advise our clients on how to vote their shares." Unlike ISS, Glass Lewis is not registered as an investment advisor and hence is not directly regulated by the SEC.

After Proxy Governance's exit at the end of 2010, there remain two other for-profit proxy advisory firms, Egan-Jones Proxy Services and Marco Consulting Group. A new firm, ProxyTell, LLC, appears to have entered the market in 2012. They collectively own less than 2% of the market share, and thus is not a part of this research.

1.3.2 Concerns over Conflicts of Interest

The most common concern about proxy advisory firms, especially ISS, is potential conflicts of interest inherent in their business model. As discussed above, ISS provides services to both institutional investors and corporate issuers on the same governance issues, while Glass Lewis serves only institutional investors. A 2007 GAO study summarizes ISS's potential conflicts of interest as follows: "For example, some industry professionals stated that ISS could help a corporate client design an executive compensation proposal to be voted on by shareholders and subsequently make a recommendation to investor clients to vote for this proposal. Some industry professionals also contend that corporations could feel obligated to subscribe to ISS's consulting services in order to obtain favorable proxy vote recommendations on their proposals and favorable corporate governance ratings."

Responding to these public charges, ISS has installed a "Chinese Wall" between its proxy advisory services and corporate consulting services, creating a separate subsidiary ICS to serve corporate issuers. According to ISS, the "Chinese Wall" includes "legal, physical and technological separations." ISS also makes substantial disclosure to its institutional clients, as well as adopts a "Code of Ethics" that applies to all employees regarding conflicts of interest. However, these measures do not solve inherent conflicts of interest embedded in its business model. This paper studies whether ISS's potential conflicts have become actual conflicts, and whether increased competition from Glass Lewis has mitigated them.

Although this type of conflict is widely considered the most damaging, by no means it is the only source of potential conflicts of interest. For example, the fact that proxy advisory firms are owned by parent companies providing other financial services to clients has drawn scrutiny. ISS is owned by MSCI Inc., a leading provider of investment decision support tools to institutions, and Glass Lewis is an indirect wholly-owned subsidiary of the Ontario Teachers' Pension Plan Board (OTPP), a large activist pension fund in Canada. These issues are beyond the scope of this paper.

1.3.3 Recent Regulatory Developments

It is surprising that proxy advisory firms are subject to little regulation despite their impact upon investors and the importance of proxy voting to corporate governance and capital markets. The principal governmental oversight for these firms is the Investment Advisers Act of 1940, but proxy advisors can easily escape such regulations. At the present time, the only real oversight comes from institutional investors, who have little incentive to monitor because proxy advisors provide cost-effective services which benefit their clients.

Concerns over conflicts of interest and other issues (such as barrier to competition, a lack of transparency, potential inaccuracies and limited engagement with issuers) have led to two GAO studies and a concept release on the U.S. proxy system issued by the SEC in July 2010. SEC Chairman Mary Schapiro noted that both companies and investors "have raised concerns that proxy advisory firms may be subject to conflicts of interest or may fail to conduct adequate research and base recommendations on erroneous or incomplete facts." According to a June 2012 article in the CFO Journal, the SEC will be issuing an interpretive guidance to advise investors about their fiduciary duties in assessing information provided by proxy advisors and potential conflicts of interest. The SEC is unlikely to address a perceived lack of competition among proxy advisors or otherwise limit the use of proxy advisors. Instead, this guidance will focus on existing rules on investor fiduciary duty and conflicts of interest.

In June 2012, the Canadian Securities Administrators (CSA) also issued a white paper on possible regulation of proxy advisory firms. The CSA aims to address regulatory concerns about the services provided by proxy advisory firms and their potential impact on the capital markets. In March 2012, the European Securities and Markets Authority (ESMA) published a discussion paper that considers possible policy options on proxy advisory firms. In the same month, the French Autorité des Marchés Financiers (AMF) proposed practice recommendations for proxy advisory firms. However, as of today, no rules have yet been adopted by any country.

1.4 Data Description

This study draws data from a number of sources. My primary datasets are ISS's Voting Analytics database and Glass Lewis's Proxy Paper database. Both datasets cover shareholder meetings during the period 2004 – 2011. Voting Analytics provides the identity of companies, description of ballot items, shareholder meeting dates, management and ISS recommendations, and the number of "For" and "Withhold/Against" votes, as well as other information. It covers all Russell 3000 companies[30] since 2005, and includes most of the Russell 3000 companies before 2005. This dataset is becoming popular among corporate governance experts, as well as academics. Most exsiting papers use only data before 2005 (see Cai, Garner and Walkling, 2009; Matvos and Ostrovsky, 2010). My dataset on Voting Analytics is comprehensive and the most up-to-date.

[30]These are the largest 3,000 publicly held U.S. companies based on total market capitalization, which represents approximately 98% of the investable U.S. equity market.

Glass Lewis's Proxy Paper database contains similar information to Voting Analytics.[31] My paper is unique in its reliance upon Glass Lewis in addition to Voting Analytics, and only this allows for the analysis of effects of competition on incumbent advisor ISS's recommendations during the period 2004-2011. Prior research has only explored a small portion of the data. Choi, Fisch and Kahan (2010) use Glass Lewis's voting recommendations for director elections at S&P 1500 companies in 2005 and 2006. Ertimur, Ferri and Oesch (2012) focus on say-on-pay recommendations at S&P 1500 companies in 2011. My own work matches these two databases using CUSIP, meeting date and ballot item number. I exclude proxy contests[32] which yields 26,304 shareholder meetings at 4,807 unique companies.

For the same period, I collect numbers of ISS and Glass Lewis's institutional clients, total client assets, as well as numbers of U.S. meetings covered, all of which are annual figures. The main sources are LexisNexis[33] and Glass Lewis's website. Evolution of Glass Lewis's market share based on client assets is plotted in Figure 1.1.

I obtain additional data from following sources: stock information from Center in Research for Security Prices (CRSP), company accounting data from Compustat, Top-5 executives' compensation and stock holdings from ExecuComp, firm governance characteristics from RiskMetrics, and institutional holdings from Thomson-Reuters (13F). Voting data are matched with these datasets on CUSIP and fiscal year. This is my main dataset. Additionally, I match director characteristics, also obtained from RiskMetrics, to the voting data, using director last name, CUSIP and year.

---

[31]In addition to Russell 3000 companies, Glass Lewis's Proxy Paper database covers smaller firms.

[32]This paper studies uncontested management proposals, both theoretically and empirically. In a proxy contest, a group of dissident shareholders seek shareholder support for their own slate of director nominees, rather than the board's nominees. The purpose of launching proxy contests is to gain corporate control. This mechanism is more complex than uncontested management proposals, and is beyond the scope of this paper. Interested readers are referred to Brav, Jiang, Partnoy and Thomas, 2008, Klein and Zur, 2009, and Fos (2011).

[33]LexisNexis maintains the world's largest electronic database for legal, news and business information.

1.4.1 Voting and Company Characteristics: 2004-2011

Every public company in the U.S. holds an annual general meeting to elect the Board of Directors and to transact other businesses such as executive compensation plans, ratification of auditors, merger and acquisition. Starting January 2011, companies are required to submit say-on-pay proposals for shareholder approval. In my sample, 90% of companies use a plurality voting system for directors under which shareholders can vote “For”, “Withhold” or “Abstain.” ISS and Glass Lewis make “For” or “Withhold” recommendations. The remaining 10% of firms use a majority voting rule[34] under which ISS and Glass Lewis recommend “For” or “Against.” Under a plurality rule, a director will be elected in uncontested meetings even if she receives less than 50% of the base. Under a majority rule, a director has to step down if she fails to receive 50% of total votes. The base for director elections is usually defined as “For+Against/Withhold.” Thus I measure director election outcomes as the number of “For” votes divided by the sum of “For” and “Against/Withhold” votes. ISS and Glass Lewis recommend “For” or “Against” for all other ballot items. The base for these items is usually “For+Against+Abstain.”

In Panel A of Table 1.2, I calculate the percent of “For” recommendations for executive compensation plans and say-on-pay proposals for each company-year pair. I also calculate the average percent of “For” recommendations for directors within each company in a given year. They are done for ISS and Glass Lewis separately for the period 2004-2011.[35]

I also control for previous-year firm performance using both market-based and accounting-based returns. The market-based return is a firm’s stock return in the 12 months prior to its annual meeting. I also use 1-year excess return, 3-year excess return or abnormal return from Fama-French (1993) three factor models. The results are similar and are not reported

[34]Many large companies, 80% of S&P 500 companies, have adopted a majority voting rule.

[35]Since say-on-pay proposals started in January 2011, associated metrics are calculated for 2011 only.

due to space.[36] For the accounting-based return, I use industry-adjusted return on assets (ROA).[37] ROA is defined as EBITDA divided by total assets.

Since 2007, RiskMetrics no longer produces the governance index of Gompers, Ishii and Metrick (2003). I use an alternative governance indicator which equals 1 if a company has both a classified board and a poison pill (see Bebchuk and Cohen, 2005; Cai, Garner and Walkling, 2009). A combination of a classified board and a poison pill makes corporate control change more difficult, and is seen as a decrease in corporate governance quality. In the sample, 29% of firms have both policies in place. I collect information on board size, the percent of independent directors and institutional and management ownerships. The median board in my sample has nine members and comprises 75% independent directors. Institutional investors hold almost three-quarters of the shares. These figures are consistent with findings in the extant literature (e.g., Cai, Garner and Walkling, 2009).

As in Walkling and Long (1984) and Hartzell, Ofek and Yermack (2004), I use abnormal executive compensation as a measure of corporate governance. I estimate abnormal compensation as the residual from a linear compensation regression of all ExecuComp firms during my sample period. I include log assets, prior-year stock return, and industry and year dummies as independent variables.

Voting mechanism is important when we analyze the effect of voting recommendations on the actual votes. For example, for firms having confidential voting in place, shareholders may be tougher with management proposals because firm policy prevents management from knowing how shareholders vote their proxy cards. Shareholders will be less concerned about retaliation from managers. Unequal voting provisions,[38] on the other hand, usually

[36] Results using 1-year excess return, 3-year excess return or abnormal return are available upon request.

[37] Following standard literature (e.g., Brav, Jiang, Partnoy and Thomas, 2008), I subtract the median ROA for all Compustat firms in the same 4-digit SIC industry and year.

[38] These provisions limit voting rights of some shareholders and expand those of others. Under time-phased voting, shareholders who have held the stock for a given period of time are given more votes per share than

benefit the management because managers are often given more votes per share than recent purchasers (time-phased voting). All the above statistics are shown in Panel B of Table 1.2.

1.5 Influence of Voting Recommendations

Before empirically analyzing potential conflicts of interest and how competition can reduce these conflicts, I provide evidence on how proxy advisors influence vote outcomes. A major concern regarding conflicts of interest is that if biased recommendations translate into actual votes, shareholder value may be adversely affected.

In this section, I examine how ISS and Glass Lewis's recommendations affect vote outcomes, and tease out their relative magnitudes. A vote outcome is a function of voting recommendations as well as firm performance and governance characteristics. I analyze compensation plans, say-on-pay proposals and director elections separately. Columns (1)-(2) of Table 1.3 show investors' reactions to compensation recommendations. A positive ISS (Glass Lewis) recommendation was associated with 23.8% (8.3%) more votes for a compensation proposal. These estimates are in line with Cai, Garner and Walkling (2009) and Choi, Fisch, and Kahan (2009).[39] In the Internet Appendix, I also show that ISS's influence had declined from the previous period (2004-2007) to the recent period (2008-2011), while Glass Lewis's influence had been on the rise. Vote results are related to voting mechanisms as well. Compensation plans at firms with unequal voting (dual class shares) received higher votes, possibly from managers.

Mandatory say-on-pay proposals began in early 2011, so with the data I have, it is

---

recent purchases. Another variety is the substantial shareholder provision, which limits voting power of shareholders who have exceeded a certain threshold of ownership. In my sample, only 1% of companies adopted unequal voting rules.

[39]Cai, Garner and Walkling (2009) first estimate a regression model of ISS's recommendations based on firm performance and governance characteristics, and then use the residuals from this model as their ISS variables.

possible only to compare influences of ISS and Glass Lewis's recommendations for that year. Consistent with results for compensation plans and director elections, columns (3)-(4) show that ISS's (Glass Lewis's) endorsement was associated with an increase of votes by 23.8% (12.6%). These magnitudes are in line with findings in Ertimur, Ferri and Oesch (2012). Interestingly, a "For" recommendation from Glass Lewis on top of ISS's endorsement only added 6% to the vote, which suggests that the marginal value of an additional "For" recommendation would be small.

Columns (5)-(6) show investors' reactions to director recommendations. A positive ISS (Glass Lewis) recommendation was related with an increase of votes by 21.7% (5.8%). As shown in the Internet Appendix, although influences of both ISS and Glass Lewis had increased since the previous period (2004-2007), the increase for Glass Lewis was more dramatic.

The aim of this section is to show that proxy advisory firms play some important role in proxy voting, though there is not necessarily a causal relationship between voting recommendations and vote outcomes. As Choi, Fisch, and Kahan (2010) point out, "investors may select a proxy advisor based on their ex-ante agreement with the bases upon which the advisor formulates its recommendations."

1.6 Effects of Competition

In the following two sections, I propose two ways to investigate impacts of competition on decisions of incumbent advisor ISS. First, I show that there was a convergence of recommendations at the firm level as Glass Lewis's market share increased. In particular, following Glass Lewis's entry the fraction of differing recommendations (ISS recommended "For," Glass Lewis recommended "Against/Withhold") went down significantly, while the fraction of differing recommendations (ISS recommended "Against/Withhold," Glass Lewis recommended

"For") barely dropped. Second, I examine whether ISS adjusted its recommendations for a company after Glass Lewis began to cover that firm for the first time. To the extent that conflicts of interest mainly arise when ISS serves both corporate clients and investors, this is a measure of the disciplinary effect of competition. Effects of competition on ISS's recommendations should be mostly felt at companies that subscribed to its consulting services. After all, ISS was more likely to be conflicted in issuing voting recommendations for these companies because these firms contributed a significant portion of its profits. Given information on ISS's corporate client base, we can test whether ISS mostly responded to rival coverage of its corporate clients rather than its non-client firms.

1.7 Competition and Convergence of Recommendations

1.7.1 Estimation Strategy: Fixed Effects

To quantitatively examine effects of competition on the convergence of recommendations for the period 2004-2011, I first regress the firm-level spread between ISS's and Glass Lewis's "For" recommendations on Glass Lewis's market share and a large number of firm observable characteristics:

$$ISS\_AvgFor_{ijt} - GL\_AvgFor_{ijt} = \alpha + \delta \cdot GL\_MktShr_t + X'_{ijt}\beta + \psi_i + \eta_t + \mu_j \cdot t + \epsilon_{ijt} \quad (1.1)$$

In equation (1.1), the dependent variable is the difference between ISS's and Glass Lewis's recommendations for firm i in industry j in year t. GL_MktShr_t is Glass Lewis's market share in year t. X_{ijt} is a vector of firm characteristics including size, performance metrics, executive compensation measures, governance indicators and institutional and management holdings. Year fixed-effects η_t control for economy-wide trends that affect recommendations, and company fixed-effects ψ_i control for all time-invariant firm-level variables. I further include industry-specific time trends $\mu_j \cdot t$ to account for differential linear trends in recommendations across industries. The coefficient of interest is δ, which measures how the spread between these two advisors' recommendations changes as Glass Lewis's market share increases. Standard errors are clustered at the firm level.

To separately identify the direction of changes in ISS's recommendations, I first replace the dependent variable in equation (1.1) by the fraction of differing recommendations at the firm level (ISS "For," Glass Lewis "Against/Withhold"). I then use the fraction of differing recommendations in the other direction (ISS "Against/Withhold," Glass Lewis "For"). Our coefficient of interest is again δ, which now gauges how the fraction of differing recommendations evolves following Glass Lewis's entry.

1.7.2 Competition Metric

The intensity of competition can be measured by Glass Lewis's market share. Since Glass Lewis entered the proxy advisory market, it has increased its market share substantially. There are potentially multiple ways to calculate its market share, and my main measure is Glass Lewis clients' total assets divided by the sum of Glass Lewis and ISS's client assets.[40] Industry experts and academics have used this ratio to gauge competition in this industry (see Belinfanti, 2010). This measure is also similar to Becker and Milbourn's (2011) measure for credit rating agencies. Figure 1.1 shows evolution of Glass Lewis's client assets (in trillions of dollars) as well as its market share. Due to the fact that ISS's client assets have remained relatively stable over the years (between 23 and 25.5 trillion dollars), Glass Lewis's market share has closely resembled its client assets. As a robustness check, I also use alternative measures for market share based on the number of institutional shareholders, as well as coverage of U.S. companies. The results are similar, as shown in Tables B.1, B.2 and B.3.

[40] For any given client asset, there is some likelihood of overlap among proxy advisors since some clients use the services of several firms.

1.7.3 Results

Before conducting regression analysis, it is useful to visualize whether there is an overall decline of differing recommendations between ISS and Glass Lewis. Figure 1.2(A) shows that the average percent of differing recommendations (ISS "For", Glass Lewis "Against/Withhold") at the firm level has trended down since 2004. I do not include year 2003 in this analysis because Glass Lewis did not provide recommendations for individual directors in 2003. On the other hand, there is no clear pattern for the average percent of differing recommendations in the opposite direction (ISS "Against/Withhold", Glass Lewis "For"), as shown in Figure 1.2(B). This supports our theory that with increased competition, ISS has substantially lowered its frequency of "For" recommendations.

To see if there indeed is a general decline of differing recommendations for the same firms since 2004, I restrict my sample to firms that never exited the Russell 3000 family. This creates a balanced panel of 2,264 companies. Running regressions for the unbalanced panel (not reported), I find similar results. First, I regress the firm-level spread between ISS's and Glass Lewis's "For" recommendations on Glass Lewis's market share. As shown in Table 1.4, the difference between ISS's and Glass Lewis's "For" recommendations dropped by nearly 1.7 percentage points for a 10 percentage points increase in Glass Lewis's market share during 2004-2011. This suggests that with Glass Lewis's entry, ISS became more likely to switch from making "For" recommendations to "Against/Withhold" than from making "Against/Withhold" recommendations to "For".

As shown in Table 1.5, there is a significant correlation between Glass Lewis's market share and the fraction of differing recommendations (ISS "For," Glass Lewis "Against/Withhold") at the firm level. For a 10 percentage points increase in competition, this fraction of differing recommendations decreased by 6 percentage points. Note that better governance quality as measured in slower growth in executive compensation or higher ratio of independent directors contributes to less dispersion in recommendations. This is intuitive in that better governance quality, especially slower growth in compensation, reduces information asymmetry. It therefore will be more costly for a conflicted proxy advisor to issue biased recommendations, leading to a higher probability of identical recommendations from both advisory firms.

However, the fraction of differing recommendations of the opposite direction (ISS recommended "Against/Withhold," Glass Lewis recommended "For") barely dropped following Glass Lewis's entry. Columns (1)-(4) of Table 1.6 shows that for a 10 percentage points increase in competition, this fraction of differing recommendations decreased by only 1 percentage point. Overall, these results suggest that competition resulted in a less friendly posture of ISS towards corporations generally.

One might be concerned that the convergence of recommendations was attributed to changes in ISS's or Glass Lewis's proxy guidelines. I check published proxy guidelines by ISS and Glass Lewis for the period 2004-2011, and find no evidence of substantial changes regarding major types of management proposals. I also calculate ISS's and Glass Lewis's overall ratios of "For" for management proposals at the firm level, and find no evidence of any trend. The range is 83%-89% for ISS, and 70%-81% for Glass Lewis (see Figure 1.3). Some scholars point out that Proxy Governance was also a credible player before it exited the market in 2010. Proxy Governance client assets were around 1 trillion dollars for my sample period. I re-run regression (1.1) taking into account Proxy Governance's market share, and obtain similar results as in Tables 1.4-1.6 (not reported). Reverse causality is less of a concern because if investors expected that ISS was going to be more truthful (less likely to inflate the quality of management proposals), they would tend to subscribe to ISS more often. This will likely inflict a positive bias on my results. In other words, absent reverse causality, the magnitudes may be even larger. However, omitted variables at the firm level may bias the results. Next I resort to a plausibly exogenous shock – the event that Glass Lewis began to cover a stock for the first time – to analyze effects of competition on ISS's recommendations.

1.8 Impact of Glass Lewis's coverage

1.8.1 Estimation Strategy: Exploring an Exogenous Shock

When a proxy advisor obtains a new institutional client, by contract it must cover all portfolio firms of the client. One Glass Lewis executive has remarked: "When we get a new client, we make reports for all the firms in their portfolio." Prior to establishing the relationship, however, the advisory firm does not know which companies are in its prospective client's portfolio.[41] Thus the very fact that Glass Lewis began to cover a company for the first time served as an exogenous shock to ISS's recommendations. Glass Lewis's coverage provided investors an alternative source of information, and ISS might adjust its recommendations in the subsequent year for that company. To evaluate this effect, I regress the change in ISS's average "For" recommendation for company i from year $t-1$ to t on a dummy indicating Glass Lewis's new coverage in year $t-1$, and a large number of firm characteristics and fixed effects:

$$\Delta ISS\_AvgFor_{ijt} = \alpha + \delta \cdot GL\_Coverage_{ijt-1} + X'_{ijt}\beta + \eta_t + \mu_j \cdot t + \epsilon_{ijt} \tag{1.2}$$

In equation (1.2), time fixed-effects η_t control for economy-wide trends, and industry-specific time trends $\mu_j \cdot t$ account for differential linear trends that might affect ISS's recommendations. Standard errors are clustered at the firm level. The coefficient of interest δ, equivalent to a difference-in-differences estimator, measures impacts of increased competition on ISS's recommendations.

---

[41] Institutional investors managing assets over $100 million must report their holdings on Form 13F with the SEC on a quarterly basis. An advisory firm could access this information through SEC's website. However, Form 13F is allowed to be filed within 45 days of the end of a calendar quarter.

The effect of competition is expected to be larger if that company was ISS's corporate client, due to conflicts of interest shown in Section 1.2. To test this hypothesis, a list of ISS's corporate clients is needed. Were the corporate client data not proprietary, I can add to equation (1.2) an interaction term $GL\_Coverage_{ijt-1} \cdot Corp\_Client_{it}$, where $Corp\_Client_{it}$ is a dummy that equals 1 if company i is ISS's corporate client in year t. The coefficient on such an interaction term captures the effect of competition on ISS's recommendations for these client firms.

1.8.2 Validity of Estimation Strategy

Glass Lewis's coverage of a firm for the first time can serve as a credible exogenous shock to ISS's recommendation for that firm. However, one may worry that that company might respond to Glass Lewis's coverage, thus might have different characteristics from companies Glass Lewis already covered. This would likely bias the results. To check this, I regress firm level characteristics at t on the dummy $GL\_Coverage_{ijt-1}$ and three basic firm controls - size, return on assets and stock return. In Table 1.7, none of the coefficients are significant at the 5% level except institutional holdings. This is intuitive because the probability that Glass Lewis had already covered a company is lower if institutional holdings of that company were lower. After all, it is institutional investors who hired proxy advisors in the first place. I control institutional holdings in Tables 1.8 and 1.9 to eleminate such potential bias.

1.8.3 Results

Columns (1)-(4) of Table 1.8 present effects of Glass Lewis coverage using the entire sample. After Glass Lewis covered a company for the first time, ISS's average "For" recommendations decreased by 1.3 to 1.9 percentage points in the following year. This translates into a decrease in "For" recommendations by 1.9% to 2.3% given that ISS's average recommendation for management proposals was 84%. It is important to note that this is only an imperfect measure of the disciplinary effect of competition. Many Russell 3000 companies did not subscribe to ISS's consulting services. The effect of competition is expected to be larger for the set of ISS's corporate clients as conflict of interest would mainly arise from serving these firms. We will be able to test this prediction with a list of ISS's corporate clients.

Table 1.9 provides robustness checks by looking at whether Glass Lewis's coverage affects ISS's recommendations, regardless of whether the firm was already covered or not. Columns (1)-(4) show that ISS's "For" recommendations were around 4 percentage points lower when the firm was covered by Glass Lewis. This suggests that ISS did not only respond to Glass Lewis's initial coverage of a firm, it might still become tougher as Glass Lewis continued to cover it.

1.9 Discussion

This paper studies conflicts of interest arising from serving both shareholders and corporate issuers, and how competition among proxy advisors can alleviate these conflicts. As mentioned in Section 1.3, although this type of conflicts are widely considered the most damaging, there exist other types of potential conflicts. These include: (1) potential conflicts related to making recommendations on proposals sponsored by institutional clients; (2) potential conflicts when owners, directors or officers of proxy advisory firms serving on public company boards that have proposals on which the proxy advisors are making voting recommendations; (3) potential conflicts when the proxy advisors or their parent companies provide other services to clients. Both ISS and Glass Lewis have all of these conflicts. Although these types of conflicts are considered much smaller than the inherent conflicts arising from serving both investors and corporations, it will be interesting to analyze these types of conflicts in future studies.

Notice that this paper does not discuss conflicts of interest for shareholder proposals, which comprise about 10.3% of my sample, excluding director elections. Shareholder proposals can be divided into three categories: corporate governance, executive compensation and social policy. Each of these categories is interesting in its own right (see e.g., Cuñat, Gine and Guadalupe, 2012; Karpoff, Malatesta and Walkling, 1996; Randall and Cotter, 2007; Gillan and Starks, 2000; Agrawal 2008; Ertimur, Ferri and Stubben, 2010). However, these proposals are considered to be more complicated than management proposals. A shareholder proposing the measure could be a client of ISS, Glass Lewis or both. Without knowing which investors use either advisory firm's services, it would be difficult to analyze the potential conflicts.

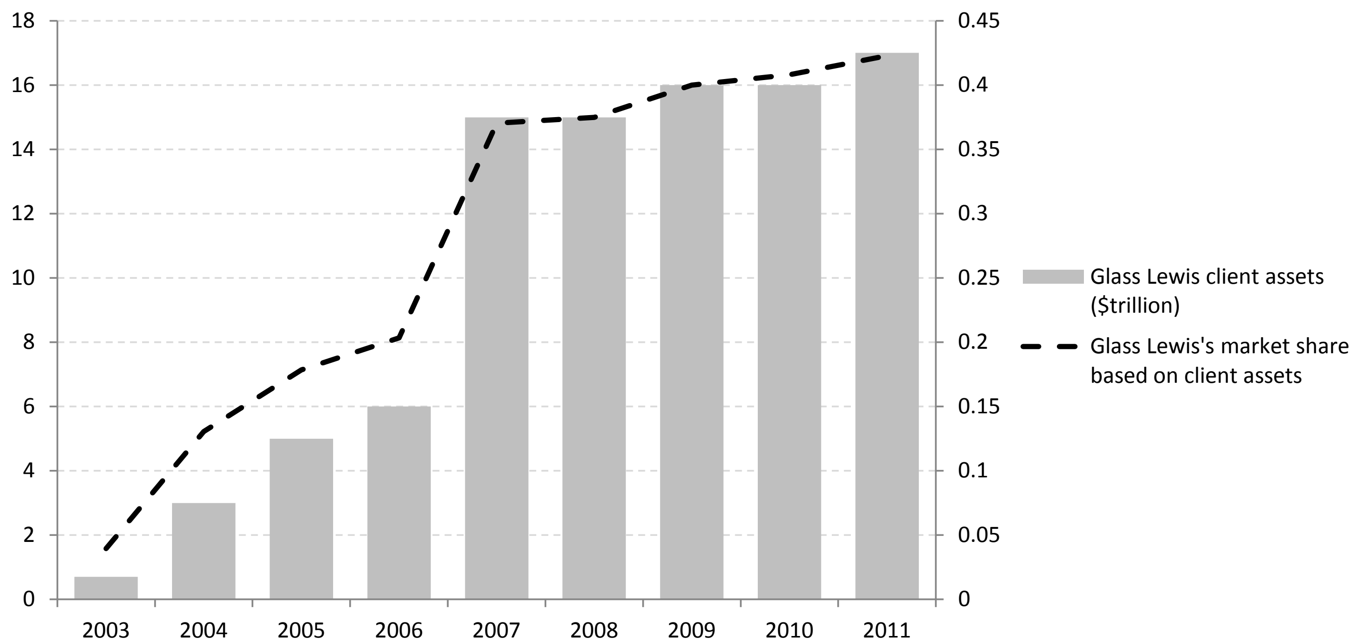
1.10 Conclusion

With ever growing institutional shareholdings and recent regulatory reforms to enhance shareholder rights, proxy advisory firms, ISS and Glass Lewis in particular, have become powerful in shaping corporate governance. Industry experts have long criticized potential conflicts of interest and a lack of competition in the business model. This paper is the first to document the fact that increased competition can alleviate ISS's potential conflicts arising from serving both investors and corporate issuers investors own. I show that ISS's and Glass Lewis's recommendations for management proposals at the firm level converged rapidly when Glass Lewis's market share grew for the period 2004-2011. This convergence was largely attributed to the fact that with Glass Lewis's entry, ISS became more likely to switch from making "For" recommendations to "Against/Withhold" than from making "Against/Withhold" recommendations to "For". Furthermore, ISS endorsed a company's proposals less frequently when Glass Lewis began to cover it for the first time. As expected, data suggest that actual vote outcomes were strongly correlated with recommendations from both proxy advisors, and Glass Lewis became more influential as it achieved higher market share.

Evidence supports the model prediction that conflicts of interest inevitably arise when a proxy advisor provides services to both shareholders and corporate issuers. Although increased competition can largely reduce the magnitude of these conflicts, competition itself may not be enough to completely eliminate them. The SEC is currently planning to issue an interpretative guidance to require proxy advisors to disclose "any significant relationship" with issuers or a shareholder proponent. This is an encouraging development. However, ultimately the SEC should ban proxy advisory firms from providing advisory services to institutional investors, while at the same time providing consulting services to corporate issuers on the matters of proxy votes. Also, proxy advisory firms should be required to fully disclose other conflicts.

It should be noted that while this paper supports the view that greater competition is desirable in the proxy advisory industry, the readiness of investors to support more than a few advisory firms remains unclear. An alternative solution is to promote a non-profit model for proxy advisors to eliminate conflicts of interest and to better serve the public interest. Other major issues in the industry include significant inaccuracies and a lack of transparency in decision-making. One example is that corporate issuers cannot access Glass Lewis's reports before they are published, increasing the chance of inaccuracies. These issues deserve careful study.

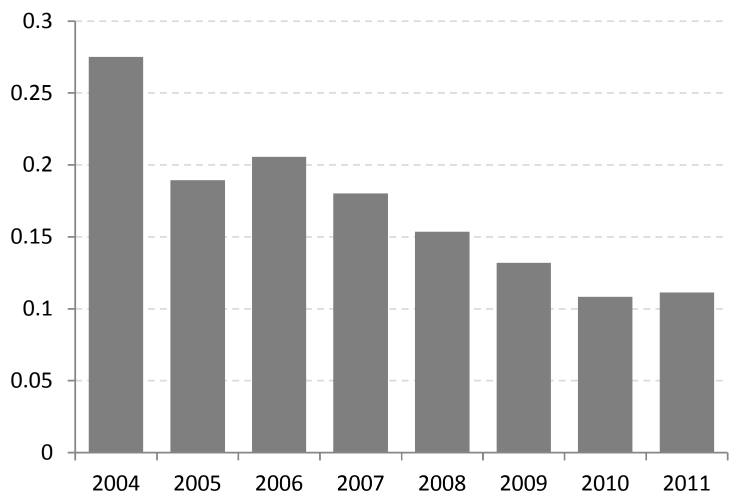
Figure 1.1: Evolution of Glass Lewis' Market Share



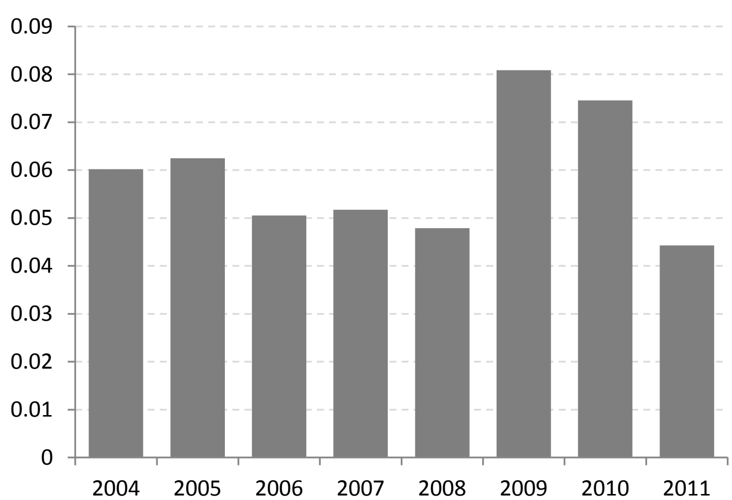
Notes: The solid bars (left axis) plot Glass Lewis client assets for the period 2003–2011. The dashed line (right axis) plots Glass Lewis' market share for the same period. It is calculated as below

$$\text{Glass Lewis' market share} = \frac{\text{Glass Lewis client assets}}{\text{ISS client assets} + \text{Glass Lewis client assets}}$$

Figure 1.2: Fraction of Differing Recommendations for Management Proposals Decreased



(A) Fraction of differing recommendations (ISS "For," Glass Lewis "Against/Withhold") decreased



(B) Fraction of differing recommendations (ISS "Against/Withhold," Glass Lewis "For") had no clear trend

Figure 1.3: Average Firm-Level "For" Recommendations for Management Proposals from ISS and Glass Lewis

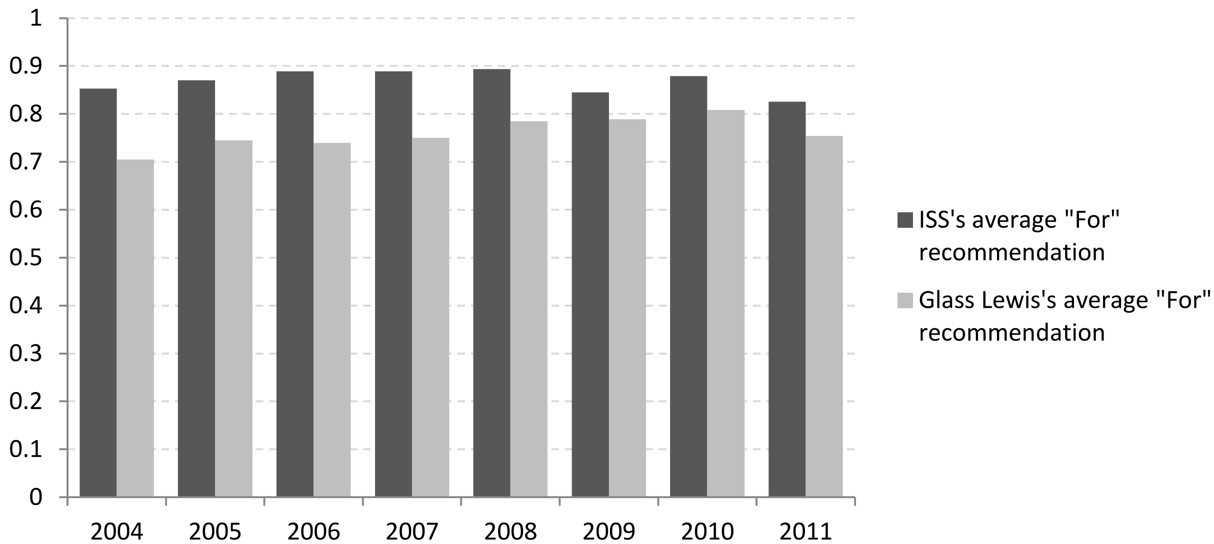


Table 1.1: Variable Definitions

Variable	Definition
Assets	Total assets in billions of dollars
Prior year industry-adjusted ROA	Earnings before interest, taxes, depreciation, and amortization divided by total assets at the end of the previous fiscal year. I adjust ROA by the industry median (all Compustat firm/year at 4-digit SIC level)
Prior year return	The 12 months buy-and-hold return prior to shareholder meeting
Book-to-market	The market value of equity divided by the book value of equity
Leverage	(Book value of debt -cash)/Total assets
Capex-to-assets	Capital expenditures less the sale of PP&E divided by total assets
Abnormal executive compensation ($million)	Residual from a compensation regression where the dependent variable is the total CEO compensation and the independent variable include log assets, prior-year stock return, industry and year dummies, estimated with all ExecuComp firms for 2004-2011
YOY change in executive compensation	Percentage change in total executive compensation year-on-year
Cash/total compensation	The ratio of salary and cash bonus to total compensation
Classified board	A Classified Board (or “staggered” board) is one in which the directors are placed into different classes and serve overlapping terms
Poison pill	It provides shareholders with special rights in the case of a triggering event such as a hostile takeover bid. Typical poison pills give the holders of the target’s stock other than the bidder the right to purchase stock in the target or the bidder’s company at a steep discount, making the target unattractive or diluting the acquirer’s voting power
Board size	The number of board members
Independent director	A director that has no material connection to the company other than a board seat
Compensation activism in past 3 years	Equals to 1 if there was a shareholder proposal targeting compensation practice in the past three years
Institutional holdings	Percent of outstanding shares held by intuitional investors
Management holdings	Percent of outstanding shares held by top-5 company executives
Confidential voting dummy	Equals one if firm policy prevents management from knowing how shareholders vote their proxy cards
Unequal voting dummy	Equals 1 if the firm has two or more classes of shares with unequal voting power, and 0 otherwise
Cumulative voting dummy	Equals 1 if the firm has a voting system whereby shareholders can cumulate votes for a single director candidate
Majority voting for directors dummy	Equals 1 if the firm’s directors are elected only if they receive more than 50% of the votes

Table 1.2: Summary Statistics

Panel A: Average "For" recommendations and votes per firm-year

	ISS "For"	Glass Lewis "For"	Average "For" Vote
Executive compensation plan	0.80	0.74	0.83
Say-on-pay proposal (2011 only)	0.88	0.79	0.91
Director election (firm level average)	0.88	0.77	0.95
All management proposals (firm level average)	0.85	0.70	0.93

Panel B: Firm characteristics, compensation and governance variables

	N	Mean	Median	SD
Firm characteristics				
Assets ($billions)	22,100	11.23	1.28	79.26
Prior year industry-adjusted ROA	19,243	0.03	0.02	0.28
Prior year stock return	22,347	0.16	0.09	0.75
Book-to-market	19,991	0.63	0.49	1.58
Leverage	22,035	0.56	0.55	0.34
Capex-to-assets	19,655	0.05	0.03	0.06
Compensation measures				
Abnormal executive compensation ($millions)	12,347	-0.31	-1.06	4.06
YOY change in executive compensation	12,663	0.40	0.10	5.59
Cash/total compensation	13,040	0.42	0.37	0.22
Governance measures				
Classified board & poison pill	11,265	0.29		
Board size	9,644	9.38	9.00	2.47
Ratio of independent directors	9,644	0.74	0.75	0.14
Institutional holdings	21,918	0.69	0.75	0.25
Management holdings	13,079	0.02	0	0.08
Voting mechanism				
Confidential voting dummy	11,265	0.13		
Cumulative voting dummy	11,265	0.08		
Unequal voting dummy	11,265	0.01		
Majority voting for directors dummy	24,954	0.10		

Notes: This table presents summary statistics for characteristics of management proposals, firm characteristics, and compensation and governance variables for Russell 3000 companies from 2004 to 2011. Details of the sample are discussed in Section 1.4. All variables are defined as in Table 1.1.

Table 1.3: Investors' Reactions to Voting Recommendations

	Dependent Variable: fraction of votes for					
	Executive compensation plan		Say-on-pay proposal (2011)		Director election (firm-level average)	
	(1)	(2)	(3)	(4)	(5)	(6)
ISS "For"	0.238***	0.238***	0.240***	0.238***	0.217***	0.217***
	(0.009)	(0.009)	(0.012)	(0.012)	(0.008)	(0.008)
Glass Lewis "For"	0.084***	0.083***	0.128***	0.126***	0.057***	0.058***
	(0.007)	(0.006)	(0.007)	(0.007)	(0.004)	(0.004)
ISS "For" * Glass Lewis "For"	0.034*	0.034*	-0.063**	-0.062**	-0.052*	-0.052*
	(0.019)	(0.020)	(0.025)	(0.025)	(0.030)	(0.034)
Firm characteristics		√		√		√
Compensation measures		√		√		√
Governance measures		√		√		√
Institutional and mgmt holdings		√		√		√
Firm FE	√	√			√	√
Industry-year trend	√	√			√	√
Observations	3,856	3,856	963	963	7,359	7,359
R-squared	0.71	0.71	0.71	0.72	0.77	0.77

Notes: The dependent variable is the fraction of favorable votes for management proposals. ISS (Glass Lewis) "For" equals 1 if ISS (Glass Lewis) recommends for a management proposal. All other variables are defined as in Table 1.1. Robust standard errors (clustered at the firm level) are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.4: Convergence of Recommendations

	Dependent Variable: fraction of ISS's "For" recommendations less fraction of Glass Lewis's "For" recommendations at firm level			
	(1)	(2)	(3)	(4)
Glass Lewis's market share (client assets)	-0.135***	-0.157***	-0.172***	-0.170***
	(0.048)	(0.055)	(0.065)	(0.066)
Log assets	-0.021**	-0.029***	-0.033**	-0.031**
	(0.008)	(0.011)	(0.015)	(0.015)
Ind-adj ROA	-0.022	-0.076*	-0.082	-0.065
	(0.018)	(0.046)	(0.066)	(0.067)
Prior-year stock return	0.002	0.002	0.011	0.013
	(0.004)	(0.006)	(0.008)	(0.009)
Leverage				0.072*
				(0.042)
Capex/Assets				0.213
				(0.136)
Abnormal executive compensation ($millions)		-0.0004	-0.001	-0.001
		(0.002)	(0.002)	(0.003)
ΔExecutive compensation YOY		-0.001***	-0.001***	-0.001***
		(0.0002)	(0.0002)	(0.0003)
Cash/total compensation		0.025	0.017	0.017
		(0.022)	(0.028)	(0.028)
Classified board*poison pill			-0.017	-0.016
			(0.014)	(0.014)
Board size			-0.0005	-0.0007
			(0.003)	(0.003)
Ratio of independent directors			-0.029	-0.043
			(0.043)	(0.043)
Institutional holdings				-0.012
				(0.046)
Management holdings				0.027
				(0.062)
Constant	0.269***	0.369***	0.446***	0.386***
	(0.060)	(0.088)	(0.126)	(0.132)
Firm FE	√	√	√	√
Fiscal-year FE	√	√	√	√
Industry-year trend	√	√	√	√
Observations	17,732	10,859	8,067	7,929
R-squared	0.39	0.33	0.33	0.33

Notes: The dependent variable is fraction of ISS's "For" recommendations minus fraction of Glass Lewis' "For" recommendations at the firm level. All independent variables are defined as in Table 1.1. Robust standard errors (clustered at the firm level) are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.5: Direction of Convergence in Recommendations

	Dependent Variable: fraction of differing recommendations at firm level (ISS "For," Glass Lewis "Against/Withhold")			
	(1)	(2)	(3)	(4)
Glass Lewis's market share (client assets)	-0.632***	-0.600***	-0.553***	-0.559***
	(0.032)	(0.041)	(0.053)	(0.054)
Log assets	-0.028***	-0.043***	-0.049***	-0.047***
	(0.008)	(0.012)	(0.014)	(0.014)
Ind-adj ROA	-0.020	-0.132***	-0.091	-0.042
	(0.019)	(0.050)	(0.066)	(0.066)
Prior-year stock return	-3.0e-4	-2.1e-05	-4.2e-4	-7.4e-4
	(0.004)	(0.005)	(0.008)	(0.008)
Leverage				0.107***
				(0.040)
Capex/Assets				-0.043
				(0.134)
Abnormal executive compensation ($millions)		0.001	0.002	0.002
		(0.001)	(0.002)	(0.002)
ΔExecutive compensation YOY		0.001***	0.001***	0.001***
		(0.0002)	(0.0002)	(0.0002)
Cash/total compensation		-0.001	-0.006	-0.004
		(0.019)	(0.024)	(0.024)
Classified board*poison pill			-0.018	-0.019
			(0.013)	(0.013)
Board size			0.001	0.001
			(0.003)	(0.003)
Ratio of independent directors			-0.084**	-0.090**
			(0.042)	(0.041)
Institutional holdings				-0.063
				(0.052)
Management holdings				0.041
				(0.055)
Constant	0.610***	0.731***	0.783***	0.758***
	(0.055)	(0.086)	(0.110)	(0.120)
Firm FE	√	√	√	√
Fiscal-year FE	√	√	√	√
Industry-year trend	√	√	√	√
Observations	17,840	10,778	7,702	7,581
R-squared	0.48	0.45	0.39	0.37

Notes: The dependent variable is fraction of differing recommendations for management proposals at the firm level (ISS "For," Glass Lewis "Against/Withhold"). All independent variables are defined as in Table 1.1. Robust standard errors (clustered at the firm level) are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.6: Direction of Convergence in Recommendations

	Dependent Variable: fraction of differing recommendations at firm level (ISS "Against/Withhold," Glass Lewis "For")			
	(1)	(2)	(3)	(4)
Glass Lewis's market share (client assets)	-0.110***	-0.131***	-0.123***	-0.124***
	(0.023)	(0.025)	(0.030)	(0.031)
Log assets	-0.001	-0.004	-0.001	-0.002
	(0.004)	(0.005)	(0.007)	(0.008)
Ind-adj ROA	0.005	0.012	0.036	0.051
	(0.010)	(0.026)	(0.034)	(0.036)
Prior-year stock return	-0.003	-0.003	-0.008*	-0.008**
	(0.002)	(0.003)	(0.004)	(0.004)
Leverage				0.015
				(0.021)
Capex/Assets				-0.165**
				(0.073)
Abnormal executive compensation ($millions)		-0.001	-0.001	-0.001
		(0.001)	(0.001)	(0.001)
ΔExecutive compensation YOY		0.001***	0.001***	0.001**
		(0.0002)	(0.0002)	(0.0002)
Cash/total compensation		-0.009	-0.010	-0.011
		(0.011)	(0.013)	(0.013)
Classified board*poison pill			-0.007	-0.007
			(0.007)	(0.007)
Board size			0.001	0.001
			(0.001)	(0.001)
Ratio of independent directors			-0.031	-0.030
			(0.021)	(0.021)
Institutional holdings				0.003
				(0.024)
Management holdings				-0.004
				(0.033)
Constant	0.075**	0.050	0.104*	0.107*
	(0.029)	(0.041)	(0.058)	(0.062)
Firm FE	√	√	√	√
Fiscal-year FE	√	√	√	√
Industry-year trend	√	√	√	√
Observations	17,732	10,859	8,067	7,929
R-squared	0.46	0.37	0.36	0.37

Notes: The dependent variable is fraction of differing recommendations for management proposals at the firm level (ISS "Against/Withhold," Glass Lewis "For"). All independent variables are defined as in Table 1.1. Robust standard errors (clustered at the firm level) are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.7: Correlations Between Firm-Level Characteristics and Glass Lewis' Coverage

	(1) Abnormal executive compensation	(2) Δexec comp. YOY	(3) Cash/total comp.	(4) Classified board* poison pill	(5) Board size	(6) % indep. directors	(7) Inst. holdings	(8) Mgmt holdings	(9) Leverage	(10) Capex
I{Glass Lewis began coverage at t-1}	0.121	0.059	0.012	-0.002	0.082	0.006	-0.033***	0.003	0.006	-0.002
	(0.108)	(0.063)	(0.008)	(0.011)	(0.069)	(0.006)	(0.006)	(0.003)	(0.005)	(0.001)
Log assets	0.967***	-0.0001	-0.042***	0.025	0.437***	-0.001	0.063***	-0.007***	-0.037***	-0.004***
	(0.165)	(0.154)	(0.008)	(0.018)	(0.085)	(0.006)	(0.006)	(0.002)	(0.011)	(0.001)
Ind-adj ROA	2.056***	1.073	-0.122***	-0.113*	0.0726	-0.001	0.014	-0.017*	-0.182***	-0.002
	(0.457)	(0.678)	(0.038)	(0.063)	(0.281)	(0.024)	(0.014)	(0.009)	(0.050)	(0.005)
Stock return	0.284***	0.167**	-0.021***	3.2e-05	-0.104***	0.002	0.005*	-0.001	-0.004	0.002***
	(0.037)	(0.071)	(0.004)	(0.006)	(0.033)	(0.003)	(0.002)	(0.001)	(0.003)	(0.001)
Constant	-4.923***	0.083	0.678***	0.206	5.730***	0.704***	0.159***	0.073***	0.883***	0.068***
	(1.290)	(1.266)	(0.061)	(0.138)	(0.659)	(0.049)	(0.036)	(0.018)	(0.072)	(0.017)
Firm FE	√	√	√	√	√	√	√	√	√	√
Year FE	√	√	√	√	√	√	√	√	√	√
Observations	12,784	12,432	12,782	11,054	9,511	9,511	19,364	12,817	20,566	20,464
R-squared	0.74	0.22	0.65	0.81	0.87	0.74	0.84	0.41	0.85	0.80

Notes: All dependent variables are defined as in Table 1.1. Robust standard errors (clustered at the firm level) are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. The values of dependent variables are taken at time t.

Table 1.8: Impact of Glass Lewis' Coverage on ISS's Recommendations

	Dependent Variable: Change in ISS's "For" recommendations at the firm level from *t-1 to t*			
	(1)	(2)	(3)	(4)
I{Glass Lewis started coverage at t-1}	-0.023**	-0.026**	-0.018*	-0.019*
	(0.010)	(0.011)	(0.010)	(0.010)
Log assets		0.033	0.044	0.048
		(0.021)	(0.028)	(0.029)
Ind-adj ROA		0.057	0.049	0.047
		(0.055)	(0.212)	(0.217)
Prior-year stock return		0.013	0.010	0.012
		(0.013)	(0.020)	(0.020)
Leverage				-0.066
				(0.086)
Capex/Assets				-0.143
				(0.410)
Abnormal executive compensation ($millions)			-0.006	-0.008
			(0.006)	(0.006)
ΔExecutive compensation YOY			-0.006***	-0.006***
			(0.001)	(0.001)
Cash/total compensation			-0.057	-0.066
			(0.063)	(0.067)
Classified board*poison pill			0.040*	0.038
			(0.024)	(0.025)
Board size			-0.006	-0.007
			(0.007)	(0.007)
Ratio of independent directors			0.276**	0.289**
			(0.116)	(0.118)
Institutional holdings				-0.078
				(0.091)
Management holdings				-0.032
				(0.124)
Constant	0.093	0.078	-0.045	0.063
	(0.071)	(0.151)	(0.249)	(0.257)
Fiscal-year FE	√	√	√	√
Industry-year trend	√	√	√	√
Observations	21,230	17,197	7,827	7,701
R-squared	0.15	0.16	0.19	0.19

Notes: The dependent variable is the change in ISS's "For" recommendations for management proposals at the firm level. All independent variables are defined as in Table 1.1. Robust standard errors (clustered at the firm level) are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.9: Persistent Effects of Glass Lewis' Coverage on ISS's Recommendations

	Dependent Variable: fraction of ISS's "For" recommendations at the firm level at *t*			
	(1)	(2)	(3)	(4)
I{Glass Lewis covered at *t*}	-0.043**	-0.045**	-0.034*	-0.041*
	(0.020)	(0.022)	(0.021)	(0.022)
Log assets		0.007	0.011	0.009
		(0.006)	(0.010)	(0.011)
Ind-adj ROA		0.047	0.033	0.042
		(0.043)	(0.045)	(0.047)
Prior-year stock return		0.002	0.013**	0.014**
		(0.003)	(0.006)	(0.006)
Leverage				-0.016
				(0.030)
Capex/Assets				0.108
				(0.093)
Abnormal executive compensation ($millions)			-0.001	-0.001
			(0.002)	(0.002)
ΔExecutive compensation YOY			-0.001***	-0.001***
			(0.0002)	(0.0002)
Cash/total compensation			-0.020	-0.022
			(0.018)	(0.018)
Classified board*poison pill			0.015	0.016
			(0.010)	(0.010)
Board size			-0.002	-0.001
			(0.002)	(0.002)
Ratio of independent directors			0.084***	0.081***
			(0.028)	(0.028)
Institutional holdings				0.011
				(0.030)
Management holdings				0.028
				(0.048)
Constant	0.782***	0.745***	0.843***	0.815***
	(0.098)	(0.106)	(0.094)	(0.101)
Firm FE	√	√	√	√
Fiscal-year FE	√	√	√	√
Industry-year trend	√	√	√	√
Observations	26,301	20,634	8,927	8,715
R-squared	0.43	0.45	0.40	0.40

Notes: The dependent variable is fraction of ISS's "For" recommendations for management proposals at the firm level. All independent variables are defined as in Table I. Robust standard errors (clustered at the firm level) are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

1.11 References

Agrawal, A. K. (2008), Corporate Governance Objectives of Labor Union Shareholders: Evidence from Proxy Voting, New York University, Stern School of Business Working Paper FIN-08-006.

Alexander, C., M. Chen, D. Seppi and C. Spatt (2010), Interim News and the Role of Proxy Voting Advice, *Review of Financial Studies* 23, 4419-4454.

Altonji, J., T. Elder and C. Taber (2005), Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools, *Journal of Political Economy* 113, 151-184.

Bebchuk, L. A. and A. Cohen (2005), The Costs of Entrenched Boards, *Journal of Financial Economics* 78, 409-433.

Becker, B. and T. Milbourn (2011), How Did Increased Competition Affect Credit Ratings? *Journal of Financial Economics* 101(3), 493-514.

Belinfanti, T. C. (2009), The Proxy Advisory & Corporate Governance Industry: The Case for Increased Oversight and Control, New York Law School Legal Studies Research Paper, No. 09/10 #18.

Bethel, J. and S. Gillan (2002), The Impact of the Institutional and Regulatory Environment on Shareholder Voting, *Financial Management* 31, 29-54.

Bew, R. and R. Fields (2012), Voting Decisions at US Mutual Funds: How Investors Really Use Proxy Advisers, The IRRC Institute and Tapestry Networks, Inc.

Biglaiser, G. (1993), Middlemen as Experts, *RAND Journal of Economics* 24, 212-223.

Bolton, P., X. Freixas and J. Shapiro (2012), The Credit Ratings Game, *Journal of Finance* 67, 85-112.

Brav, A., W. Jiang, F. Partnoy and R. Thomas (2008), Hedge Fund Activism, Corporate Governance, and Firm Performance, *Journal of Finance* 63(4), 1729-1775.

Brickley, J., R. Lease and C. Smith (1988), Ownership Structure and Voting on Antitakeover Amendments, *Journal of Financial Economics* 20, 267-291.

Brickley, J., R. Lease and C. Smith (1994), Corporate Voting: Evidence from Charter Amendment Proposals, *Journal of Corporate Finance* 1, 5-31.

Cai, J., J. L. Garner and R. A. Walkling (1999), Electing Directors, *Journal of Finance* 64(5), 2389-2421.

Camanho, N., P. Deb and Z. Liu (2012), Credit Rating and Competition, mimeo, London School of Economics and Political Science.

Canadian Securities Administrators (2012), Potential Regulation of Proxy Advisory Firms, Consultation Paper, 25-401.

Center on Executive Compensation (2011), A Call for Change in the Proxy Advisory Industry Status Quo: The Case for Greater Accountability and Oversight? January 2011.

Choi, S., J. Fisch and M. Kahan (2009), Director Elections and the Role of Proxy Advisors, *Southern California Law Review* 82, 649.

Choi, S., J. Fisch and M. Kahan (2010), The Power of Proxy Advisors: Myth or Reality? *Emory Law Journal* 59, 101-151.

Cremers, M. and R. Romano (2007), Institutional Investors and Proxy Voting: The Impact of the 2003 Mutual Fund Voting Disclosure Regulation, Yale Law and Economics Research Paper, No. 349.

Cuñat, V., M. Gine and M. Guadalupe (2012), The Vote is Cast: The E?ect of Corporate Governance on Shareholder Value, *Journal of Finance* 67(5), 1943-1977.

Daines, R., I. Gow, and D. Larcker (2010), Ratings the Ratings: How Good are Commercial Governance Ratings? *Journal of Financial Economics* 98(3), 439-461.

Davis, F. G. and E. H. Kim (2012), Business Ties and Proxy Voting by Mutual Funds, *Journal of Financial Economics* 85, 552-570.

Dechow, P., A. Hutton, and R. G. Sloan (1999), The Relation between Affiliated Analysts' Long-Term Earnings Forecasts and Stock Price Performance Following Equity Offerings, *Contemporary Accounting Research* 17, 1-32.

Dehejia, R. and S. Wahba (1999), Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs, *Journal of the American Statistical Association* 94(488), 1053-1062.

Dugar, A. and S. Nathan (1995), The Effect of Investment Banking Relationships on Financial Analysts' Earnings Forecasts and Investment Recommendations, *Contemporary Accounting Research* 12, 131-160.

Ertimur, Y., F. Ferri and S. Stubben (2010), Board of Directors' Responsiveness to Shareholders: Evidence from Shareholder Proposals, *Journal of Corporate Finance* 16, 53-72.

Ertimur, Y., F. Ferri and V. Muslu (2011), Shareholder Activism and CEO Pay, *Review of Financial Studies* 24, 535-592.

Ertimur, Y., F. Ferri and D. Maber (2012), Reputation Penalties for Poor Monitoring of Executive Pay: Evidence from Option Backdating, *Journal of Financial Economics* 104, 118-144.

Ertimur, Y., F. Ferri and D. Oesch (2012), Shareholder Votes and Proxy Advisors: Evidence from Say on Pay, mimeo, Columbia University.

European Securities Markets Authority (2012), An Overview of the Proxy Advisory Industry. Considerations on Possible Policy Options, European Securities Markets Authority

Discussion Paper, March 2012.

Fos, V. (2011), The Disciplinary Effects of Proxy Contests, mimeo, University of Illinois at Urbana-Champaign, August 2011.

Gillan, S. and L. Starks (2000), Corporate Governance Proposals and Shareholder Activism: The Role of Institutional Investors, *Journal of Financial Economics* 57, 275-305.

Gompers, P. A., J. L. Ishii and A. Metrick (2003), Corporate governance and equity prices, *Quarterly Journal of Economics* 118, 107-155.

Gordon, L. and J. Pound (1993), Information, Ownership Structure, and Shareholder Voting: Evidence from Shareholder-sponsored Corporate Governance Proposals, *Journal of Finance* 48, 697-718.

Government Accountability Office (2007), Corporate Shareholder Meetings: Issues Relating to Firms That Advise Institutional Investors on Proxy Voting, June 29.

Griffin, J. and D. Y. Tang (2012), Did Subjectivity Play a Role in CDO Credit Ratings? *Journal of Finance* 67, 1293-1328.

Hartzell, J., E. Ofek, and D. Yermack (2004), What's In It For Me? CEOs Whose Firms Are Acquired, *Review of Financial Studies* 17, 37-61.

Heckman, J., R. LaLonde and J. Smith (1999), The economics and econometrics of active labor market programs. In: Ashenfelter, O., Card, D. (Eds.), *Handbook of Labor Economics*, Vol. 3A. NorthHolland, Amsterdam, 1865-2097.

Hong, H. and M. Kacperczyk (2010), Competition and Bias, *Quarterly Journal of Economics* 125 (4), 1683-1725.

Hörner, J. (2002), Reputation and Competition, *American Economic Review* 92 (3), 644-663.

Imbens, G. (2004), Nonparametric estimation of average treatment effects under exogeneity: A review, *Review of Economics and Statistics* 86(1), 4-29.

Karpoff, J., P. Malatesta and R. Walkling (1996), Corporate Governance and Shareholder Initiatives: Empirical Evidence, *Journal of Financial Economics* 42, 365-395.

Klein, A. and E. Zur (2009), Entrepreneurial shareholder activism: Hedge funds and other private investors, *Journal of Finance* 64(1), 187-229.

Larcker, D. , A. McCall and G. Ormazabal (2012), Proxy Advisory Firms and Stock Option Exchanges, mimeo, Stanford University.

Lin, H. and M. F. McNichols (1998), Underwriting Relationships, Analysts' Earnings Forecasts and Investment Recommendations, *Journal of Accounting and Economics* 25, 101-127.

Lizzeri, A. (1999), Information Revelation and Certification Intermediaries, *RAND Journal of Economics* 30, 214-231.

Mathis, J., J. McAndrews and J.C. Rochet (2009), Rating the Raters: Are Reputational Concerns Powerful Enough to Discipline Rating Agencies?*Journal of Monetary Economics* 56(5), 675-677.

Matvos, G. and M. Ostrovsky (2010), Heterogeneity and peer effects in mutual fund proxy voting, *Journal of Financial Economics* 98, 90-112.

Maug, E. and K. Rydqvist (2009), Do Shareholders Vote Strategically? Voting Behavior, Proposal Screening, and Majority Rules, *Review of Finance* 13, 47-79.

Michaely, R. and K. L. Womack (1999), Conflict of Interest and the Credibility of Underwriter Analyst Recommendations, *Review of Financial Studies* 12, 653-686.

Morgan, A. and A. Poulsen (2001), Linking Pay to Performance-Compensation Proposals in the S&P 500, *Journal of Financial Economics* 62, 489-523.

Morgan, J. and P. C. Stocken (2003), An analysis of stock recommendations, *RAND Journal of Economics* 34, 183-203.

Morgenson, G. (2004), How to Succeed on Wall Street, Conflict-Free, The New York Times, December 19.

Pound, J. (1988), Proxy Contests and the Efficiency of Shareholder Oversight, *Journal of Financial Economics* 20, 237-265.

Rabimov, S. and M. Tonello (2010), The 2010 Institutional Investment Report: Trends in Asset Allocation and Portfolio Composition, The Conference Board Research Report, No. R-1468-10-RR.

Rose, F. (1999), The economics, concept, and design of information intermediaries: A theoretic approach, Physica-Verlag, Heidelberg, Germany.

Rosenbaum, P. and D. Rubin (1983), The central role of the propensity score in observational studies for causal effects, *Biometrika* 70(1), 41-55.

Sangiorgi, F., J. Sokobin and C. Spatt (2009), Credit-Rating Shopping, Selection and Equilibrium Structure of Ratings, mimeo, Carnegie Mellon University.

Securities and Exchange Commission (2010), Concept Release, No. 34-62495, July 14.

Skreta, V. and L. Veldkamp (2009), Ratings Shopping and Asset Complexity: A Theory of Ratings Inflation, *Journal of Monetary Economics* 56(5), 678-695.

Smith, J. and P. Todd (2005), Does matching overcome Lalonde's critique of nonexperimental estimators, *Journal of Econometrics* 125, 305-353.

Strobl, G.and H. Xia (2012), The Issuer-Pays Rating Model and Ratings Inflation: Evidence from Corporate Credit Ratings, mimeo, University of North Carolina at Chapel Hill.

Stroebel, J. (2012), The Impact of Asymmetric Information about Collateral Values in Mortgage Lending, mimeo, Stanford University.

Thomas, R. and J. F. Cotter (2007), Shareholder Proposals in the New Millennium: Shareholder Support, Board Response and Market Reaction, *Journal of Corporate Finance* 13, 368-391.

Verdam, A. (2007), An Exploration of the Role of Proxy Advisors in Proxy Voting, mimeo, University of Amsterdam.

Walkling, R. and M. Long (1984), Agency Theory, Managerial Welfare and Takeover Bid Resistance, *RAND Journal of Economics* 15, 54-68.

White, L. (2002), The credit rating industry: An industrial organization analysis, in Richard M. Levich, Giovanni Majnoni, and Carmen Reinhart, eds.: *Ratings, Rating Agencies and the Global Financial System* (Kluwer Academic Publishers, Boston).

Chapter 2

The Value of Access to Public Transportation in a Congested City – Evidence from Housing Prices in Beijing

2.1 Introduction

Modern urban economic theory predicts a trade-off between property values and commuting costs. This can be illustrated by the classic monocentric city model (Alonso, 1964; Muth 1969; Mills, 1967) in which all employment is concentrated in the Central Business District (CBD), and residents commute to the CBD through radial roads with no congestion. The fact that residents are identical implies an equal utility at all locations since any resident can replicate the location and consumption bundle of anyone else. To achieve the same utility with an increase in the distance to the CBD, and thus with a greater commuting cost (opportunity cost in lost wages), the rental price must fall to reach the same consumption level.

This stylized model typically assumes constant travel speed. However, in fast-growing megacities like Beijing, Mumbai and Shanghai, severe traffic congestion causes substantial delay time for commuters, lowering consumer welfare. As one of the most important types of negative externality in cities, congestion costs are reflected in lower housing prices (Solow and Vickrey, 1971; Fujita, 1989). Governments in large cities invest heavily in public transportation in an effort to minimize traffic congestion.[1] However, not all public transit modes are created equal. Although an increase in bus routes would better connect residential areas and the workplace, buses themselves often cause traffic delays, potentially decreasing consumer welfare. On the other hand, metro lines do not compete with vehicles for space, and could be more effective in reducing traffic congestion. Therefore, aggregate quantity of public transit services does not necessarily increase consumer welfare.

I begin the study of the impact of public transportation by analyzing a monocentric model. Before introducing a metro line, the model features a rent "gradient" between the CBD and urban fringe. Because of increased commuting costs, housing prices decrease with the distance from the CBD. With traffic congestion, the model implies a negative association between congestion and housing prices. The introduction of a metro line reduces commuting costs along the line, resulting in an increase in residential rents in the proximity of the line. The differential jump is more pronounced for areas that originally had more severe congestion. Besides reducing negative effects of congestion, the metro line also causes disposable income to increase and the urban fringe to expand.

These predictions are tested in the context of Beijing, a rapidly-growing megacity that has had severe traffic congestion. I find that road congestion has a large negative impact on housing prices in Beijing. A half-hour increase in daily delay time[2] translates into a 3.8% decrease in housing prices. The order of magnitude appears to be large, compared to that

[1] According to Beijing Municipal Commission of Transportation, governments in Beijing, the largest city in China, spent 5.3% of its GDP on improving its transportation infrastructure annually between 1993 and 2003.

[2] Daily delay hours = (Daily morning travel hours – One-way travel hours at the speed limit) $\times 2$. I assume workers drive to work.

in Brounen, Neuteboom and Xu (2010). It is important to note that I make an assumption that all workers drive to work, so that to the extent that many people in Beijing still take buses to work, the variable – travel delay time – may be underestimated.[3] The instrumental variable estimation yields consistent but somewhat larger magnitudes, suggesting that reverse causality could be present in the baseline results.

There are multiple approaches that cities could use to combat traffic congestion. These include developing public transit systems, expanding road capacity as well as charging peak-hour tolls or restricting vehicle usage. Although the Beijing Municipal Government has used all of the three ways to alleviate congestion, it has relied the most on developing an expansive metro system that serves the entire city. The second part of the paper examines whether the presence of public transportation has indeed mitigated the adverse effect of traffic congestion. I find that the negative effect of congestion on housing prices is smaller when there are metro stations within walking distance. I further show that construction announcements of new metro lines have a strong effect on prices of properties in the proximity of the stations. The difference-in-differences coefficient of 3.3% translates to a difference of 560 yuan per square meter ($8.0 per square foot) within a six-month window. Furthermore, the effect is more pronounced in more congested places, suggesting the metro system is expected to alleviate the negative impact of traffic congestion. However, additional bus routes are not capitalized into prices since buses move slowly in the gridlocked city, often exacerbating rather than alleviating congestion.

The literature on anticipated effects of public infrastructure in developed countries gives mixed results. Several papers find significant capitalization effects (see McMillen and McDonald, 2004; Damm et al., 1980; McDonald and Osuji, 1995), while Gatzlaff and Smith

---

[3]By assuming that every worker rides the bus to work, and that bus rides double the delay time I calculate, the effect is still economically significant – half an hour delay per day translates into a 1.9% decrease in housing prices.

(1993) find no announcement effect of a new train system in Miami. The positive magnitudes found in these studies are typically smaller than those in this paper. For example, McMillen and McDonald (2004) show that within 1.5 years following the announcement of Chicago's Midway Line, houses closer to the line were 4.2% higher than those further away. My study yields a 3.3% difference within 6 months following the announcement, a higher magnitude.[4] This more significant result found in Beijing may be due to lower car ownership as well as higher density (see Zheng and Kahn, 2008).

This is one of the first papers that use micro-level congestion data to assess property values. This enables me to calculate delay time at the property level, utilizing a series of city-wide speed maps for the period 2005-2011. Existing empirical studies in urban economics involving commuting time usually gather survey data either at the city level (e.g., Coulson and Engle, 1987; Gordon, Kumar and Richardson, 1991) or at the metropolitan statistical area (MSA) level (e.g., Glaeser, Kahn and Rappaport, 2000).

This paper also contributes to the policy debate regarding efficient mass transit systems in fast-growing megacities. In these congested cities, workers seem to attach different premiums to different public transit modes. My estimates confirm that housing prices are sensitive to new metro lines but not to additional bus routes. The reason may be that in extremely congested cities, additional bus routes exacerbate rather than alleviate congestion. Although further cost-benefit analysis needs to be applied to both transit modes (e.g., Winston, 2007), it is reasonable to argue that the metro system may be more desirable in megacities. Note that this paper does not discuss direct measures that could curb traffic congestion, such as congestion pricing and highway investment (see e.g., Vickrey, 1969 and Leape, 2006 for a discussion). At of late 2012, the Beijing Municipal Government was discussing a possible congestion charge in the urban area, and I plan to analyze it in a follow-up study.

[4] I also show that the difference amounts to 10.4% within 2.5 years after the announcement.

2.1.1 Related Literature

Although there exists a large theoretical literature on traffic congestion and congestion pricing (e.g., Vickrey, 1969; Arnott, de Palma and Lindsey, 1993), the empirical literature in this area remains less developed. A growing body of policy papers have attempted to evaluate the welfare costs of congestion by measuring the direct monetary costs (e.g., Mun and Yonekawa, 2006; Hartgen and Fields, 2009). The Foton Motor Report (2010) find that economic costs of congestion in Beijing are 4,027 yuan per capita in 2009 ($631), which is somewhat larger than my findings.[5] Note that these papers could only measure current cost of congestion, while my paper does a reasonable job in calculating all future costs as reflected in differential property prices.

Another strand of literature looks at the relationship between traffic congestion and land use patterns (e.g., Solow and Vickrey, 1971; Fujita, 1989; Arnott and MacKinnon, 1978). It is important to note that a decrease in traffic congestion does not necessarily lead to a reduction in housing price, since traffic congestion is often temporary (most severe during rush hours) and difficult to predict. A few papers have explored the relationship between traffic congestion and housing prices, either at the city level or district level. Coulson and Engle (1987) find that transport costs, especially time and gasoline costs, were capitalized into housing prices across major U.S. cities. Using Dutch traffic data, Brounen, Neuteboom and Xu (2010) find a negative relationship between traffic congestion and house prices at the district level in the Netherlands. Zhang and Shing (2006) compare house prices inside and outside London's congestion-charge zone, and find that the gap in house prices reduced after London imposed the charge. Although other factors may potentially affect the price gap, their paper provides indirect evidence that traffic congestion and house prices may be

[5] Beijing's population in 2009 was 17.55 million. This implies that the annual monetary cost of congestion was around 70.7 billion yuan, which amounted to 4.4% of Beijing's Gross Metropolitan Product in 2011.

negatively correlated.

This paper also relates to the vast empirical literature that studies the effects of improved accessibility, such as new transit systems, on house values (see Vessali, 1996 for a comprehensive review). Smersh and Smith (2000) and Agostini and Palmucci (2008) study price impacts of a new bridge in the U.S. and a metro line in the Chile, respectively. They find that the completion of these infrastructure projects had a positive effect on local house prices. At the macro level, various studies have shown a linkage between infrastructure and GDP or population growth. Atack, Bateman, Haines and Margo (2010) examine the effect of railroads on urbanization and population growth in the U.S. Banerjee, Duflo and Qian (2009) estimate the effect of railroad networks on regional growth in China. Both of these papers find a positive relationship between regional growth and the distance to major railroads. My paper also relates to the literature that studies the effect of infrastructure on trade costs and its consequences (see Michaels, 2008; Donaldson, 2010).

Other authors have evaluated the costs and benefits of infrastructure investments. Duranton and Turner (2011) conclude that increased provision of roads or public transit is unlikely to relieve congestion, which is consistent with Winston (2006) who finds that investment in highways is ineffective in alleviating congestion. Winston (2007) further shows that investment in rapid transit lines in the U.S. is not socially efficient due to low ridership. This paper does not discuss agglomeration economies, such as how traffic congestion affects productivity in cities (e.g., Mun and Yonekawa, 2006; Weisbrod, Vary and Treyz, 2007; Hartgen and Fields, 2009)

The rest of the paper is organized as follows. In Section 2.2, I presents a conceptual framework to motivate the empirical analysis. Section 2.3 provides a history of Beijing's urban form, housing market, as well as its transportation system. Section 2.4 describes the various data sources and presents summary statistics. In Section 2.5, I estimate the impact of road congestion on housing prices. Section 2.6 estimates the announcement effect of new metro lines on housing prices. Section 2.7 offers further discussions, and Section 2.8 concludes. References are in Section 1.9.

2.2 Conceptual Framework

To inform the empirical estimation I provide a simple model of transportation infrastructure and urban growth in which a city can grow by drawing population from rural areas. In the spirit of Alonso (1964), Mills (1967), and Muth (1969), I consider an open city version of the monocentric model in which residents commute to a central business district (CBD) to work. To ease exposition, I deliberately abstract from a number of important details: agglomeration effects (endogenous wages), the production of housing and endogenous lot size. These features can be incorporated into the basic monocentric city model described here at some cost in complexity.

All residents receive a wage w by working at the CBD. All city residents occupy one unit of land at endogenously determined rental price $R(x)$ where x is the distance to the CBD. Residents spend their income, net of rent and commuting, on a numéraire good c. To form the baseline scenario, I assume that the city initially has no mass transit system, and residents drive[6] to the CBD along linear "rays" that assemble surface roads.[7] The commuting cost is b per unit of distance. b reflects the opportunity cost in lost wages, and is inversely proportional to the travel speed.

Residents derive utility from consumption according to an increasing concave utility function $U(c)$. Because all individuals are identical, in equilibrium everybody has the same

[6]To keep the model tractable, I do not consider commuters who take the bus. See Baum-Snow (2005) for more details.

[7]For simplicity I do not consider heterogeneous travel speeds on roads, such as those of surface streets, expressways and highways. In Beijing, most expressways and highways are equally congested as surface streets during peak hours.

utility and cannot gain higher utility by moving to an alternate location. Let the reservation utility be $\underline{u}$. With free migration, we have

$$U(w - bx - R(x)) = \underline{u}$$

Inverting $U(\cdot)$, we have $w - bx - R(x) = k$ for $k = U^{-1}(\underline{u})$. Thus, the rent for any occupied location in the city is

$$R(x) = w - bx - k \tag{2.1}$$

It is straightforward to observe that land rent decreases in the distance to the CBD, due to increased transportation cost. As in Duranton and Turner (2012), I assume that land outside of the city is employed in agriculture and generates reservation land rent $\underline{r}$. Thus, at the boundary of the city $\overline{x}$, we must have $\underline{r} = w - b\overline{x} - k$. This condition determines the city boundary

$$\overline{x} = \frac{1}{b}(w - \underline{r} - k) \tag{2.2}$$

Since each resident consumes one unit of land, this condition also determines equilibrium city population

$$N^* = \pi(\overline{x})^2 = \frac{\pi}{b^2}(w - \underline{r} - k)^2 \tag{2.3}$$

2.2.1 Introducing A Transit Line

As in Baum-Snow (2007), I introduce one transit line that is modeled as a linear "ray" emanating from the CBD. The speed ratio on surface streets to the transit line is $\gamma < 1$,[8] so unit cost of taking the transit is $b\gamma$, which is smaller than that of driving. Given severe congestion in cities like Beijing, especially during the peak hours, this seems to be a realistic assumption. For simplicity, I normalize fixed pecuniary costs of driving and taking the

[8]Baum-Snow (2005) discusses the case in which $\gamma > 1$.

transit to zero.[9]

Space is indexed in polar coordinates (x, ϕ), where ϕ is the angle to the transit line. It is convenient to break the urban area up into two areas: one in which residents do not use the transit line for any part of their commutes and another in which residents commute at least partly via the transit. In the former region, commuting cost remains as bx. In the transit commuting region, I restrict to the case where residents drive to the transit line via concentric circles around the origin.[10] Now the unit cost of commuting is $b(\gamma + \phi)$ from location (x, ϕ). $b\phi$ is the commuting cost of reaching the transit line, and $b\gamma$ is the commuting cost of riding the transit. A resident living at location (x, ϕ) chooses the commuting route to achieve the minimum traveling cost: $\min[bx, bx(\gamma + \phi)]$.

Define $\overline{\phi}$ as the angle separating the two regions specified above. It is straightforward to show that $\overline{\phi} = 1 - \gamma$. Residents living beyond $\overline{\phi}$ still commute only on surface streets, while those living at angles $\phi \leq \overline{\phi}$ will use both modes of commuting (See Figure 2.1). The rental price for the latter is

$$R(x) = w - bx(\gamma + \phi) - k \tag{2.4}$$

Comparing equation (2.3) with equation (2.1), rental price at angles $\phi \leq \overline{\phi}$ is higher than the baseline case after the introduction of the transit line. The reason again is lower cost associated with transportation. Similar to equation (2.2), the boundary at $\phi \leq \overline{\phi}$ is

$$\overline{x}(\phi) = \frac{1}{b(\gamma + \phi)}(w - \underline{r} - k)$$

As shown in Figure 2.1, the urban area expands at angles $\phi \leq \overline{\phi}$, comparing with the

[9]See Baume-Snow(2005) for a detailed discussion.

[10]There are other reasonable assumptions on how residents access the transit line using surface streets, such as perpendicularly via linear streets, in which unit cost is $b(\gamma \cos\phi + \sin\phi)$. See Anas and Moses (1979) and Baum-Snow (2007) for detailed discussions.

baseline case. This is caused by rising net income brought by the transit line, pushing residents away from the CBD. Since each resident consumes one unit of land, the equilibrium urban population in this area is

$$N_{(\phi \leq \overline{\phi})} = 2\int_0^{\overline{\phi}} \int_0^{\overline{x}(\phi)} x \mathrm{d}x \mathrm{d}\phi \tag{2.5}$$

2.2.2 Incorporating Traffic Congestion

Thus far, I have assumed that the city is free of traffic congestion, so transport cost is a linear function of commuting distance x. However, traffic congestion is probably one of the most important type of negative externality in large cities like Beijing. Let $N(x)$ be the number of households residing beyond distance x, thus $N(x)$ equals the number of commuters passing through radius x. Now the marginal cost of travel has a fixed component b, as well as endogenous congestion. Here I use an amended version of the Solow (1973) congestion function where the cost per unit distance due to congestion is proportional to demand $N(x)$ and inverse to road capacity $L(x)$

$$\tilde{b}(x) = \left(\frac{N(x)}{L(x)}\right)^{\alpha} + b$$

As in Wheaton (2004), this model assumes that transportation capacity can be provided without using up land, and so $L(x)$ implicitly reflects only road "capital." This simplifies the model. Thus before the construction of the transit line, rent at location x is given as below

$$R(x) = w - \left[bx + \lambda \int_0^x \left(\frac{\overline{x} - t}{L(t)}\right)^{\alpha} \mathrm{d}t\right] - k \tag{2.6}$$

Traffic congestion increases commuting costs at each location x, reducing disposable income located to housing. The transit line does not affect rent prices in the area at angles

$\phi > \overline{\phi}$ because residents will not take advantage of the transit. However, in the area at angles $\phi \leq \overline{\phi}$, rent becomes

$$R(x) = w - \left[bx(\gamma + \phi) + \lambda \int_0^{\overline{\phi}x} \left(\frac{\overline{\phi}x - t}{L(x)} \right)^{\alpha} \mathrm{d}t \right] - k \tag{2.7}$$

A few assumptions need to be discussed. First, I assume that there is no congestion associated with the transit line. Although this is a somewhat extreme assumption, in practice transit companies regularly dispatch more trains during peak hours, easing congestion along the transit. The second term in the bracket represents traffic congestion along the concentric circles.

2.2.3 Empirical Implications

The stylized model allows me to address how changes in transport accessibility affect housing prices, as well as disposable income and urban population. Consider an urban location x at angle $\phi \leq \overline{\phi}$. The transit project is announced at time $t = 0$, and is scheduled to be completed at time $t = T$. The price of the house at x equals to the expectation of the present value of all future rent streams, that is, $P_0(x) = E_0 \left[\sum_0^{\infty} \beta^t R_t(x) \right]$. Since rental price after $t = T$ will increase as shown in equation (2.4), house price at the announcement date $t = 0$ will jump. Since prices already reflect future information, the completion of the transit will have a smaller effect on prices. This leads to the following prediction:

Prediction 5 *The announcement of a new transit line leads to a jump in prices of houses along the line (at angles $\phi \leq \overline{\phi}$), compared with those in other parts of the city. The actual completion of the line may have a smaller effect on prices.*

Equations (2.6) and (2.7) show that congestion increases commuting costs and leads to lower rents. Absent future government measures to curb congestion, house prices will be lower as well. Although it is difficult to compare equations (2.6) and (2.7) directly, if we assume that road capacity is uniform across the city, then we can show that congestion cost is lower after the transit line is built. Coupled with lower commuting costs unrelated to congestion, rents will be higher in the transit commuting region.

In more congested places, where $L(\cdot)$ is smaller, the difference in rents between equations (2.6) and (2.7) is greater. This implies that house prices respond more strongly in places that are more congested.

Prediction 6 *Traffic congestion adversely affects housing prices in a city. The completion of a transit line alleviates traffic congestion along the line, and pushes up house prices. Prices are more responsive in more congested locations.*

The remaining sections will take these predictions to the data. Before doing so, it is worth noting one disconnect between the theory and the empirical study. In the model, the transit commuting region is defined as the region $\phi \leq \overline{\phi}$, while in reality it is likely to be areas close to individual stations. So to more closely match the empirical setting, I would need a model with discrete transit stations. In this paper, I work with the model as described, and test the predictions using distance measures outlines in the empirical section, while acknowledging that the latter is a proxy for the transit commuting region.

2.3 Empirical Setting

In this section, I provide background information on Beijing's urban form and housing market, as well as its transportation system. It highlights the interlink between urban development and transportation accessibility.

2.3.1 Beijing's Urban Form

Figure 2.2 shows Beijing's metropolitan area, which consists of 18 districts. Tiananmen Square and the surrounding commercial, cultural, and administrative areas are traditionally regarded as the city center. The city center roughly spans four districts: Dongcheng, Xicheng, Chongwen, and Xuanwu. Throughout this paper I define Tiananmen Square as the city center. Since China adopted its Economic Reform in 1978, Beijing has expanded rapidly, fueled by strong economic growth. By 2000, the urban area had included four more districts: Chaoyang, Haidian, Fengtai and Shijingshan. This is the area roughly within the 5th Ring Road. Since the early 2000s, the city has seen continued growth in the suburban area, and even the rural area, encouraged by an improved transportation system and government policy to quickly achieve urbanization.

Despite its rapid urban expansion, Beijing remains a monocentric city. Its city center continues to attract a large share of the metropolitan area's economic activities, mainly due to the concentration of urban amenities, as well as government agencies which is a major sector in Beijing. According to Zheng and Khan (2008), over 70% of Beijing's total employment and 65.2% of its population are concentrated within 10 kilometers of Tiananmen Square. This monocentric structure is in stark contrast to many cities in the United States, where employment has spread to the suburban areas (see Glaeser and Kahn, 2001).

Unlike those in many American cities, wealthier residents in Beijing live closer to the city center (see Zheng, Fu and Liu, 2006). For example, residents in Dongcheng District, near the city center, earned disposable income of 34,626 yuan ($5,361) in 2011, 23% more than those living in Shunyi, a suburban district. There are potentially many reasons that have contributed to this pattern. Among them are a concentration of high-paying jobs and amenities near the city center, as well as extremely high opportunity costs related to transportation.

2.3.2 Beijing's Housing Market

Ever since China unveiled its market-based housing reform in 2003, Beijing's property market has developed rapidly. Residential properties have dominated the market, accounting for more than 70% of the total sold floor areas. Due to limited land supply by the municipal government and increased housing demand from an influx of migrants, Beijing's real estate market has become a sellers' market. The average selling price of new housing units has risen rapidly since 2004, from 4,700 yuan to over 17,000 yuan in 2010, an increase of more than three-fold (see Figure 2.3). Pressures on housing prices will continue to exist given rapid economic development and population growth. The number of permanent residents was 19.6 million in 2011, an increase of 44.5% from 2000.

Given limited land supply in the urban areas, recent new residential land supply has concentrated in the suburban districts outside the 5th Ring Road, such as Fangshan, Shunyi, Changping. The construction of suburban metro lines further fueled continued development in these districts.

Since late 2010, the municipal government has introduced a series of restrictive measures to stabilize prices. It has restricted the number of housing units a family can purchase, introduced more government-subsidized housing units, and tightened mortgage standards. These policies have led to lower market turnover. However, the effect on housing prices has been limited. Prices have actually picked up since the summer of 2012.

2.3.3 Beijing's Transportation System

Beijing offers a sophisticated transportation network that comprises of roads, metro lines, railways and an international airport. To meet the needs of rapid urbanization, Beijing's transportation system is under constant growth and reconstruction. This section provides background information on the road network and public transit system.

The Road Network

Beijing possesses a vast network of roads - five ring roads, nine radial toll expressways that connect the city center and suburban areas, eleven China National Highways that depart from Beijing in all directions, and thousands of local roads and streets. The 2nd[11] to 5th Ring Roads were built before 2003, and the 6th Ring Road was completed in September 2009. Most of the expressways and China National Highways were built in the early 2000s or even earlier.

With so much new road and transit capacity every year, Beijing remains the second most congested city in the world. With more than 5 million registered cars on its roads, Beijing's city center is often gridlocked - traffic congestion is widespread within the 5th Ring Road which divides the urban and suburban areas. According to Beijing Transportation Research Center, within the 5th Ring Road, the average speed during peak hours in 2010 was only 20 kilometers per hour (12.4 miles per hour). To combat congestion, the municipal government has implemented a number of policy rules:

1. The government continues to invest heavily in transportation infrastructure, with a focus on expressways and metro lines.

2. The government reduced public transportation fares in an attempt to encourage greater ridership of public transportation.

3. Since April 2010, the government has adopted odd-even traffic restriction on alternative weekdays. Respectively from Monday to Friday, cars with number plates ending with 1&6, 2&7, 3&8, 4&9, and 5&0 were restricted within the 5th Ring Road. This policy would have banned one-fifth of the cars on Beijing's roads if there had been no change in demand.

[11] The 2nd Ring Road is the shortest ring road in Beijing. Within it lies the Forbidden City.

4. Since 2011, the government has imposed a quota of 240,000 new car licenses each year. Residents have to participate in monthly lotteries in order to obtain a new license.

5. In April 2011, the municipal government substantially increased parking fees in the urban area, which may have further reduced traffic congestion.

However, these policy changes have only achieved a limited success in curbing traffic congestion. With ever growing population and car ownership, Beijing's roads remain gridlocked during peak hours, especially around the city center.

The Metropolitan Metro System

Beijing's metro system was proposed in 1953, and its first metro line opened in 1969. The network has grown to 15 lines, 192 stations and 372 kilometers (231 miles) of track in operation. Among the world's metro systems, Beijing's metro network ranks fourth in track length after those of Shanghai, London and Seoul, and fourth in annual ridership after those of Tokyo, Seoul, and Moscow. Although it is the oldest metro system in mainland China, Beijing's metro system did not grow quickly until 2001, when the city won the bid to host the 2008 Summer Olympic Games. As of 2001, the city only had two metro lines in operation, mainly serving the city center. From 2002 and 2008, the municipal government invested 63.8 billion yuan ($7.69 billion) in new metro projects. Total ridership reached 2.18 billion in 2011 from 1.2 billion in 2008. The municipal government plans to achieve a total of 19 lines and 561 kilometers (349 miles) by 2015.

Responding to the global recession of 2008-2009, the Chinese government on November 10, 2008 announced a four-trillion-yuan ($586 billion) economic stimulus package. After the announcement, Beijing's Urban Planning Commission further expedited its metro plans, with a focus on suburban metro lines (surface light rails) that connect urban fringes and suburban districts. In November 2008, three such lines - Lines 15, the Changping Line, and the Fangshan Line - were announced (see Figure 2.3). They collectively added 108 kilometers (67 miles) of tracks, a near 50% increase, making Beijing's metro system the fourth longest in the world, just behind Seoul's Metropolitan Subway. Part of the paper studies the announcement effect of these suburban lines on housing prices.

The City Bus Network

The first bus route in Beijing was established in 1935. The City Bus system has developed rapidly since then, reaching 96 routes in 1975. In 1980, Beijing Public Transport Holdings, Ltd. (BPT) was established, and this state-owned enterprise became the main bus and trolleybus operator in the city. As of 2011, BPT operated over 28,000 buses (including trolleybuses) on 948 bus routes and delivered 4.89 billion rides in 2011. The bus network covers all districts in Beijing.

2.4 Data Description

This study relies upon longitudinal data on Beijing's residential property market as well as detailed information on transport infrastructure and congestion patterns. In this section I describe the steps in constructing the dataset, and provide relevant summary statistics.

2.4.1 Property Prices, Attributes and Location

To kick off the sale of a new residential property in Beijing, a developer is required to submit a presale application to Beijing Municipal Commission of Housing and Urban-Rural Development ("the Commission" thereafter). Upon approval, the Commission publishes on its website basic information about the property, which contains the property's name and location, developer's name, land lease term,[12] and approved date for sale. The Commission maintains a list of all new residential properties since 1996.

Using this database, I identify all residential properties going on sale between 2003 and 2005. The sample starts in 2003, when Beijing's residential real estate market began to take off, and ends in 2005 since transportation data are not available for years before 2005. This yields a balanced panel dataset. Excluding low-income and rental housing projects, the sample yields 1,152 unique commercial residential properties.

From SouFun Holdings Limited (NYSE: SFUN), the largest real estate data vendor in China, I obtain monthly average list prices for each of these properties for the period January 2007 – June 2011. Given that Beijing's housing market remained red hot during this period, the list prices were likely to be quite close to transaction prices. I construct annual list prices for each property for the period 2005–2011. Major property-level attributes are also obtained from SouFun, which are cross-checked with information from the Commission and other data vendors. These property-level attributes include GPS location, total floor area,[13] average size per housing unit, floor area ratio[14] (FAR), green plot ratio[15] (GPR), monthly maintenance fee, land-lease term, among other variables.

For local amenities, I utilize a list of 169 major municipal parks from Beijing Municipal Administration Center of Parks, as well as their GPS locations from Google Earth. 64 city-wide core high schools[16] are also located using information released by Beijing Municipal

[12]All land in China is owned by the government, and can be leased for development. The land lease terms are 70 years for residential use, 40 years for commercial use, 50 years for industrial or mixed use.

[13]Total floor area is the square footage that is located completely above grade, developed on a permanent foundation, is heated for use in all four seasons and has electrical service.

[14]Floor area ratio = (total floor areas of property)/(land areas of property).

[15]Green plot ratio = (total greenery areas of property)/(land areas of property).

[16]Chinese high schools are grouped categorized into two groupings: core high schools and common high schools. Core high schools were designated before the market reforms of the 1980s. Approximately 15% of Beijing's high schools are core schools. Core high schools receive more funding from local governments and

Commission of Education. All geographical information is imported into ArcGIS in order to measure relevant distances.

Figure 2.5 shows that the 1,152 new housing properties are spatially distributed quite evenly across the entire urban area. Table 2.2 provides descriptive statistics of this dataset. The average list price is about 16,961 yuan per square meter ($248.3 per square foot in December 2011) for the period January 2007 – June 2011, with an average unit size of 170 square meters (15.8 square feet). The average property has 991 housing units and is 18.8 kilometers (11.7 miles) from the City Center.

2.4.2 Road Congestion

From Beijing Transportation Research Center, I obtain average morning peak-hour speeds (kilometers per hour) for major roads in Beijing. These are a series of spatial maps spanning the period 2005-2011 (see Figure 2.7 for the 2009 map). The 2008 map only spans three months around Beijing Olympics and does not reflect traffic conditions throughout the year, thus I exclude data in that year. I also use speed-limit maps for the City's major roads from Beijing Traffic Management Bureau (see Figure 2.8 for the 2009 map). In these maps, travel speeds are color coded. Purple represents the slowest speed (0-10 kilometers per hour), while blue renders the fastest speed (greater than 60 kilometers per hour). Importing these spatial data into ArcGIS, I am able to calculate the shortest commute time[17] from each property to the city center. Assuming that there are 250 workdays annually, and morning and evening commute times are roughly the same, I compute annual delay hours as below:

Annual delay hours = (Daily morning travel hours – One-way travel hours at the speed limit)×2×250

are allowed to set more selective entry requirements.

[17] Alternatively, I can use the shortest route from each property to the city center, and then calculate the associated travel time.

To facilitate comparison of congestion across properties, for each property I scale the delay time by the corresponded travel time without delay.[18] This so-called "congestion index" represents how much time is lost due to traffic congestion per hour of drive without congestion.

$$\textit{Congestion Index} = \frac{\textit{Annual delay hours}}{\textit{Annual travel hours at the speed limit}}$$

Table 2.2 shows that for the period 2005-2011, workers living in these properties experienced an annual delay of 285.9 hours during their commutes. The average congestion index is 1.9.

It is worth noting that I make an important assumption that workers drive to work. To the extent that many people in Beijing take the bus to work, the variable – annual delay time – may be underestimated. This is because buses make stops along the way, and workers have to walk to bus stops, all of which take extra time. This has implications for my estimation, and will be discussed in Subsection 2.5.2.

2.4.3 Metro Lines and Bus Routes

To study the announcement effect of new metro lines, I identify the locations of three suburban metro lines whose construction was announced in November 2008. Each line connects the city center and a suburban commercial hub. From Google Earth, I obtain GPS locations of metro stations along the Changping Line, Fangshan Line and Line 15 (see Figure 2.4). There are a total of 29 stations. Furthermore, I locate planned stations for the three metro lines (see Figure 2.5). Beijing had two city-wide transportation plans – in 2000 and June 2007, respectively. The 2000 Plan focuses on inner city metro lines, while the 2007 Plan provides a detailed plan of suburban metro lines (see Figure 2.6 for the 2007 plan). For

[18] I also use an alternative "congestion index," which is calculated by scaling delay time by travel distance. Using this index yields similar results.

this study, I use the latter plan. Comparing the built stations and the planned ones, one discovers that some stations deviated substantially from their planned locations. Section 2.5 explains why these deviations can be considered exogenous to properties in my sample. I also locate all existing metro stations in Beijing. The average property is 11.4 kilometers (7.1 miles) to the closest suburban metro station, and is 5.8 kilometers (3.6 miles) to the closest existing metro station.

The number of bus routes at each property is acquired from SouFun. The average property has access to 7.6 bus routes.

2.5 Traffic Congestion, Public Transport and Housing Prices

2.5.1 Econometric Framework

This section studies whether road congestion, as well as the access to public transportation, has been capitalized into the value of residential properties in Beijing. I measure traffic congestion by annual delay hours, which is the difference between actual annual commute time and annual commute time at the speed limit. Again, let's assume that Beijing is a monocentric city and all workers travel to the CBD to work. Following Coulson and Engle (1987), I express the annual cost of commuting to the city center by

$$C = N\left[(D+L)K + \frac{Q}{M}X + OX\right] + F \tag{2.8}$$

where

$$
\begin{aligned}
N &= \text{number of (one-way) commuting journeys per year} \\
D &= \text{daily delay time to the city center (in hours)} \\
L &= \text{daily commuting time to the city center at limit-speed (in hours)} \\
K &= \text{value of time (yuan per hour)} \\
Q &= \text{price of gasoline (in yuan per liter)} \\
M &= \text{automobile mileage (in kilometers per liter)} \\
X &= \text{distance to the city center (in kilometers)} \\
O &= \text{other monetary costs of commuting (in yuan per kilometer)} \\
F &= \text{Annual fixed costs of owning a car such as taxes and insurance}
\end{aligned}
$$

The first term in the bracket represents the time cost of commuting and the second term represents the gasoline cost. The third term expresses miscellaneous costs associated with commuting, such as tolls and gasoline costs during delays. Since $L = \frac{X}{S}$ where S is the average top-speed, we further write equation (2.8) as

$$C = N\left[KD + \frac{K}{S}X + \frac{Q}{M}X + OX\right] + F \tag{2.9}$$

In equation (2.9), the annual commuting cost C is a linear function in annual delay hours and distance to the city center. In the hedonic pricing function (Rosen, 1974), expected commuting cost should be capitalized into housing prices. Since workers form expectations of future congestion based on current congestion situation (in a linear fashion), we can write the hedonic pricing function as

$$\log(P_{ijt}) = \alpha + \theta \cdot C_{it} + \gamma \cdot Dist\_Metro_{it} + \rho \cdot Bus_{it} + X_i'\beta_t + \psi_{jt} + \epsilon_{ijt}$$

where P_{ijt} represents the logarithm of the list price for property i in district j in year t. C_{it} represents the total commuting cost at property i in year t. $Dist\_Metro_{it}$ is a dummy variable that equals 1 if the distance between property i and the closest existing metro station is less than one kilometer.[19] Bus_{it} represents the number of bus routes available at property i in year t. In addition, X_i is a vector of major property-level time-invariant attributes such as average unit size and floor area ratio. District-year fixed-effects ψ_{jt} control for time-varying district-level characteristics. ϵ_{ijt} is an error term, which represents other unobservable costs. Plugging in equation (2.9), we obtain

$$\log(P_{ijt}) = \alpha + \delta \cdot Delay_{it} + \kappa \cdot Dist\_Center_i + \gamma \cdot Dist\_Metro_{it} + \rho \cdot Bus_{it} + X_i'\beta_t + \psi_{jt} + \epsilon_{ijt} \tag{2.10}$$

Equation (2.10) is the main empirical equation, in which total commuting cost is decomposed into two parts – commuting costs associated with traffic delays and those associated with travel distance (travel at the speed limit). I use time-varying coefficients β_t in order to obtain consistent estimates. Standard errors are clustered at the property level.

I fit OLS regressions for equation (2.10), and the coefficient on $Delay_{it}$ is expected to be negative, reflecting the prediction that costs associated with traffic delays are factored into housing prices. I also replace $Delay_{it}$ with an alternative measure of traffic congestion, the congestion index constructed in Subsection 2.4.2. For robustness checks, I replace $Delay_{it}$ with annual total commuting hours. On the other hand, the coefficient on $Dist\_Metro_{it}$ is expected to be positive, reflecting the fact that households are willing to pay a premium for access to rapid transit. However, it is *a priori* not clear whether more bus connections

[19]One kilometer is the maximum walking distance from properties to metro stations in Beijing, according to studies on Beijing's metro system (see He, 2005).

necessarily translate into a higher housing price. As discussed in Subsection 2.5.4, although an increasing number of bus routes gives local residents more transport choices, the buses themselves may cause additional traffic delays that could have an adverse effect on housing prices.

For the effects of congestion on housing prices, unbiased OLS estimates require no correlation between the error term ϵ_{ijt} and the regressor $Delay_{it}$, conditional on controls and the property and district-year fixed effects. The district-year fixed effects control for regional factors that may systematically affect delay times. However, idiosyncratic factors that affect travel time from one property to the city center cause less concern. It is because traffic congestion depends on the dynamics of the entire road network, so any property level attributes are unlikely to systematically affect delay times. Numerous studies (e.g., Malalur, 2011) have shown that congestion spots on a large road network are stochastic in nature. Nevertheless, as robustness checks, I present additional evidence in Subsection 2.5.3 using the instrumental variable approach.

2.5.2 Effects of Traffic Congestion

Before conducting the regression analysis, it is useful to visualize the relationship between delay time and housing prices over time. I estimate equation (2.10) year by year. Figure 2.9 plots the series of coefficients on $Delay_{it}$. Note that the coefficient is fairly stable over the years, suggesting people's expectation of future congestion is stable.

I then estimate how annual morning delay time affects housing prices. To precisely estimate the effect of congestion on housing prices, I restrict the sample to houses within 50 kilometers of the city center. Table 2.3 shows the OLS estimates from equation (2.10). In column (1), I only include property and district-year fixed effects. The result shows that one standard deviation, 104.5 hours annually, increase in peak-hour delay time leads to a 3.1% decrease in housing prices. The estimate is statistically significant at the 5% level. In column

(2), I add the Euclidean distance from property i to the city center. Now the property fixed effects disappear because of multicollinearity. The coefficient on annual delay hours remains positive and significant, and the magnitude is larger than that in column (1). However, omitted property heterogeneity may potentially bias the finding. To address such a concern, in columns (3) to (5) I add an increasing number of property-level control variables. Column (3) includes property characteristics such as unit size, the floor area ratio, and maintenance fees. Column (4) includes additional property attributes such as amenities and transport accessibility. The estimated effect of delay time now is close to that in column (1), and it is significant at the 1% level.

Overall, these results indicate that a half-hour delay per day translates into a 3.8% decrease in housing prices. This order of magnitude seems large. However, the estimate is based on the assumption that all workers drive to work in Beijing. To the extent that some people in Beijing take the bus to work (bus rides often increase the delay time on the same route given same traffic conditions), I may have underestimated the average delay time in the sample. Even so, by assuming that every worker rides the bus to work, and that bus rides double the delay time I calculate, the effect is still economically significant (half an hour delay per day translates into a 1.9% decrease in housing prices).

It is worth noting that most control variables also capture variations in housing prices. For example, increasing distance to the city center by one kilometer (0.62 miles) is associated with a decline in housing prices of about 1.3%. This order of magnitude is consistent with the figures found in Subsection 2.6.2. Clearly, traffic congestion exerts additional costs on workers other than the regular traffic costs which are approximated by the distance to the city center.

Alternatively, I use the congestion index (defined in Subsection 2.4.2) to estimate how traffic congestion affects housing prices. Table 2.4 presents the results. As before, the coefficient on $Congestion\_Index$ is not significantly affected by the addition of a large set of control variables between columns (1) and (4). The magnitude of the estimates suggests that a one unit increase in the congestion index tends to lower housing prices by 8.1%-13.2%, other things being equal. Across the four regression specifications, most control variables also capture variations in housing prices.

Third, instead of using annual delay time, I use annual total commuting hours as the independent variable of interest. As shown in Table 2.5, total commuting time across different specifications is negatively associated with housing prices. This is not surprising in that the theory predicts that commuting costs are capitalized into housing prices. Interestingly, the coefficient on the distance to the city center is not statistically significant at 10%. Since the distance is a proxy for the regular commuting hours without congestion (which is part of total commuting time), the distance effect is completely absorbed by the total-commuting effect.

2.5.3 The Instrumental Variable Approach

The OLS estimates in Subsection 2.5.2 demonstrate a negative relationship between delay time and housing prices. Although we have argued that any property level attributes are unlikely to systematically affect delay time, as a robustness check, I use the completion of Metro Line 4[20] (a major urban subway line completed in 2009) as an instrument for the delay time. This produces the following first-stage equation:

$$\begin{aligned} Delay_{it} = &\alpha + \delta_1 \cdot Dist\_L4_i \cdot Post\_Constr_t + \delta_2 \cdot Dist\_L4_i + \delta_3 \cdot Post\_Constr_t \\ &+ \kappa \cdot Dist_i + \gamma \cdot Dist\_Metro_{it} + \rho \cdot Bus_{it} + X_i' \beta_t + \psi_{jt} + \epsilon_{ijt} \end{aligned}$$

in which $Dist\_L4_i$ equals 1 if property i is located within one kilometer of the nearest

[20]Line 4 Subway is the 9th subway line in Beijing's mass transit network. It entered into operation on September 28, 2009, and runs from north to south in the western part of the city. The green line in figure 2.4 represents the No. 4 Subway.

Line 4 metro station, and the indicator $Post\_Constr_t$ equals 1 for years after 2009, and equals 0 otherwise. All other variables are defined as in equation (2.10). The identification assumption is that the completion of Metro Line 4 affected delay hours for locations along the line, but were uncorrelated with property level characteristics that affected housing prices. I show strong evidence to justify the first assumption, and argue that the completion of the metro line is an exogenous event which is unlikely to be correlated with property level attributes. Similar to findings in Section 2.5, the distance dummy $Dist\_L4_i$ is unlikely to be correlated with property level attributes.

The instrumental variable estimates are predicted to be larger than the OLS estimates if housing prices are negatively correlated with delay time due to reverse causality. It could be the case that car ownership at expensive properties was higher which led to more local congestion. The regressions in Table 2.6 replicate those in Table 2.3 but replace OLS results with two-stage least squares estimates. The first stage F-statistics are always larger than 10 (see e.g., Staiger and Stock, 1997). Across the four specifications, the instrument is negatively correlated with delay time (statistically significant at 1% level). This indicates that the completion of Line 4 reduced delay hours for properties along the line, due to the possibility that few residents drive or take the bus to work in these neighborhoods. In column (1), with only property and district-year fixed effects, the estimated coefficient is -0.0012 (significant at the 1% level), which is larger than the corresponding OLS estimate. In columns (2) to (4), results with additional controls are reported. The coefficients are consistently larger than the OLS counterparts. It is likely that some form of endogeneity or measure error in delay time is operating against the results. This implies that the true effect of traffic congestion may be larger than the OLS estimates suggest. However, the magnitudes in the IV estimates are almost four times as large. An appropriate interpretation of the IV results may be to confirm the existence of a negative causation from traffic congestion to housing prices, without reading much into the magnitudes.

2.5.4 The Role of Public Transportation

Prediction 5 shows that access to metro lines is priced in property values. Since these rapid transit lines reduce the opportunity cost involved in travel, more disposal income could be used to purchase housing services. This implies that γ in equation (2.10) is expected to be positive. Column (4) of Table 2.3 shows that properties within one kilometer of the closest metro station are 3.9% more expensive than those further away from metro stations, other things being equal. Replacing delay hours with the congestion index or total commute hours, we have similar results, as shown in Tables 2.4 and 2.5. This difference is comparable to about 30 minutes of delay time each day.

To see the heterogeneous effect of metro lines on values of properties with different degrees of congestion, I further include an interaction term $Delay_{it} \cdot Dist\_Metro_{it}$ in equation (2.10). The coefficient on this interaction term is expected to be positive because in more congested residential areas, the value of accessibility to metro lines is greater. In other words, the effect of congestion on housing prices would be smaller. In column (5) of Table 2.5 I re-run regression (2.10) with the inclusion of this interaction term. As predicted, the estimated coefficient is positive.

On the other hand, additional bus routes do not necessarily lead to higher housing prices. Due to the severe congestion situation in Beijing, the benefit of additional bus routes may be minimal. As shown in column (4) of Table 2.3, the coefficient on the number of bus routes is economically small, and is not statistically significant. With alternative measures of congestion, the coefficient remains small and statistically insignificant (Tables 2.4 and 2.5).

2.6 The Announcement Effect on Housing Prices

In Subsection 2.5.4, I show that proximity to existing metro lines is captilized into housing prices in Beijing. However, the distance between a property and its closest metro station may not be exogenous. Several existing metro lines – most of the gray lines in Figure 2.4 – were either built before or during the construction of the housing properties in my sample.[21] Developers could potentially choose to purchase land close to these existing transit lines,[22] rendering the distance endogenous. To overcome this obstacle, I utilize three of the new suburban metro lines built between April 2009 and December 2010, to study the effect of access to public transportation on traffic congestion and housing prices.

2.6.1 Econometric Framework: Difference-in-Differences

This section examines whether the announcement of three suburban metro lines, the Changping Line, Fangshan Line and Line 15, triggered a jump in prices of properties along the lines, compared with those in other parts of Beijing. As mentioned in Subsection 2.3.3, the November 2008 announcement was part of China's stimulus package responding to the global recession of 2008-2009. The announcement of these three lines was widely unexpected.[23] This enables me to make causal claims regarding these construction announcements.

[21]Properties in my sample went on sale between 2003 and 2005. No.1 and No.2 Subway Lines were built before 2003, and several other lines started construction between 1999 and 2005.

[22]Although Beijing Municipal Commission of Urban Planning claims to follow strict urban planning guidelines, developers occasionally could influence the zoning process.

[23]The 2008-2009 Chinese economic stimulus plan was an attempt to minimize the impact of the global financial crisis on the world's second largest economy. To Chinese leaders, the financial crisis, as well as the ensuing economic slowdown was unexpected. The Economic Times reported on November 10, 2008 that "The plan follows an unexpectedly sharp slowdown in economic growth that has raised the prospect of job losses and unrest."

As a baseline, I estimate the following hedonic pricing function:

$$\log(P_{ijt}) = \alpha + \delta_1 \cdot D_i \cdot Post_t + \delta_2 \cdot D_i + \delta_3 \cdot Post_t + X'_{1i}\beta_t + X'_{2it}\gamma + \psi_{jq} + \epsilon_{ijt} \qquad (2.11)$$

The logarithm of the price[24] for property i in district j in month t is a function of: a binary variable D_i, which equals 1 under either of two conditions: (1) if property i is located within one kilometer of the nearest new metro station, or (2) property i is in a suburban commercial hub; an indicator $Post_t$, which equals 1 for months after the announcement month; major property-level time-invariant attributes X_{1i} and time-varying attributes X_{2it} such as distance to the nearest existing metro station. I also include district-quarter fixed-effects ψ_{jq} which control for time-varying district-level characteristics. I use time-varying coefficients β_t in order to obtain consistent estimates. Standard errors are clustered at the property level. δ_1 is a difference-in-differences estimator. If prices of properties along the metro lines jumps after the announcement relative to those in other parts of the city, $\delta_1 > 0$.

The treatment group includes properties within one kilometer of the nearest metro station, given that one kilometer is the maximum walking distance, according to various studies on Beijing's metro system (see He, 2005). Beyond one kilometer, people would take the bus or drive to work. The treatment group also contains properties in suburban commercial hubs (even if they are not within walking distance to the metro stations) insofar as the metro lines will be their primary mode of transportation to reach the city center.[25]

I begin by estimating equation (2.11) using OLS. Unbiased OLS estimates require no correlation between the error term ϵ_{ijt} and the regressor D_i, conditional on controls and

[24] I use nominal prices throughout the paper because the difference-in-differences coefficient estimates the differential trend in housing prices regardless of nominal or real prices. Time fixed effects also take into account the effect of inflation.

[25] Residents in suburban commercial hubs living beyond walking distance to metro stations often take the bus to the stations, and use the metro as their main transportation mode.

the property and district-year fixed effects. This requirement would fail if metro stations were constructed in areas that were expected to experience high growth in housing prices. However, this type of reverse causality is less of a concern as it will likely inflict a negative bias on my results. In other words, absent reverse causality, the magnitudes may be even larger.

2.6.2 Baseline Results

Before we conduct the regression analysis, it is useful to visualize whether the announcement of these suburban lines triggers a jump in prices of properties along the lines. I first regress the logarithm of monthly prices $\log(P_{ijt})$ on the distance dummy D_i, property fixed-effects and district-quarter fixed-effects. The regressions are carried out month by month for the period July 2006 – June 2011. Figure 2.10 plots the series of coefficients on the distance dummy. Note that there is a sharp increase in the coefficients in the months following November 2008. This is consistent with the notion that the announcement has a larger positive effect on prices of properties in the proximity of the metro stations. It is worth mentioning that the coefficients experience a sharp decline around September 2009, when the urban Metro Line 4 began its operation. The coefficients also has a sizable increase after December 2010 when the three suburban lines were completed.

To precisely estimate the announcement effect on housing prices, I restrict the sample period to six months before and after the announcement, which is May 2008 to May 2009. Table 2.7 shows the OLS estimates from equation (2.11). In column (1), I only include property and district-quarter fixed-effects. The result shows that within a six-month window, the construction announcement leads to an increase in prices of properties along the metro lines that are 3.3% higher than those of the rest. This translates to a difference of 560 yuan per square meter ($8.0 per square foot in January 2009). In columns (2) to (5), results with additional controls are reported. The estimated announcement effect (3.1%-3.6%) on housing prices is stable to the addition of these control variables, and it remains significant at the 1% level.

It is worth noting that most control variables also capture variations in housing prices. For example, increasing the distance to the city center by one kilometer (0.62 miles) correlates with a decline in housing prices of about 1.8%. When the floor area ratio increases by 1 unit, housing prices tend to fall by 1.7%.

2.6.3 Effects During the Whole Sample Period

In addition to the announcement effect, it is interesting to explore whether the issuance of the June 2007 Metro Plan or the completion of the metro lines (December 2010) had any effects on housing prices. The sample is extended to the period January 2007 – June 2011.[26] To test these potential effects, I add two interactions terms to equation (2.11): $E(D_i) \cdot Post\_Plan_t$ and $D_i \cdot Post\_Constr_t$, where $E(D_i)$ denotes the distance (a dummy variable as D_i) between a property and its closest metro station on the 2007 Plan. Insofar as prices of properties along the metro lines jump after the 2007 Metro Plan was released, the coefficient on $E(D_i) \cdot Post\_Plan_t$ is positive. Similarly, as property prices responded differentially to the completion of the metro lines, the coefficient on $D_i \cdot Post\_Constr_t$ is positive.

Table 2.8 shows the associated estimates. In column (1), I only include property and district-quarter fixed-effects. The result shows that between January 2007 and June 2011, the construction announcement (project completion) leads to an increase in prices of properties along the metro lines that is 10.4% (4.1%) higher than those of the rest. On the other hand, the release of the Metro Plan has little effect on housing prices. Between columns (2) and

[26] The beginning of this series is six months before the plan was issued, and six months after the lines were completed.

(4), I add an increasing number of property-level control variables,[27] such as the distance to the city center and unit size. The coefficients on the three interaction terms remain stable and consistent with those found in column (1).

2.6.4 Heterogeneous Announcement Effects

Housing prices in more congested areas are predicted to be more responsive to announcements of new metro lines (Prediction 6). To test this hypothesis, I first divide the sample into two groups. Properties in Group A are located in congested areas, whose congestion indices (defined in Subsection 2.4.2) in 2007 are above the median index. Properties in Group B are located in less congested places, whose congestion indices in 2007 are below the median one. I then estimate equation (2.11) for these two groups separately. The estimated coefficient δ_1 for Group A is predicted to be larger than that for Group B.

Panel A in Table 2.9 shows the estimates of announcement effects for Group A. Between columns (1) to (4), I add an increasing number of property-level controls. The result indicates that for congested areas, the construction announcement leads to an increase in prices of properties along the metro lines that are 3.7% higher than those of the rest. However, for less congested places, the corresponding coefficients are neither economically nor statistically significant at the 5% level, as shown in Panel B. These results suggest that the announcement effect is larger for more congested areas.

2.7 Discussion

The empirical results in Subsection 2.5.2 are appealing in that they provide necessary information for calculating the cost of traffic congestion in Beijing. A lower bound of this cost is the market price of avoiding delay as reflected in higher housing prices. This lower bound

[27]The property fixed-effects disappear because of multicollinearity.

should largely capture the private costs of congestion, while public costs, such as increased air pollution, are not internalized by homeowners. There are two major categories of private costs due to congestion – monetary costs and non-pecuniary costs. Monetary costs include lost wages and business opportunities associated with delay time, extra gasoline expenditures, and wear and tear on vehicles as a result of frequent acceleration and braking while idling. Non-pecuniary costs, on the other hand, include health issues resulting from higher levels of stress and frustration. According to IBM's 2011 Global Commuter Pain Survey, 42% of people say their stress level has increased and 35% say they get angry because of traffic congestion.

This paper studies costs of traffic congestion, as well as benefits of new transit lines, as reflected in housing prices. As shown in the model, however, there is a more direct link between rents and transport accessibility. For example, rents should not respond to the announcement of the new transit lines; rather they will increase only when the line is built. Due to data constraint, in this paper I study how housing prices respond to infrastructure. Studying how rents and rent-to-price ratios react to congestion and transport infrastructure is left to future work.

I use nominal prices throughout the paper since time fixed effects take into account the effect of inflation. In addition, the difference-in-differences coefficients estimate the differential trend in housing prices regardless of nominal or real prices. However, it will be interesting to measure how prices respond to congestion and accessibility in real terms. After all, Beijing experienced high inflation during the sample period.

The paper studies the interplay between transport accessibility and property values. In reality, transportation modes may have agglomeration effects: they may affect firm productivity and flows of goods. Cost-benefit analysis of the effects of transportation on productivity in urban areas is worth further study (see e.g., Mun and Yonekawa, 2006; Weisbrod, Vary and Treyz, 2007; Hartgen and Fields, 2009).

2.8 Conclusion

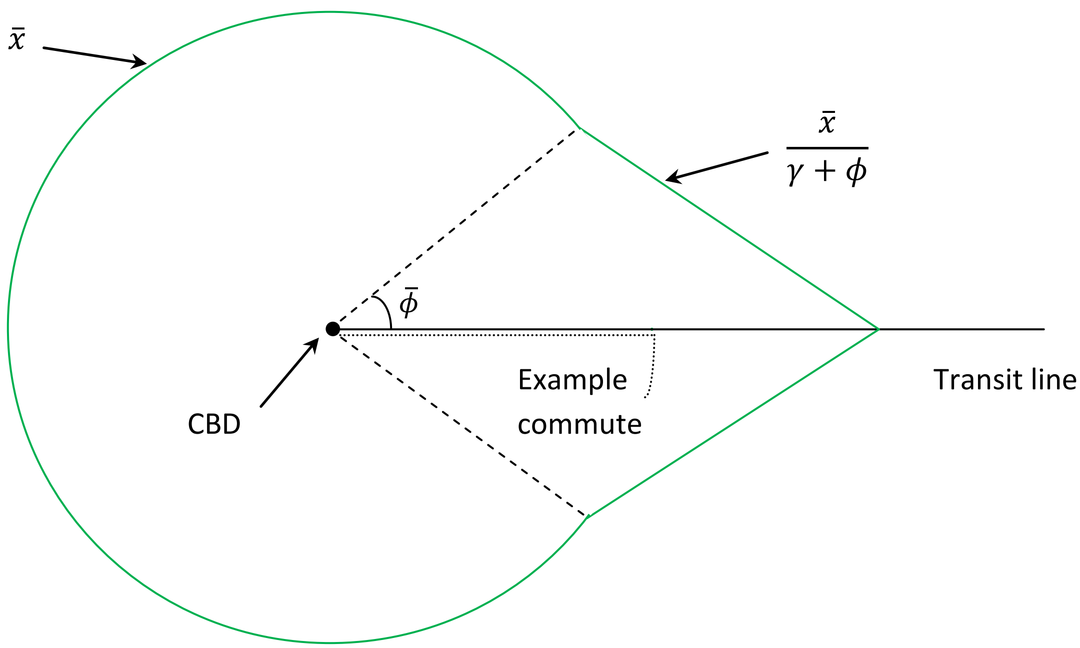
Existing urban economic theory states that traffic congestion is welfare reducing. In practice, policymakers in congested cities invest heavily in public transit systems to reduce transportation costs. However, not all public transit modes are created equal – those that reduce traffic congestion are the most desirable. Using a unique panel dataset of all residential properties on sale between 2003 and 2005 in Beijing, I first show that traffic congestion has a large adverse effect on housing prices. A half-hour increase in daily delay time (drive time) is associated with a 3.8% decrease in housing prices. This implies that there may be a huge welfare cost to congestion. first papers that use micro-level congestion data to assess property values.

Furthermore, I show that announcements of metro line construction inflate prices of properties near future stations – the difference-in-difference coefficient of 3.3% translates to a difference of 560 yuan per square meter ($8.0 per square foot). The increase is even more staggering for more congested areas. This suggests that metro lines are effective in reducing the impact of congestion. The magnitude of capitalization effects appears to be larger than findings in developed countries (e.g., McMillen and McDonald, 2004; Gatzlaff and Smith, 1993), possibly due to lower car ownership as well as higher density. However, additional bus routes are not capitalized into prices because buses move slowly in the gridlocked city, often exacerbating rather than alleviating congestion.

I conclude that the mere quantity of public transit services may not increase welfare. In congested cities, workers seem to attach different premiums to different public transit modes. My results show that housing prices are sensitive to new metro lines but not to additional bus routes. The reason may be that in extremely congested cities, additional bus routes exacerbate rather than alleviate congestion. Although further cost-benefit analysis needs to be applied to both transit modes (e.g., Winston, 2007), it is reasonable to argue that the metro system may be more desirable in megacities. Note that this paper does not discuss direct measures that could curb traffic congestion, such as congestion pricing and highway investment (see e.g., Vickrey, 1969 and Leape, 2006 for a discussion). As of November 2012, the Beijing Municipal Government was discussing a possible congestion charge in the urban area, and I plan to analyze it in a follow-up study.

While the findings in this paper are promising, it does not address the interplay between mass transit and traffic congestion, as well as the welfare implications of unpredictable congestion. Promising areas for future work include the following: how transport accessibility affects workers' employment decisions as well as their location choice; and how different types of property developers (state-owned enterprises versus private developers) react to changes in transport accessibility.

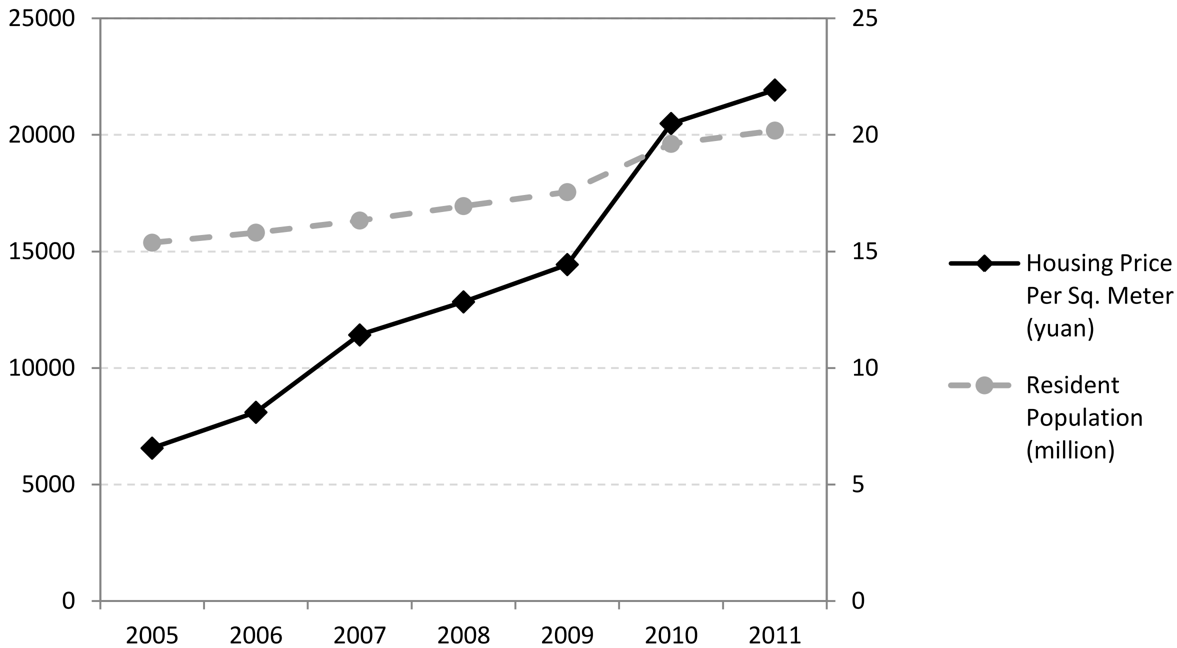
Figure 2.1: The Effect of a New Transit Line on Urban Form



Notes: This figure is similar to Figure 1 in Baum-Snow (2007). $\bar{\phi}$ is the angle separating the transit commuting region and the rest. $\bar{x}$ is the original urban fringe, and $\frac{\bar{x}}{\gamma+\phi}$ is the expanded urban fringe due to the transit line. An example commute is a case in which residents drive to the transit line via concentric circles around the origin, and then take the transit to the CBD.

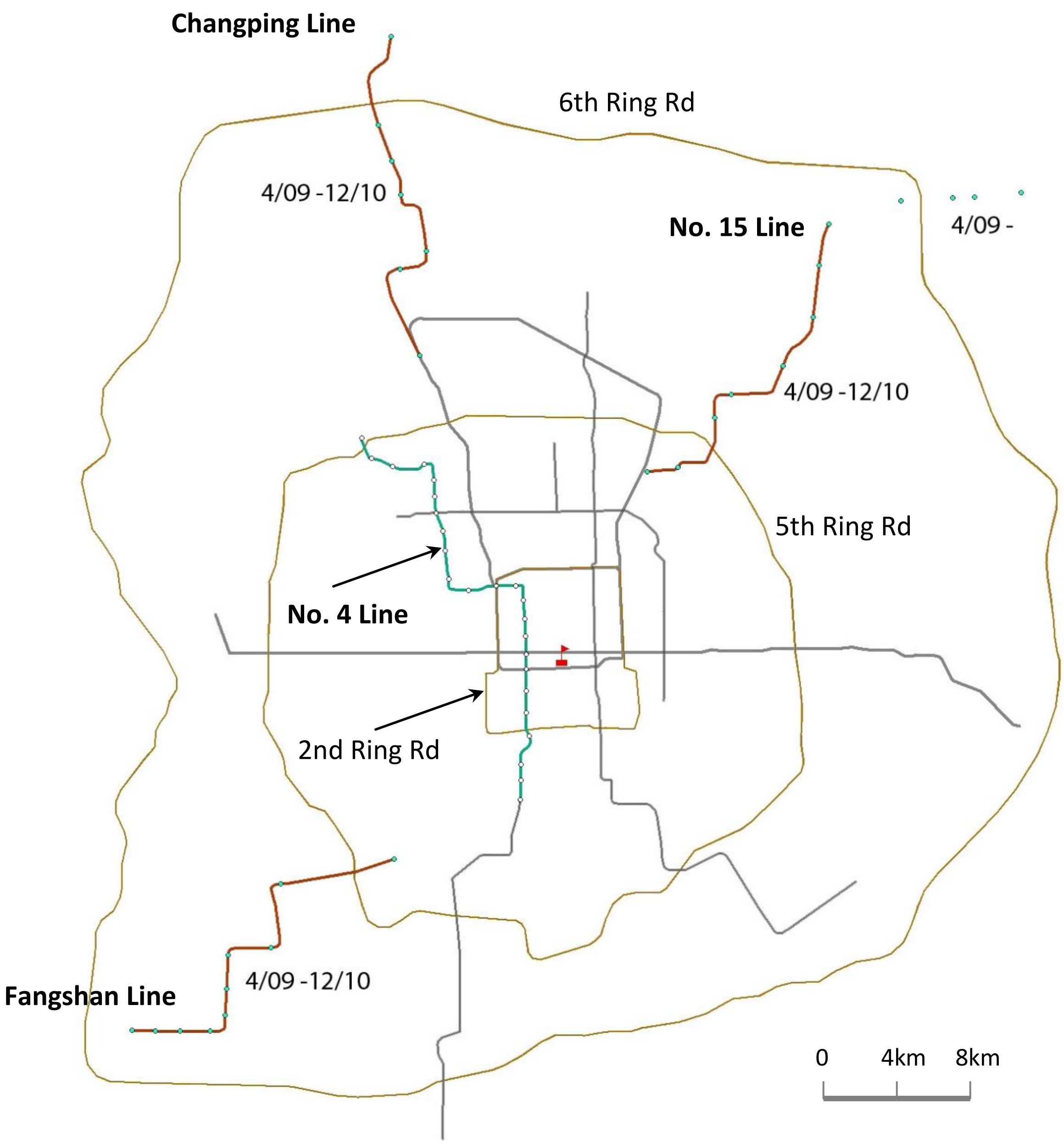
Figure 2.2: Beijing District Level Map

Figure 2.3: Housing Prices and Population Growth in Beijing



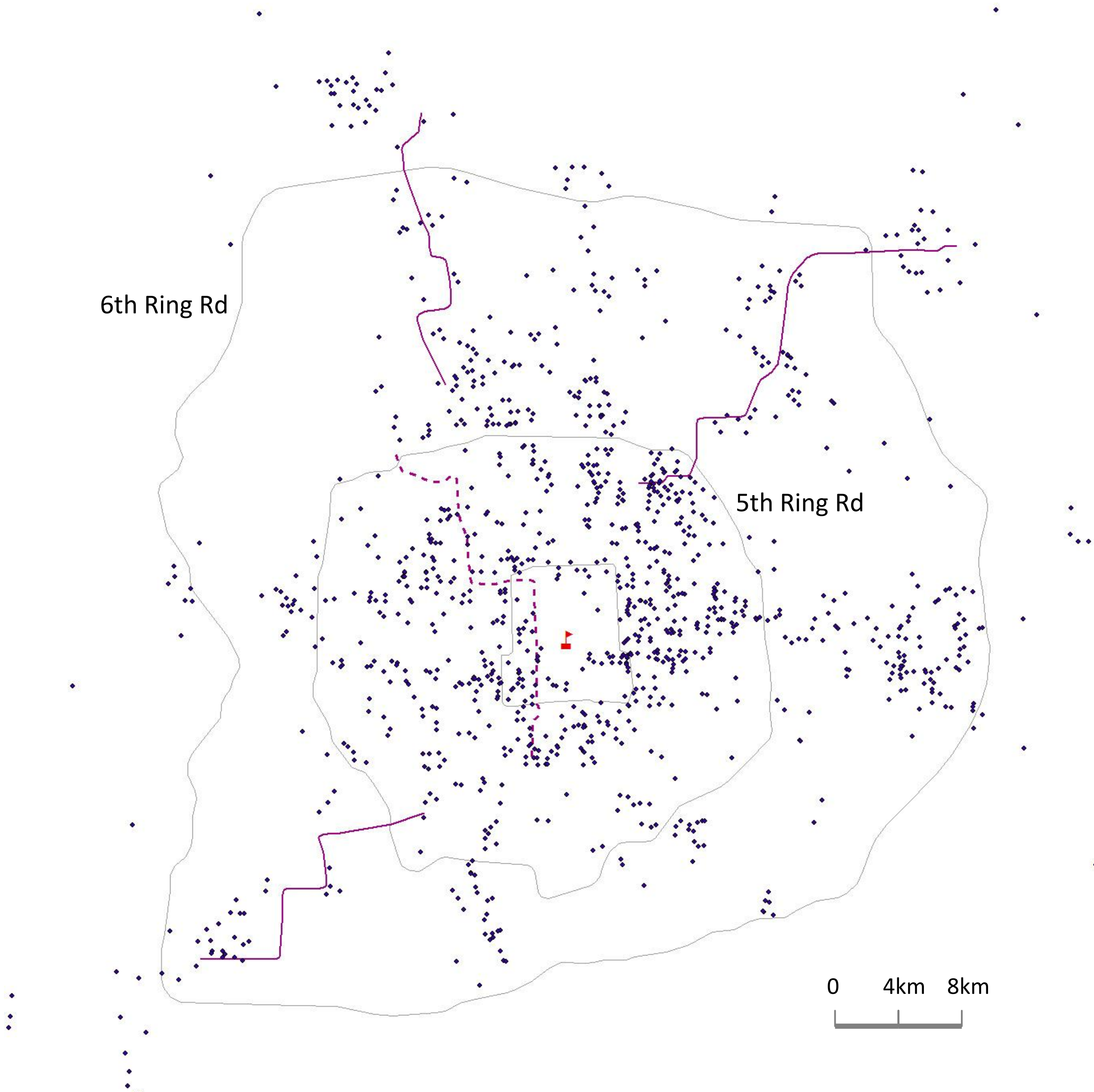
Notes: This figure plots transaction prices for commercial housing properties and resident population in Beijing. The solid line (left axis) plots housing prices per square meter. The dashed line (right axis) plots total resident population in Beijing. Housing prices are from SouFun Holdings Limited, and population data are from Beijing Statistical Yearbook (2006-2012).

Figure 2.4: Three Suburban Metro Lines in Beijing



Notes: The three suburban metro lines Changping Line, Fangshan Line and No. 15 Line are in brown. The lines were announced in November 2008. Construction began in April 2009 and ended in December 2010. The No. 4 Line is in green. The flag represents the city center.

Figure 2.5: Locations of Properties and Metro Lines



Notes: Each dark dot represents a residential property. The three suburban metro lines Changping Line, Fangshan Line and No. 15 Line are in purple. The dashed line represents the No. 4 Metro Line. The flag represents the city center.

Figure 2.6: The June 2007 Metro Plan (Features Suburban Lines)

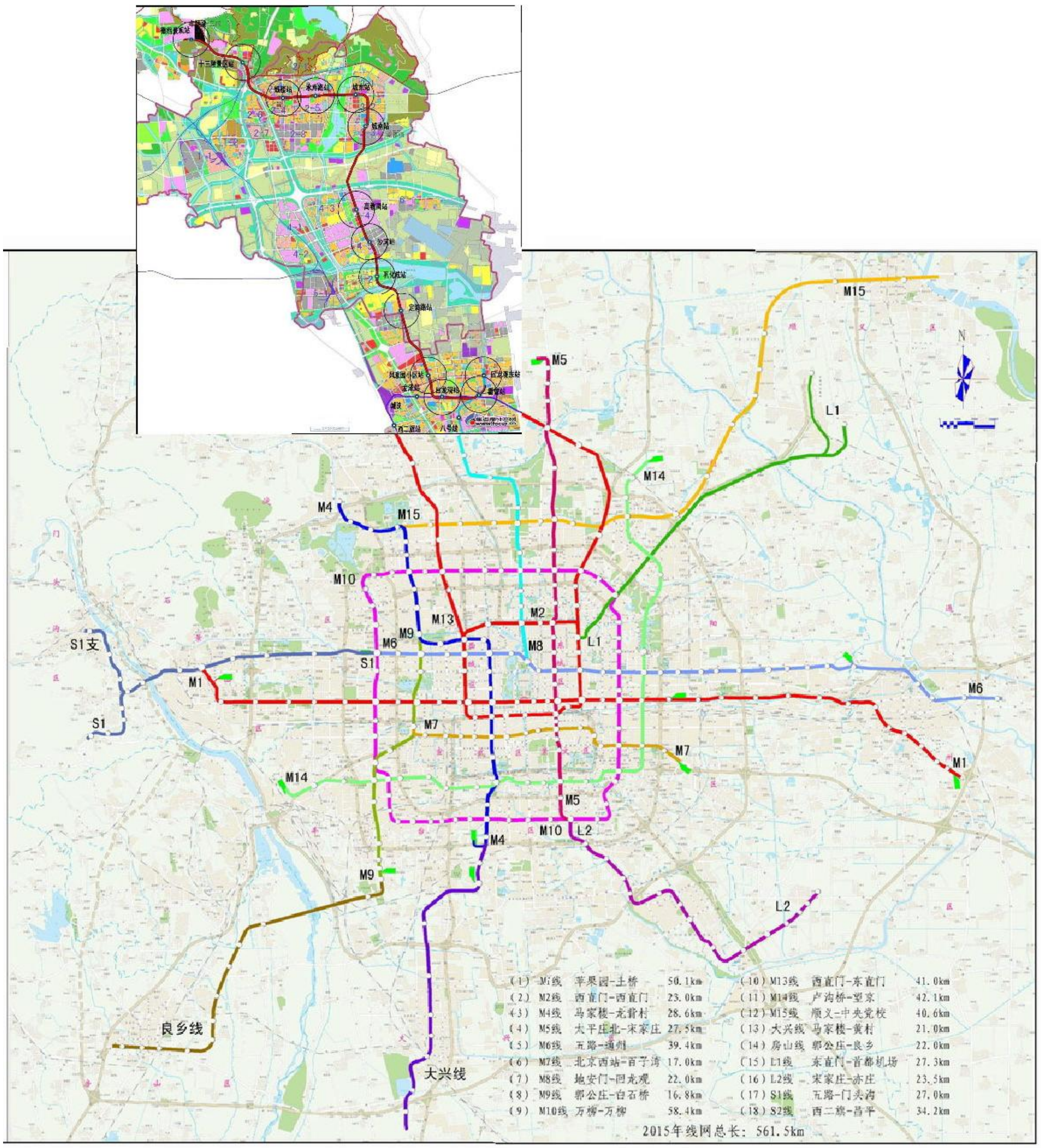


Figure 2.7: Congestion Map for Beijing (Morning Peak Hours), 2009

Figure 2.8: Speed-Limit Map for Beijing, 2009

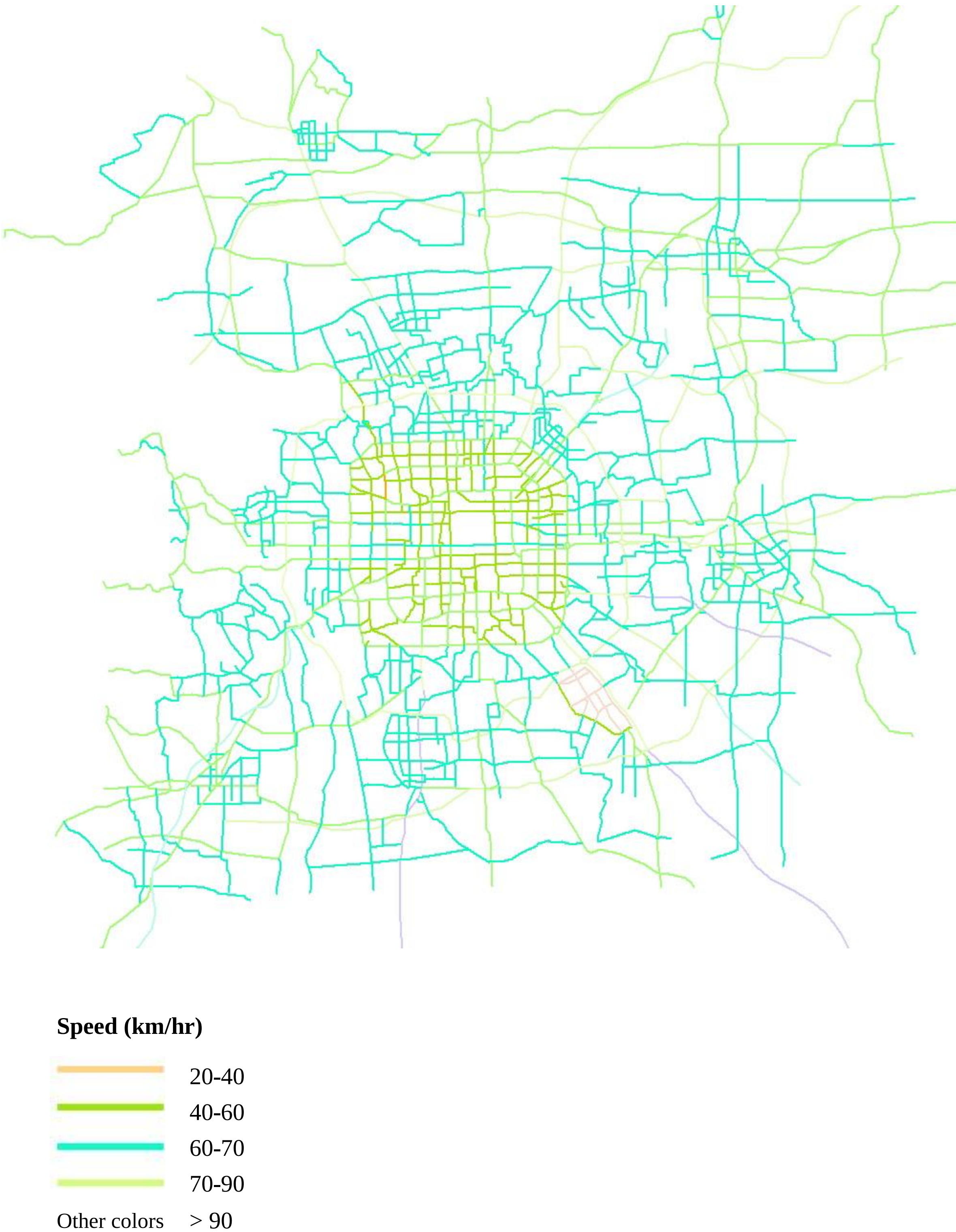
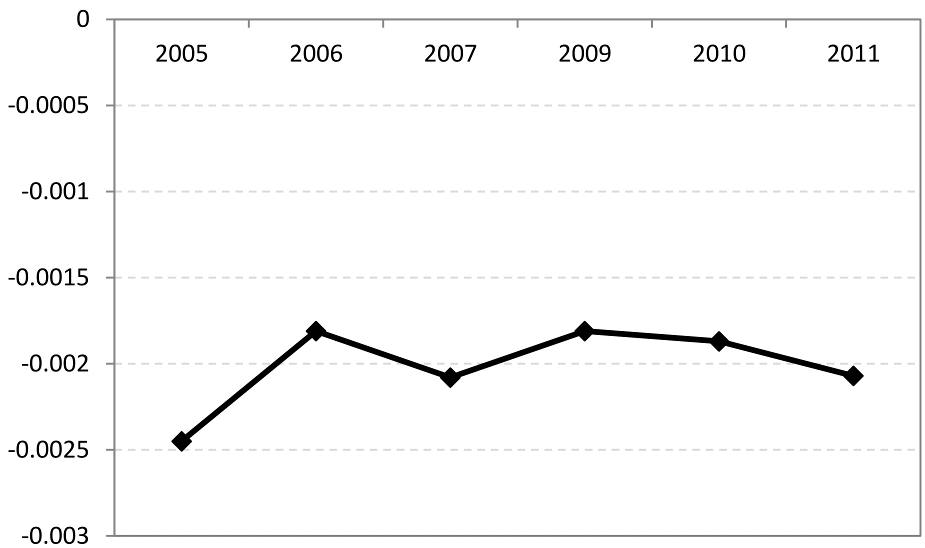
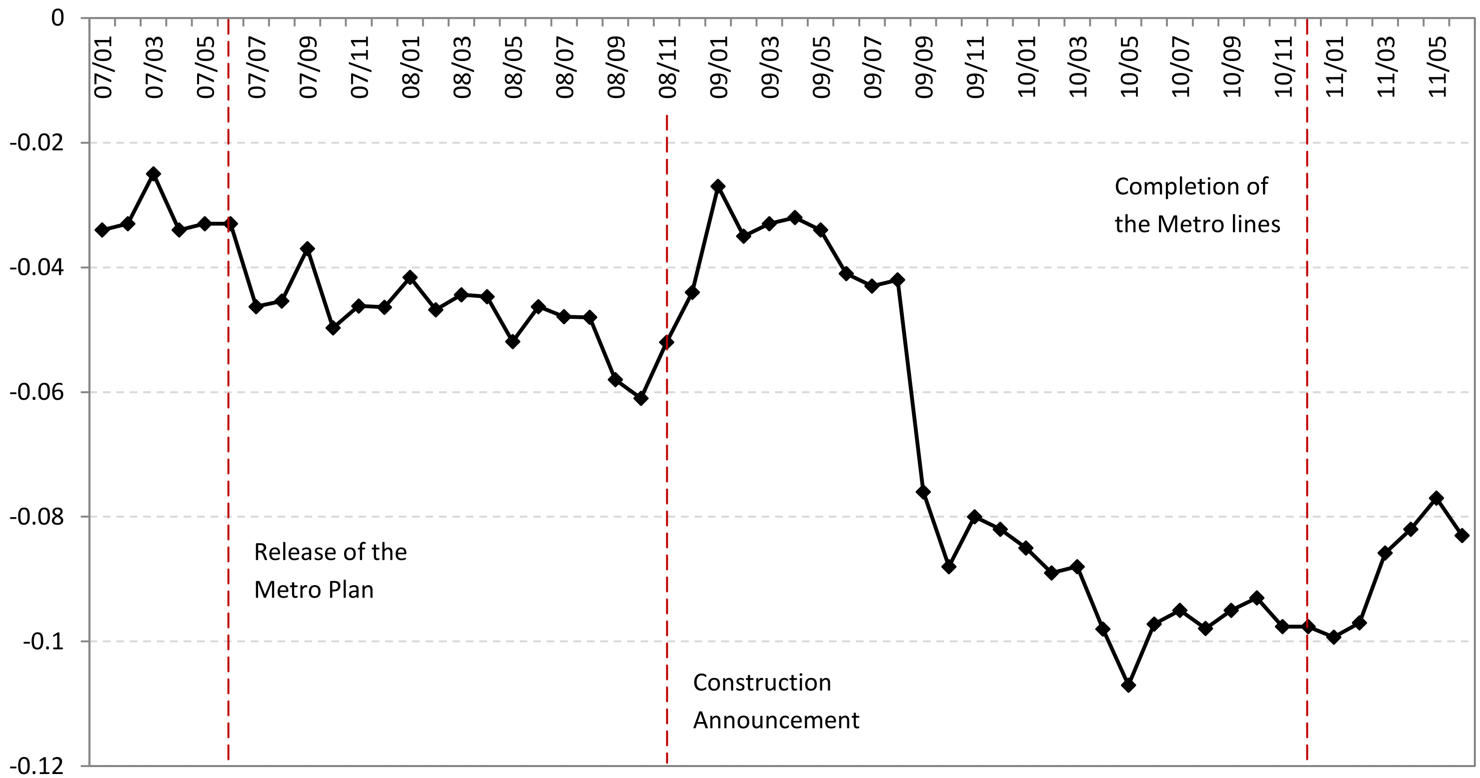


Figure 2.9: The Effect of Delay Time on Housing Prices



Notes: I regress the logarithm of annual housing prices on annual delay hours and district fixed-effects, in which annual delay hours = (daily morning travel hours one-way travel hours at speed limit)$\times$500. The regressions are carried out year by year for the period 2005-2011 (excluding 2008). The coefficient on annual delay hours is then plotted out for this period.

Figure 2.10: Difference in Monthly Housing Prices (Properties Along the Metro Lines Versus the Rest)



Notes: I regress the logarithm of monthly housing prices on 1{along suburban line} and district fixed-effects, in which 1{along suburban line} equals 1 if the property is located with one kilometer of the nearest suburban metro station, or it is in a suburban commercial hub. The regressions are carried out month by month for the period July 2006 June 2011. The coefficient on 1{along suburban line} is then plotted out for this period.

Table 2.1: Metro Lines in Operation as of December 31, 2011

Line	Opened	Construction Started	Length (km)	Stations (surface)
1	1999	--	30.4	23 (2)
2	1987	--	23.1	18
Batong Suburban Line	12/17/2003	12/28/2001	18.9	13(13)
13 Suburban Line	1/28/2003	12/1/1999	40.9	16(15)
5	10/7/2007	12/28/2002	27.6	23(7)
8 (Phase 1)	7/19/2008	5/1/2005	4.5	4
10 (Phase 1)	7/19/2008	12/27/2003	24.7	22
4	9/28/2009	8/1/2004	28.2	24(1)
Yizhuang Suburban Line	12/30/2010	10/16/2008	23.3	14(8)
Daxing Suburban Line	12/30/2010	10/28/2008	21.7	11(1)
15 (Phase 1, Sec 1)	12/30/2010	4/11/2009	20.2	8(4)
Changping Suburban Line	12/30/2010	4/2/2009	21.2	7(6)
Fangshan Suburban Line	12/30/2010	4/9/2009	24.7	11(9)
15 (Phase 1, Sec 2)	12/30/2011	4/11/2009	10	5
9 (South Sec)	12/31/2011	Apr 2007	10.8	9

Notes: The shaded lines are the focus of this study.

Table 2.2: Summary Statistics for 1,152 Properties

Variable	Mean	Median	Max	Min	Std Dev
Monthly list price (yuan/m^2), Jan 07 – Jun 11	16,961	14,778	130,399	1,907	9,587
Annual list price (yuan/m^2), 2005-2011	15,969	13,500	95,115	1,810	10,376
Location and Accessibility					
Distance to city center (km)	18.8	15.6	49.9	2.2	11.1
Distance to nearest park (km)	2.8	1.9	25.2	0.08	3.0
Distance to nearest core high school (km)	4.1	3.0	20.8	0.12	3.7
Distance to nearest suburban-line station (km)	11.4	11.3	47.9	0.13	7.3
Distance to existing metro line station (km)	5.8	2.9	42.4	0.10	7.0
Congestion Indicators					
Annual commute hours	449.9	437.6	928.6	63.4	173.1
Annual delay hours	285.9	281.6	622.4	39.0	104.5
Congestion index	1.9	1.8	5.4	0.52	0.56
Property Attributes					
Total floor areas of property (km^2)	0.14	0.10	2.5	0.002	0.17
Housing units within a property	991	650	11,026	1	1,162
Unit size (square meters)	170.0	136.0	3,000	25	150.8
Floor area ratio	2.7	2.4	19.6	0.08	1.9
Green plot ratio	0.37	0.35	0.89	0.10	0.10
1{interior construction included}	0.26	0	1	0	0.44
Maintenance fee (yuan/month/m^2)	2.7	2.3	18	0	2.1
No. of parking lot per household	0.88	0.98	3	0.07	0.41
No. of bus lines	7.6	6	38	0	5.1
1{central heat}	0.77	1	1	0	0.42
Frequency					
Properties going on sale in 2003	19.7%				
Properties going on sale in 2004	32.5%				
Properties going on sale in 2005	47.8%				

Notes: The sample includes all commercial residential properties going on sale between 2003 and 2005 in Beijing. It excludes low-income housing projects and those missing all prices. Annual commute hours = Daily morning travel hours $\times$ 500. Annual delay hours = (Daily morning travel hours One-way travel hours at the speed limit) $\times$ 500. Congestion index = (Annual delay hours)/(Annual travel hours at the speed limit). Floor area ratio = (total floor areas of property)/(land areas of property). Green plot ratio = (total greenery areas of property)/(land areas of property). 1{central heat} equals 1 if the housing units are centrally heated.

Table 2.3: The Effect of Delay Time on Housing Prices

	Dependent Variable: Log (price)				
	(1)	(2)	(3)	(4)	(5)
Annual delay hours	-0.0003**	-0.0007***	-0.0004***	-0.0004***	-0.0005***
	(0.0001)	(0.0003)	(0.0002)	(0.0002)	(0.0003)
1{dist to existing metro<1km}				0.039**	
				(0.017)	
1{dist to existing metro<1km}*					0.0003***
Annual delay hours					(0.0001)
No. of bus lines				0.001	0.001
				(0.001)	(0.001)
Dist to City Center (km)		-0.010	-0.012***	-0.013***	-0.012***
		(0.006)	(0.004)	(0.004)	(0.004)
Dist to nearest park (km)				-0.002	-0.002
				(0.003)	(0.003)
Dist to nearest core school (km)				-0.005	-0.004
				(0.003)	(0.003)
Unit size (square meters)			0.0003*	0.0002*	0.0002*
			(0.0001)	(0.0001)	(0.0001)
Floor area ratio			-0.016**	-0.014***	-0.014***
			(0.005)	(0.004)	(0.004)
Green plot ratio			0.384***	0.323***	0.322***
			(0.076)	(0.082)	(0.081)
1{interior construction included}			0.045*	0.044**	0.045**
			(0.023)	(0.021)	(0.022)
Maintenance fee (yuan/month/m^2)			0.080***	0.072***	0.073***
			(0.011)	(0.008)	(0.008)
No. of parking lot per household				0.097***	0.097***
				(0.015)	(0.015)
1{central heat}				-0.021	-0.020
				(0.020)	(0.020)
Constant	8.892***	9.342***	8.880***	8.872***	8.879***
	(0.030)	(0.086)	(0.099)	(0.116)	(0.114)
Property FE	Yes	No	No	No	No
District-year FE	Yes	Yes	Yes	Yes	Yes
Observations	6,223	6,223	6,172	6,166	6,166
R-squared	0.95	0.79	0.88	0.88	0.88

Notes: The sample period is 2005 2011 (excludes 2008). It includes all properties within 50 kilometers of the city center. Annual delay hours = (Daily morning travel hours One-way travel hours at the speed limit) × 500. 1{dist to existing metro < 1km} equals 1 if the property is located with one kilometer of the nearest existing metro station. Please refer to Table 2.2 for details on the construction and definition of other controls. Coefficients are reported with robust standard errors in parentheses. Standard errors are clustered on the property level. Time-varying coefficients are not reported due to space. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels.

Table 2.4: The Congestion Index and Housing Prices

	Dependent Variable: Log (price)				
	(1)	(2)	(3)	(4)	(5)
Congestion index	-0.132**	-0.081**	-0.084**	-0.086*	-0.020
	(0.067)	(0.043)	(0.041)	(0.048)	(0.027)
1{dist to existing metro<1km}				0.040**	
				(0.015)	
1{dist to existing metro<1km}*					0.023***
Congestion index					(0.007)
No. of bus lines				0.001	0.001
				(0.001)	(0.001)
Dist to City Center (km)		-0.022***	-0.018***	-0.018***	-0.018***
		(0.007)	(0.005)	(0.005)	(0.006)
Dist to nearest park (km)				-0.001	-0.001
				(0.002)	(0.002)
Dist to nearest core school (km)				-0.005*	-0.005*
				(0.003)	(0.003)
Unit size (square meters)			0.0003*	0.0002*	0.0002*
			(0.0001)	(0.0001)	(0.0001)
Floor area ratio			-0.013**	-0.012***	-0.012***
			(0.005)	(0.004)	(0.004)
Green plot ratio			0.373***	0.310***	0.310***
			(0.074)	(0.084)	(0.084)
1{interior construction included}			0.051**	0.049**	0.049**
			(0.023)	(0.022)	(0.022)
Maintenance fee (yuan/month/m^2)			0.081***	0.074***	0.073***
			(0.011)	(0.008)	(0.008)
No. of parking lot per household				0.104***	0.104***
				(0.017)	(0.017)
1{central heat}				-0.017	-0.017
				(0.019)	(0.019)
Constant	8.817***	9.299***	8.801***	8.804***	8.813***
	(0.031)	(0.192)	(0.156)	(0.157)	(0.155)
Property FE	Yes	No	No	No	No
District-year FE	Yes	Yes	Yes	Yes	Yes
Observations	6,223	6,223	6,172	6,166	6,166
R-squared	0.95	0.78	0.87	0.88	0.88

Notes: The sample period is 2005 2011 (excludes 2008). It includes all properties within 50 kilometers of the city center. Congestion index = (Annual delay hours)/(Annual travel hours at the speed limit). 1{dist to existing metro $<$ 1km} equals 1 if the property is located with one kilometer of the nearest existing metro station. Please refer to Table 2.2 for details on the construction and definition of other controls. Coefficients are reported with robust standard errors in parentheses. Standard errors are clustered on the property level. Time-varying coefficients are not reported due to space. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels.

Table 2.5: The Effect of Commute Hours on Housing Prices

	Dependent Variable: Log (price)				
	(1)	(2)	(3)	(4)	(5)
Annual commute hours	-0.001**	-0.004***	-0.003**	-0.002**	-0.001**
	(0.0004)	(0.001)	(0.001)	(0.001)	(0.0005)
1{dist to existing metro<1km}				0.038**	
				(0.016)	
1{dist to existing metro<1km}* Annual commute hours					0.0002***
					(0.0001)
No. of bus lines				0.001	0.001
				(0.001)	(0.001)
Dist to City Center (km)		-0.005	-0.008	-0.009	-0.009
		(0.008)	(0.005)	(0.005)	(0.006)
Dist to nearest park (km)				-0.002	-0.002
				(0.003)	(0.003)
Dist to nearest core school (km)				-0.003	-0.003
				(0.003)	(0.002)
Unit size (square meters)			0.0003*	0.0002*	0.0002*
			(0.0001)	(0.0001)	(0.0001)
Floor area ratio			-0.016**	-0.014***	-0.014***
			(0.006)	(0.005)	(0.005)
Green plot ratio			0.385***	0.320***	0.320***
			(0.078)	(0.085)	(0.084)
1{interior construction included}			0.044*	0.044**	0.045**
			(0.023)	(0.021)	(0.022)
Maintenance fee (yuan/month/m^2)			0.080***	0.073***	0.073***
			(0.010)	(0.008)	(0.008)
No. of parking lot per household				0.098***	0.098***
				(0.015)	(0.015)
1{central heat}				-0.021	-0.020
				(0.020)	(0.020)
Constant	8.939***	9.348***	8.893***	8.880***	8.888***
	(0.047)	(0.053)	(0.091)	(0.113)	(0.112)
Property FE	Yes	No	No	No	No
District-year FE	Yes	Yes	Yes	Yes	Yes
Observations	6,228	6,228	6,177	6,171	6,166
R-squared	0.95	0.79	0.88	0.89	0.88

Notes: The sample period is 2005 2011 (excludes 2008). It includes all properties within 50 kilometers of the city center. Annual commute hours = Daily morning travel hours 500. 1{dist to existing metro < 1km} equals 1 if the property is located with one kilometer of the nearest existing metro station. Please refer to Table 2.2 for details on the construction and definition of other controls. Coefficients are reported with robust standard errors in parentheses. Standard errors are clustered on the property level. Time-varying coefficients are not reported due to space. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels.

Table 2.6: The Effect of Delay Time on Housing Prices – Instrumental Variable

	Dependent Variable: Log (price)			
	(1)	(2)	(3)	(4)
Annual delay hours	-0.001***	-0.0017***	-0.0014**	-0.0014**
	(0.0003)	(0.0007)	(0.0007)	(0.0007)
1{dist to existing metro<1km}				0.039***
				(0.008)
No. of bus lines				0.001
				(0.001)
Dist to City Center (km)		-0.006	-0.008*	-0.011**
		(0.006)	(0.005)	(0.006)
Dist to nearest park (km)				-0.002
				(0.002)
Dist to nearest core school (km)				-0.004***
				(0.001)
Unit size (square meters)			0.0003***	0.0002***
			(5.5e-05)	(4.9e-05)
Floor area ratio			-0.018***	-0.015***
			(0.003)	(0.003)
Green plot ratio			0.390***	0.327***
			(0.049)	(0.050)
1{interior construction included}			0.040***	0.043***
			(0.010)	(0.009)
Maintenance fee (yuan/month/m^2)			0.078***	0.072***
			(0.003)	(0.003)
No. of parking lot per household				0.094***
				(0.015)
1{central heat}				-0.023***
				(0.009)
Constant	9.422***	9.454***	8.994***	8.955***
	(0.121)	(0.099)	(0.119)	(0.125)
Property FE	Yes	No	No	No
District-year FE	Yes	Yes	Yes	Yes
First-stage F-Statistics	21.0	23.5	16.9	17.6
Observations	6,223	6,223	6,172	6,166
R-squared	0.79	0.79	0.88	0.89

Notes: The sample period is 2005 2011 (excludes 2008). It includes all properties within 50 kilometers of the city center. Annual delay hours = (Daily morning travel hours One-way travel hours at the speed limit) × 500. 1{dist to existing metro < 1km} equals 1 if the property is located with one kilometer of the nearest existing metro station. Please refer to Table 2.2 for details on the construction and definition of other controls. Coefficients are reported with robust standard errors in parentheses. Standard errors are clustered on the property level. Time-varying coefficients are not reported due to space. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels.

Table 2.7: The Effect of Construction Announcement on Housing Prices

	Dependent Variable: Log (price)			
	(1)	(2)	(3)	(4)
1{along suburban line}*Post	0.033***	0.036***	0.031***	0.032**
	(0.010)	(0.009)	(0.010)	(0.013)
1{along suburban line}		0.079**	0.055	0.053
		(0.034)	(0.040)	(0.045)
Post	-0.051***	-0.049***	-0.050***	-0.061***
	(0.005)	(0.005)	(0.005)	(0.004)
Dist to City Center (km)		-0.019***	-0.016**	-0.018***
		(0.005)	(0.004)	(0.003)
Location and Accessibility				
Annual delay hours				-0.0005**
				(0.0002)
1{dist to existing metro<1km}				0.069***
				(0.015)
No. of bus lines				0.002
				(0.002)
Dist to nearest park (km)				-0.003
				(0.003)
Dist to nearest core school (km)				-0.004
				(0.003)
Property Attributes				
Unit size (square meters)			0.0004*	0.0003
			(0.0002)	(0.0002)
Floor area ratio			-0.014***	-0.017***
			(0.004)	(0.004)
Green plot ratio			0.250**	0.233**
			(0.112)	(0.098)
1{interior construction included}			0.041	0.032
			(0.030)	(0.027)
Maintenance fee (yuan/month/m^2)			0.081***	0.073***
			(0.016)	(0.013)
No. of parking lot per household				0.086***
				(0.021)
1{central heat}				-0.011
				(0.029)
Constant	9.447***	9.805***	9.409***	9.508***
	(0.002)	(0.101)	(0.104)	(0.115)
Property FE	Yes	No	No	No
District-quarter FE	Yes	Yes	Yes	Yes
Observations	12,188	12,188	12,100	11,990
R-squared	0.97	0.57	0.76	0.77

Notes: This table reports OLS estimates of equation (2.11). The sample period is May 2008 - May 2009. 1{along suburban line} equals 1 if the property is located with one kilometer of the nearest suburban metro station, or it is in a suburban commercial hub. *Post* equals 1 for months after the announcement month November 2008. Please refer to Table 2.2 for details on the construction and definition of other controls. Coefficients are reported with robust standard errors in parentheses. Standard errors are clustered on the property level. Time-varying coefficients are not reported due to space. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels.

Table 2.8: Effects of the Metro Plan, Announcement and Completion on Housing Prices

	Dependent Variable: Log (price)			
	(1)	(2)	(3)	(4)
1{along suburban line}*Post	0.113***	0.112***	0.109***	0.104***
	(0.024)	(0.025)	(0.024)	(0.017)
1{along planned line}*Post_Plan	-0.040	-0.047*	-0.044	-0.045
	(0.030)	(0.028)	(0.033)	(0.028)
1{along suburban line}*Post_Constr	0.048**	0.044**	0.044*	0.041*
	(0.024)	(0.022)	(0.024)	(0.025)
1{along suburban line}		0.045*	0.096***	0.098***
		(0.023)	(0.024)	(0.024)
1{along planned line}		0.073	0.035	0.013
		(0.058)	(0.059)	(0.069)
Post	0.299***	0.302***	0.301***	0.249***
	(0.012)	(0.011)	(0.011)	(0.021)
Post_Plan	0.169***	0.168***	0.169***	0.173***
	(0.007)	(0.008)	(0.007)	(0.008)
Post_Constr	0.398***	0.399***	0.398***	0.385***
	(0.008)	(0.008)	(0.009)	(0.010)
Dist to City Center (km)		-0.018***	-0.015***	-0.012***
		(0.005)	(0.005)	(0.003)
Location and Accessibility	No	No	No	Yes
Property Attributes	No	No	Yes	Yes
Property FE	Yes	No	No	No
District-quarter FE	Yes	Yes	Yes	Yes
Observations	56,240	56,240	55,731	55,128
R-squared	0.84	0.61	0.71	0.73

Notes: The sample period is January 2007 - June 2011. 1{along suburban line} equals 1 if the property is located with one kilometer of the nearest suburban metro station, or it is in a suburban commercial hub. 1{along planned line} equals 1 if the property is located with one kilometer of the nearest planned suburban station, or it is in a suburban commercial hub. *Post* equals 1 for months after the announcement month November 2008. *Post_Plan* equals 1 for months after June 2007. *Post_Constr* equals 1 for months after December 2010.Please refer to Table 2.2 for details on the construction and definition of other controls. Coefficients are reported with robust standard errors in parentheses. Standard errors are clustered on the property level. Time-varying coefficients are not reported due to space. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels.

Table 2.9: Heterogeneous Effects of Construction Announcement on Housing Prices

Panel A: Properties that had greater than median congestion index (2007)

	Dependent Variable: Log (price)			
	(1)	(2)	(3)	(4)
1{along suburban line}*Post	0.037***	0.044***	0.040***	0.036**
	(0.010)	(0.009)	(0.011)	(0.015)
1{along suburban line}		0.024	0.013	0.012
		(0.022)	(0.027)	(0.0359)
Post	-0.048***	-0.048***	-0.053***	-0.061***
	(0.009)	(0.008)	(0.009)	(0.008)
Dist to City Center (km)		-0.021***	-0.017***	-0.015***
		(0.006)	(0.004)	(0.004)
Location and Accessibility	No	No	No	Yes
Property Attributes	No	No	Yes	Yes
Property FE	Yes	No	No	No
District-quarter FE	Yes	Yes	Yes	Yes
Observations	5,743	5,743	5,703	5,691
R-squared	0.97	0.54	0.76	0.77

Panel B: Properties that had less than median congestion index (2007)

	Dependent Variable: Log (price)			
	(1)	(2)	(3)	(4)
1{along suburban line}*Post	0.009	0.008	0.006	0.004
	(0.028)	(0.027)	(0.027)	(0.028)
1{along suburban line}		0.461**	0.278***	0.265***
		(0.207)	(0.0831)	(0.0874)
Post	-0.046***	-0.045***	-0.044***	-0.051***
	(0.004)	(0.004)	(0.004)	(0.004)
Dist to City Center (km)		-0.027***	-0.024***	-0.025***
		(0.007)	(0.004)	(0.003)
Location and Accessibility	No	No	No	Yes
Property Attributes	No	No	Yes	Yes
Property FE	Yes	No	No	No
District-quarter FE	Yes	Yes	Yes	Yes
Observations	6,445	6,445	6,397	6,299
R-squared	0.96	0.64	0.77	0.78

Notes: This table reports OLS estimates of equation (2.11). The sample period is May 2008 - May 2009. 1{along suburban line} equals 1 if the property is located with one kilometer of the nearest suburban metro station, or it is in a suburban commercial hub. *Post* equals 1 for months after the announcement month November 2008. Please refer to Table 2.2 for details on the construction and definition of other controls. Coefficients are reported with robust standard errors in parentheses. Standard errors are clustered on the property level. Time-varying coefficients are not reported due to space. *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels.

2.9 References

Agostini, C. and G. Palmucci (2008), The Anticipated Capitalization Effect of a New Metro Line on Housing Prices, *Fiscal Studies* 29 (2), 233-256.

Ahlfeldt, G., S. Redding, D. Sturm and N. Wolf (2012), The Economics of Density: Evidence from the Berlin Wall, CEP Discussion Paper No 1154.

Alonso, W. (1964), Location and Land Use Cambridge, Harvard University Press.

Anas, A. and L. Moses (1979), Mode Choice, Transport Structure and Urban Land Use, *Journal of Urban Economics* 6(2), 228-246.

Arnott, R. and J. MacKinnon, (1978), Market and Shadow Land Rents with Congestion, *American Economic Review* 58, 588-600.

Arnott, R., A. de Palma and R. Lindsey, (1993), A Structural Model of Peak-Period Congestion: A Traffic Bottleneck with Elastic Demand, *American Economic Review* 83, 161-179.

Atack, J., F. Bateman, M. Haines and R. Margo (2010), Did Railroads Induce or Follow Economic Growth? Urbanization and Population Growth in the American Midwest, 1850-1860, *Social Science History* 34(2), 171-97.

Baum-Snow, N. and M. Kahn (2005), Effects of Urban Rail Transit Expansions: Evidence from Sixteen Cities, *Brookings-Wharton Papers on Urban Affairs*, 88-135.

Baum-Snow, N. (2007), Suburbanization and Transportation in the Monocentric Model, *Journal of Urban Economics* 62(3), 405-423.

Brounen, D., P. Neuteboom and J. Xu (2010), Traffic Jamming House Prices, mimeo, Tilburg University.

Chinese Academy of Sciences (2012), The 2012 China Urbanization Report, November 2012.

Coulson, E.and R. Engle (1987), Transportation Costs and the Rent Gradient, *Journal of Urban Economics* 21, 287-297.

Damm, D., S. Lerman, E. Lerner-Lam, and J. Young (1980), Response of Urban Real Estate Values in Anticipation of the Washington Metro, *Journal of Transport Economics and Policy* 14(3), 315.

Davis, W., S. Diegel and R. Boundy (2011), Transportation Energy Data Book: Edition 30, Office of Energy Efficiency and Renewable Energy, U.S. Department of Energy.

Donaldson, D. (2010), Railroads of the Raj: Estimating the Impact of Transportation Infrastructure, mimeo, Massachusetts Institute of Technology.

Duranton, G. and M. Turner (2011), The Fundamental Law of Road Congestion: Evidence from US cities, *American Economic Review* 101(6), 2616-2652.

Foton Motor Co., Ltd. (2010), The 2009 Foton Motor Report, January 2010.

Fujita, M. (1989), Urban Economic Theory: Land Use and City Size. Cambridge University Press, Cambridge.

Gatzlaff, D., and M. Smith (1993), The Impact of the Miami Metrorail on the Value of Residences near Station Locations, *Land Economics* 69(1), 5466.

Glaeser, E. and M. Kahn (2001), Decentralized Employment and the Transformation of the American City, *Brookings-Wharton Papers on Urban Affairs*, 1-63.

Gordon, P., A. Kumar and H. Richardson (1989), The Influence of Metropolitan Spatial Structure on Commuting Time, *Journal of Urban Economics* 26, 138-151.

Hartgen, D. and M. G. Fields (2009), Gridlock and Growth: The Effect of Traffic Congestion on Regional Economic Performance, mimeo, Reason Foundation.

He, J. (2005), The Capitalization Effect of Beijing's No. 13 Subway Line: A Hedonic Study, mimeo, Tsinghua University.

Leape, J. (2006), The London Congestion Charge, *Journal of Economic Perspectives* 20(4), 157-176.

Malalur, P. (2011), Traffic Delay Prediction from Historical Observations, mimeo, Massachusetts Institute of Technology.

McDonald, J., and C. Osuji (1995), The Effect of Anticipated Transportation Improvement on Residential Land Values, *Regional Science and Urban Economics* 25(3), 26178.

McMillen, D., and J. McDonald (2004), Reaction of House Prices to a New Rapid Transit Line: Chicago's Midway Line, 19831999, *Real Estate Economics* 32(3), 46386.

Michaels, G. (2008), The Effect of Trade on the Demand for Skill – Evidence from the Interstate Highway System, *Review of Economics and Statistics* 90(4), 683-701.

Mills, E. S. (1967), An Aggregative Model of Resource Allocation in a Metropolitan Area, *American Economic Review Papers and Proceedings* 57(2), 197-210.

Mun, S. and M. Yonekawa (2006), Flextime, Traffic Congestion and Urban Productivity, *Journal of Transport Economics and Policy* 40(3), 329-358.

Muth, R. F. (1969), Cities and Housing: the Spatial Pattern of Urban Residential Land Use, University of Chicago Press, Chicago.

Qian, N., A. Banerjee and E. Duflo (2012), On the Road: Transportation Infrastructure and Economic Development, mimeo, Yale University.

Rosen, S. (1974), Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition, *Journal of Political Economy* 82(1), 34-55.

Schrank, D., T. Lomax and B. Eisele (2011), TTIs 2011 Urban Mobility Report, Texas Transportation Institute, September 2011.

Smersh, G. and M. Smith (2000), Accessibility Changes and Urban House Price Appreciation: A Constrained Optimization Approach to Determining Distance Effects, *Journal of Housing*

Economics 9, 187-196.

Staiger, D. and J. Stock (1997), Instrumental Variables Regression with Weak Instruments, *Econometrica* 65(3), 557-586.

Weisbrod, G., D. Vary and G. Treyz (2007), Measuring Economic Costs of Urban Traffic Congestion to Business, *Journal of the Transportation Research Board*, 98-106.

Vessali, K. V. (1996), Land Use Impacts of Rapid Transit: A Review of Empirical Literature, *Berkeley Planning Journal* 11, 71-105.

Vickrey, W. S. (1969), Congestion Theory and Transport Investment, *American Economic Review* 59(2), 251-260.

Wheaton, W. (2004), Commuting, Congestion and Employment Dispersal in Cities with Mixed Land Use, *Journal of Urban Economics* 55(3), 417-439.

Winston, C. and A. Langer (2006), The Effect of Government Highway Spending on Road Users, *Journal of Urban Economics* 60(3), 463-483.

Winston, C. and V. Maheshri (2007), On the Social Desirability of Urban Rail Transit Systems, *Journal of Urban Economics* 62(2), 362-382.

Yi, D., K. Chen and G. Chang (2008), The Dilemma of Migrant Workers – The Case of Beijing, *Chinese Cadres Tribune* 9.

Zheng, S., Y. Fu and H. Liu (2006), Housing-Choice Hindrances and Urban Spatial Structure: Evidence from Matched Location and Location-Preference Data in Chinese Cities, *Journal of Urban Economics* 60 (3), 535-557.

Zheng, S. and M. Kahn (2008), Land and Residential Property Markets in a Booming Economy: New Evidence from Beijing, *Journal of Urban Economics* 63(2), 743-757.

Zhang, Y. and H. Shing (2006), The London Congestion Charge And Property Prices: An Evaluation Of The Impact On Property Prices Inside And Outside The Zone, MPRA Paper

4050, University Library of Munich, Germany.

A. Theoretical Appendix

Proof of Proposition 1

First, consider the case in which rational investors believe that the PA is conflicted, and therefore do not buy its report. Each of these investors votes for management with probability $\frac{1}{2}$. Litigation-averse investors would like to buy the report given the informational gain is larger than price of the report. With probability β the proposal passes. β is as defined in footnote 17.

When the PA receives signal $s = o$ and reports $m = A$, its profit is

$$\begin{aligned}\pi(A|o) &= \beta(\phi + f\alpha - ep\rho\alpha) + (1-\beta)(\phi + f\alpha) \\ &= \phi + f\alpha - \beta ep\rho\alpha\end{aligned}$$

If the proposal is approved, with probability ep a loss occurs (from PA's point of view) and the PA incurs a reputation cost from litigation-averse investors. With probability $1-\beta$ the proposal stalls, and the PA therefore is not liable. When the PA receives signal $s = o$ and reports $m = O$, it loses its corporate client. Its profit is

$$\pi(O|o) = f\alpha$$

The PA always reports $m = A$ when $\pi(A|o) > \pi(O|o)$ which yields the condition $\phi > \beta ep\rho\alpha$. The PA reports truthfully when $\phi < \beta ep\rho\alpha$. Note that when receiving signal $s = a$ the PA will not be conflicted because recommending against management in this case does not yield additional profit.

Now we pin down fee conditions. With probability $\frac{1}{2}$, the PA receives an "o" signal. When $\phi > \beta ep\rho\alpha$, it reports $m = A$ with probability 1. With probability β the proposal passes and leads to a loss with probability ep. So with probability $\frac{1}{2}\beta ep$ litigation-averse investors suffer a loss of c from clients. The expected total cost of purchasing the report is $f + \frac{1}{2}\beta epc$. Similarly, if a litigation-averse investor deviates to vote based on ex ante belief, with probability $\frac{1}{2}\frac{1}{2}$ it votes for an "o" proposal. The cost of making a wrong vote therefore is $\frac{1}{2}\frac{1}{2}\beta pC$. The profit maximizing PA charges a fee $f = \frac{1}{2}\frac{1}{2}\beta pC - \frac{1}{2}\beta epc$. Since we normalize $c = 0$, $f = \frac{1}{2}\frac{1}{2}\beta pC$. Realizing this, a rational investor does not want to deviate to buy the report. The expected total cost of buying the report is $f + \frac{1}{2}\beta ep\tilde{C}$, which is greater than the cost of voting based on ex ante belief $\frac{1}{2}\frac{1}{2}\beta p\tilde{C}$. Note that $e > \frac{1}{2}$.

Second, when $\phi < \beta ep\rho\alpha$, assume there is an equilirbium in which the PA reports truthfully and only litigation-averse investors buy the report. Litigation-averse investors understand that they will not incur a cost following the PA since the PA is truthful. When $s = o, m = O$, the expected cost is $\frac{1}{2}\gamma epC$ if a litigation-averse investor votes based on ex ante belief. When $s = a, m = A$, the expected cost of voting based on ex ante belief is $\frac{1}{2}\beta(1-e)pC$. Thus the PA charges a fee $f = \frac{1}{2}[\frac{1}{2}\gamma epC + \frac{1}{2}\beta(1-e)pC]$. The corresponding fee for rational investors is $f = \frac{1}{2}[\frac{1}{2}\gamma ep\tilde{C} + \frac{1}{2}\beta(1-e)p\tilde{C}]$. In the general case in which C and $\tilde{C}$ are close enough, the PA will lower its fee to woo both types of investors in order to boost its profit. This therefore can not be an equilibrium.

Finally, we show that there is another equilibrium in which both types of investors buy PA's report. Now If the PA receives signal $s = o$ and reports $m = A$, its profit is

$$\pi(A|o) = \phi + f - ep\rho$$

If the PA receives signal $s = o$ and reports $m = O$, it loses the corporate client. Its profit is

$$\pi(O|o) = f$$

So the PA reports truthfully when $\pi(A|o) < \pi(O|o)$ which yields the condition $\phi < ep\rho$. Now consider the fee condition. If a rational investor deviates and does not buy the PA's report, it will vote for management with probability $\frac{1}{2}$. Since all other investors vote with the PA, the voting result will not change. The cost of deviating is $\frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}$. Similarly, if a litigation-averse investor deviates, the cost is $\frac{1}{2}\frac{1}{2}(1-e)pC$. So PA sets $f = \min[\frac{1}{2}\frac{1}{2}(1-e)pC, \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}]$.

Proof of Lemma 1

First consider the case where $\alpha < \frac{1}{2}$. Rational investors are the majority. They will determine vote outcomes if they follow PA 2. For a rational investor, voting based on ex ante belief is costly only when PA 2's signal s_2 is "a" because the proposal passes. This happens with probability $\frac{1}{2}$, and the investor votes for management with probability $\frac{1}{2}$. When $s_2 = a$, the probability that a loss occurs is $(1-e)p$, so cost for the rational investor is $\frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}$. If PA 2 only retains rational investors, it sets fee $f_2 = \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}$.

Now PA 2 wants to lower its fee to attract litigation-averse investors as well. Let the lowest fee PA 2 is willing to charge be $\tilde{f}_2$. This fee makes PA 2 indifferent towards serving both types of investors or just rational investors. It leads to the following relation

$$\tilde{f}_2 \cdot N = f_2 \cdot (1-\alpha)N$$

Rearranging, we obtain $\tilde{f}_2 = (1-\alpha)\frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}$. However, PA 1 has an incentive to charge a fee $f_1 = \tilde{f}_2 - \epsilon$ where ϵ is a small number. At this price, litigation-averse investors find it less expensive to purchase PA 1's report. Recall that for litigation-averse investors, following PA 1 does not incur a cost imposed by clients (c is normalized to 0). For these investors, the cost of voting based on ex ante belief is $\frac{1}{2}\frac{1}{2}(1-e)pC$ which will be larger than f_1, in the general case when C and $\tilde{C}$ are close. Therefore litigation-averse investors will switch to PA 1.

Anticipating this, PA 2 raises its fee to $f_2 = \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}$ to serve only rational investors. In turn, PA 1 sets its fee $f_1 = \min[\frac{1}{2}\frac{1}{2}(1-e)pC, \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}]$ to serve litigation-averse investors.

The case where $\alpha > \frac{1}{2}$ is similar. Fees charged by both PAs are derived in Proposition 4. We therefore have shown that there exists no equilibrium in which both types of investors buy a report from PA 2.

Proof of Proposition 3

This is the case where $\alpha < \frac{1}{2}$. Rational investors will determine the vote outcome if they follow PA 2. I show there exists an equilibrium in which litigation-averse investors buy PA 1's report and rational investors purchase PA 2's report.

When $s_1 = o, m_1 = A, s_2 = a, m_2 = A$, all investors vote for the management, and the proposal therefore passes. PA 1 does not incur a reputation cost when a loss occurs because PAs' recommendations are the same. It is difficult to determine which PA is conflicted. The profit for PA 1 is

$$\pi_1(A|o) = \phi + f_1\alpha$$

When $s_1 = o, m_1 = A, s_2 = o, m_2 = O$, litigation-averse investors vote for the proposal while rational investors vote against it. Since there are more rational ones ($\alpha < \frac{1}{2}$), the proposal fails to pass. Again PA 1 does not suffer a reputation cost. The profit for PA 1 is

$$\pi_1(A|o) = \phi + f_1\alpha$$

When $s_1 = o, m_1 = O$, no matter what m_2 is, PA 1 does not incur a reputation cost because it is truthful. The profit for PA 1 is

$$\pi_1(O|o) = f_1\alpha$$

Since $\pi_1(A|o) > \pi_1(O|o)$, PA 1 always reports $m_1 = A$.

Now we pin down the fee conditions. As shown in the proof of Lemma 1, PA 2 raises its fee to $f_2 = \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}$ to serve only the rational investors. PA 1 sets its fee $f_1 = \min[\frac{1}{2}\frac{1}{2}(1-e)pC, \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}]$ to serve the litigation-averse investors. Note that litigation-averse investors will not deviate to vote based on their ex ante belief because the cost of doing so is $\frac{1}{2}\frac{1}{2}(1-e)pC$, which is larger or equal to f_1.

Proof of Proposition 4

This is the case where $\alpha > \frac{1}{2}$. Litigation-averse investors will determine the vote outcome. I show there exists an equilibrium where litigation-averse investors buy PA 1's report and rational investors purchase PA 2's report.

When $s_1 = o, m_1 = A, s_2 = a, m_2 = A$, all investors vote for management, and the proposal therefore passes. Again, PA 1 does not suffer a reputation cost when a loss occurs because PAs' recommendations are the same. The profit for PA 1 is

$$\pi_1(A|o) = \phi + f_1\alpha$$

When $s_1 = o, m_1 = A, s_2 = o, m_2 = O$, litigation-averse investors vote for the proposal while rational investors vote against it. Since there are more litigation-averse ones ($\alpha > \frac{1}{2}$), the proposal is approved. The probability that the state is "o" equals $\frac{e^2}{e^2+(1-e)^2}$. For simplicity, define this probability as q. Now PA 1 suffers a reputation cost. The profit for PA 1 is

$$\pi_1(A|o) = \phi + f_1\alpha - q\tilde{p}\rho_1\alpha$$

The expected profit for PA 1 is the weighted average profits under these two scenarios:

$$\mathrm{E}[\pi_1(A|o)] = \phi + f_1\alpha - \frac{1}{2}q\tilde{p}\rho_1\alpha$$

When $s_1 = o, m_1 = O$, no matter what m_2 is, PA 1 does not suffer a reputation loss because it is truthful. The profit for PA 1 is

$$\pi_1(O|o) = f_1\alpha$$

So the PA always reports $m_1 = A$ when $\mathrm{E}[\pi_1(A|o)] > \pi_1(O|o)$ which yields the condition $\phi > \frac{1}{2}q\tilde{p}\rho_1\alpha$.

Now we pin down the fee conditions. Since PA always recommends $m_1 = A$, the proposal will be approved for sure. If a rational investor deviates and votes for the management based on ex ante belief, then with probability $\frac{1}{2}\frac{1}{2}$ it votes for an "o" proposal. The cost of making a wrong vote is $\frac{1}{2}\frac{1}{2}p\tilde{C}$. PA 2 therefore sets fee $f_2 = \frac{1}{2}\frac{1}{2}p\tilde{C}$ to serve rational investors. PA 1 sets its fee $f_1 = \min[\frac{1}{2}\frac{1}{2}pC, \frac{1}{2}\frac{1}{2}p\tilde{C}]$ to serve the litigation-averse investors.

Next I show that when $\phi < \frac{1}{2}q\tilde{p}\rho_1\alpha$, the fee conditions are identical to those in Proposition 3. PA 2 sets the fee to $f_2 = \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}$ to serve only the rational investors. PA 1 sets its fee $f_1 = \min[\frac{1}{2}\frac{1}{2}(1-e)pC, \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}]$ to serve the litigation-averse investors.

Other Equilibria Under Competition

Recall that if litigation-averse investors obtain PA 1 and follow its voting recommendation, the litigation cost is $c < C$ when the PA is found out to be conflicted. Note that we normalize

c to zero in the main text. Here I do not make this assumption. Litigation-averse investors now may have an incentive to purchase an additional report from PA 2. The reason is that if an institutional investor buys reports from both PAs, it may be more difficult for its clients to win a lawsuit when a loss occurs. Because the investor could argue that she acted in good faith by subscribing to both reports in order to make the best judgment. In practice, some institutional investors subscribe to multiple PAs in order to make a more informed vote.

Note that in the case of conflicting reports, litigation-averse investors are assumed to follow PA 1's recommendation. This is because PA 1 has been in business longer, and is regarded as the more established advisor. It is natural for litigation-averse investors facing conflicting recommendations to follow its recommendation. For simplicity, I assume that litigation-averse investors are immune from clients' lawsuits when obtaining both PAs. I proceed to derive the following symmetric equilibrium under each informational regime:

Proposition 5 *When $\alpha < \frac{1}{2}$, the equilibrium of the subgame is:*

1. *PA 1 always reports $m_1 = A$, and sets $f_1 = \frac{1}{2}\frac{1}{2}(1-e)pC - \frac{1}{2}\frac{1}{2}\frac{1}{2}pc$. Litigation-averse investors buy both reports and rational investors buy only PA 2's report. PA 2 reports truthfully and sets $f_2 = \min[\frac{1}{2}\frac{1}{2}\frac{1}{2}pc, \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}]$.*

Proof. This is the case where $\alpha < \frac{1}{2}$. Rational investors will determine the vote outcome if they follow PA 2. I show that there exists an equilibrium where litigation-averse investors buy both reports and rational investors buy only PA 2's report. The proof is identical to the proof of Proposition 3 except the fee conditions. PA 1 always reports $m_1 = A$.

We now pin down the fee conditions. For a litigation-averse investor, voting based on ex ante belief is costly only when $s_2 = a$ because the proposal passes. This happens with probability $\frac{1}{2}$, and the investor votes for the management with probability $\frac{1}{2}$. When $s_2 = a$, the probability that a loss occurs is $(1-e)p$, so the expected cost to the litigation-averse investor is $\frac{1}{2}\frac{1}{2}(1-e)pC$.

When the litigation-averse investor only buys a report from PA 1, she incurs a cost when $s_1 = o, s_2 = a$, and the probability that the state is "o" equals $\frac{e(1-e)}{e(1-e)+(1-e)e} = \frac{1}{2}$. So the total cost is $f_1 + \frac{1}{2}\frac{1}{2}\frac{1}{2}pc$.

When the litigation-averse investor buys both reports, she is immune from any costs imposed by her clients, and total cost is $f_1 + f_2$. Then PA 1 will charge a fee $f_1 = \frac{1}{2}\frac{1}{2}(1-e)pC - \frac{1}{2}\frac{1}{2}\frac{1}{2}pc$. PA 2 sets $f_2 \leq \frac{1}{2}\frac{1}{2}\frac{1}{2}pc$.

For a rational investor, voting based on ex ante belief is costly only when $s_2 = a$ because the proposal passes. The cost is $\frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}$. Following PA 2 only leads to total cost f_2. So PA 2 sets its fee $f_2 = \min[\frac{1}{2}\frac{1}{2}\frac{1}{2}pc, \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}]$.

It is optimal for litigation-averse investors to purchase both reports because this minimizes the cost. Rational investors will buy only PA 2's report because PA 1 is conflicted.

Proposition 6 *When $\alpha > \frac{1}{2}$, the equilibrium of the subgame is:*

1. *If $\phi > \frac{1}{2}q\tilde{p}\rho_1\alpha$, PA 1 always reports $m_1 = A$, and sets $f_1 = \frac{1}{2}\frac{1}{2}pC - \frac{1}{2}\frac{1}{2}pc$. Litigation-averse investors buy both reports and rational investors buy only PA 2's report. PA 2 reports truthfully and sets $f_2 = \min[\frac{1}{2}\frac{1}{2}pc, \frac{1}{2}\frac{1}{2}p\tilde{C}]$.*

2. *If $\phi < \frac{1}{2}q\tilde{p}\rho_1\alpha$, PA 1 reports truthfully, and sets $f_1 = \frac{1}{2}\frac{1}{2}(1-e)pC$. Litigation-averse investors purchase PA 1's report, and rational investors buy PA 2's report. PA 2 is truthful and sets $f_2 = \min[\frac{1}{2}\frac{1}{2}(1-e)pC, \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}]$.*

Proof This is the case where $\alpha > \frac{1}{2}$. Litigation-averse investors will determine the vote outcome. I show there exists an equilibrium where PA 1 always reports $m_1 = A$, and litigation-averse investors buy both reports and rational investors buy only PA 2's report. The proof is identical to the proof of Proposition 4 except the fee conditions.

Now we pin down the fee conditions. When PA 1 always recommends $m_1 = A$, the proposal will be approved for sure. If a litigation-averse investor deviates and votes for the management based on ex ante belief, then with probability $\frac{1}{2}\frac{1}{2}$ it votes for an "o" proposal. The cost of making a wrong vote is $\frac{1}{2}\frac{1}{2}pC$. If a litigation-averse investor buys only a report from PA 1, the total cost is $f_1 + \frac{1}{2}\frac{1}{2}pc$. The total cost is $f_1 + f_2$ if she purchases both reports. Therefore PA 1 sets $f_1 = \frac{1}{2}\frac{1}{2}pC - \frac{1}{2}\frac{1}{2}pc$ and PA 2 sets $f_2 \leq \frac{1}{2}\frac{1}{2}pc$.

Since rational investors expect that PA 1 is conflicted, they purchase only PA 2's report. The cost of deviating to vote based on ex ante belief is $\frac{1}{2}\frac{1}{2}p\tilde{C}$. PA 2 sets $f_2 = \min[\frac{1}{2}\frac{1}{2}pc, \frac{1}{2}\frac{1}{2}p\tilde{C}]$.

Next I derive the fee conditions when $\phi < \frac{1}{2}q\tilde{p}\rho_1\alpha$. PA 1 will be truthful in this case. For a litigation-averse investor, deviating to vote based on ex ante belief is costly when $s_1 = a$. The associated cost is $\frac{1}{2}\frac{1}{2}(1-e)pC$. Since now PA 1 is truthful, litigation-averse investors purchase only its report. PA 1 sets $f_1 = \frac{1}{2}\frac{1}{2}(1-e)pC$. Similarly, the cost of deviating for rational investors is $\frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}$. Therefore PA 2 sets $f_2 = \min[\frac{1}{2}\frac{1}{2}(1-e)pC, \frac{1}{2}\frac{1}{2}(1-e)p\tilde{C}]$. Since rational investors are indifferent as which report to buy, we assume they purchase PA 2's report.

B. Empirical Appendix

Table B.1: Robustness Checks: Convergence in Recommendations

	Dependent Variable: fraction of ISS's "For" recommendations less fraction of Glass Lewis's "for" recommendations at firm level					
	(1)	(2)	(3)	(4)	(5)	(6)
GL's mkt share (# of clients)	-0.140***	-0.163***	-0.177***			
	(0.050)	(0.057)	(0.068)			
GL's mkt share (# firms covered)				-0.145***	-0.169***	-0.183***
				(0.052)	(0.059)	(0.070)
Log assets	-0.021**	-0.029***	-0.030**	-0.021**	-0.029***	-0.030**
	(0.008)	(0.011)	(0.015)	(0.008)	(0.011)	(0.015)
Ind-adj ROA	-0.022	-0.075*	-0.065	-0.022	-0.076*	-0.065
	(0.018)	(0.046)	(0.067)	(0.018)	(0.046)	(0.067)
Prior-year stock return	0.002	0.002	0.013	0.002	0.002	0.013
	(0.004)	(0.006)	(0.008)	(0.004)	(0.006)	(0.009)
Leverage			0.072*			0.072*
			(0.042)			(0.042)
Capex/Assets			0.213			0.213
			(0.136)			(0.136)
Abnormal exec. compensation ($m)		-0.0004	-0.001		-0.0004	-0.001
		(0.002)	(0.003)		(0.002)	(0.003)
ΔExecutive compensation YOY		-0.001***	-0.001***		-0.001***	-0.001***
		(0.0002)	(0.0003)		(0.0002)	(0.0003)
Cash/total compensation		0.025	0.017		0.025	0.017
		(0.022)	(0.028)		(0.022)	(0.028)
Classified board*poison pill			-0.016			-0.016
			(0.014)			(0.014)
Board size			0.001			0.001
			(0.003)			(0.003)
Ratio of independent directors			-0.043			-0.043
			(0.043)			(0.043)
Institutional holdings			-0.012			-0.012
			(0.046)			(0.046)
Management holdings			0.026			0.027
			(0.062)			(0.062)
Constant	0.269***	0.369***	0.386***	0.280***	0.382***	0.400***
	(0.060)	(0.088)	(0.132)	(0.061)	(0.089)	(0.132)
Firm FE	√	√	√	√	√	√
Fiscal-year FE	√	√	√	√	√	√
Industry-year trend	√	√	√	√	√	√
Observations	17,732	10,859	7,929	17,732	10,859	7,929
R-squared	0.39	0.33	0.33	0.39	0.33	0.33

Notes: The dependent variable is fraction of ISS's "For" recommendations minus fraction of Glass Lewis' "For" recommendations at the firm level. All independent variables are defined as in Table 1.1. Robust standard errors (clustered at the firm level) are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table B.2: Robustness Checks: Direction of Convergence in Recommendations

	Dependent Variable: fraction of differing recommendations at firm level (ISS "For," Glass Lewis "Against/Withhold")					
	(1)	(2)	(3)	(4)	(5)	(6)
GL's mkt share (# of clients)	-0.500***	-0.465***	-0.427***			
	(0.028)	(0.036)	(0.046)			
GL's mkt share (# firms covered)				-1.120***	-1.042***	-0.732***
				(0.063)	(0.081)	(0.103)
Log assets	-0.031***	-0.047***	-0.044***	-0.031***	-0.047***	-0.044***
	(0.008)	(0.011)	(0.013)	(0.008)	(0.011)	(0.013)
Ind-adj ROA	-0.027	-0.132***	-0.063	-0.027	-0.132***	-0.063
	(0.017)	(0.044)	(0.059)	(0.017)	(0.044)	(0.059)
Prior-year stock return	0.003	0.002	0.008	0.003	0.002	0.008
	(0.004)	(0.005)	(0.007)	(0.004)	(0.005)	(0.007)
Leverage			0.096***			0.096***
			(0.036)			(0.036)
Capex/Assets			0.093			0.093
			(0.120)			(0.120)
Abnormal exec. compensation ($m)		-0.0003	-0.001		-0.0003	-0.001
		(0.001)	(0.002)		(0.001)	(0.002)
ΔExecutive compensation YOY		-0.0001	-0.0001		-0.0001	-0.0001
		(0.0002)	(0.0002)		(0.0002)	(0.0002)
Cash/total compensation		0.016	0.018		0.016	0.018
		(0.019)	(0.024)		(0.019)	(0.024)
Classified board*poison pill			-0.016			-0.016
			(0.012)			(0.012)
Board size			0.001			0.002
			(0.003)			(0.003)
Ratio of independent directors			-0.063*			-0.063*
			(0.038)			(0.038)
Institutional holdings			-0.072			-0.072
			(0.049)			(0.049)
Management holdings			0.043			0.043
			(0.051)			(0.051)
Constant	0.532***	0.661***	0.632***	0.856***	0.963***	0.844***
	(0.055)	(0.086)	(0.114)	(0.053)	(0.083)	(0.108)
Firm FE	√	√	√	√	√	√
Fiscal-year FE	√	√	√	√	√	√
Industry-year trend	√	√	√	√	√	√
Observations	17,864	10,803	7,607	17,864	10,803	7,607
R-squared	0.48	0.46	0.35	0.48	0.46	0.35

Notes: The dependent variable is fraction of differing recommendations for management proposals at the firm level (ISS "For," Glass Lewis "Against/Withhold"). All independent variables are defined as in Table 1.1. Robust standard errors (clustered at the firm level) are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table B.3: Robustness Checks: Direction of Convergence in Recommendations

	Dependent Variable: fraction of differing recommendations at firm level (ISS "Against/Withhold," Glass Lewis "For")					
	(1)	(2)	(3)	(4)	(5)	(6)
GL's mkt share (# of clients)	-0.100***	-0.136***	-0.129***			
	(0.024)	(0.026)	(0.032)			
GL's mkt share (# firms covered)				-0.104***	-0.141***	-0.133***
				(0.024)	(0.027)	(0.033)
Log assets	-0.001	-0.004	-0.002	-0.001	-0.004	-0.002
	(0.004)	(0.005)	(0.008)	(0.004)	(0.005)	(0.008)
Ind-adj ROA	0.005	0.012	0.052	0.005	0.012	0.052
	(0.010)	(0.026)	(0.036)	(0.010)	(0.026)	(0.036)
Prior-year stock return	-0.003	-0.003	-0.009**	-0.003	-0.003	-0.009**
	(0.002)	(0.003)	(0.004)	(0.002)	(0.003)	(0.004)
Leverage			0.015			0.015
			(0.022)			(0.022)
Capex/Assets			-0.165**			-0.165**
			(0.073)			(0.073)
Abnormal exec. compensation ($m)		-0.001	-0.001		-0.001	-0.001
		(0.001)	(0.001)		(0.001)	(0.001)
ΔExecutive compensation YOY		0.001***	0.001**		0.001***	0.001**
		(0.0002)	(0.0002)		(0.0002)	(0.0002)
Cash/total compensation		-0.009	-0.011		-0.009	-0.011
		(0.011)	(0.014)		(0.011)	(0.014)
Classified board*poison pill			-0.007			-0.007
			(0.007)			(0.007)
Board size			0.001			0.001
			(0.002)			(0.002)
Ratio of independent directors			-0.030			-0.030
			(0.022)			(0.022)
Institutional holdings			0.003			0.003
			(0.024)			(0.024)
Management holdings			-0.004			-0.004
			(0.033)			(0.033)
Constant	0.075**	0.050	0.107*	0.075**	0.050	0.107*
	(0.030)	(0.042)	(0.062)	(0.030)	(0.042)	(0.062)
Firm FE	√	√	√	√	√	√
Fiscal-year FE	√	√	√	√	√	√
Industry-year trend	√	√	√	√	√	√
Observations	17,732	10,859	7,929	17,732	10,859	7,929
R-squared	0.46	0.37	0.37	0.46	0.37	0.37

Notes: The dependent variable is fraction of differing recommendations for management proposals at the firm level (ISS "Against/Withhold," Glass Lewis "For"). All independent variables are defined as in Table 1.1. Robust standard errors (clustered at the firm level) are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.